

T.C.
GÜMÜŞHANE ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

YAPAY ZEKA VE AKILLI SİSTEMLER ANA BİLİM DALI

FONOKARDİYOGRAM VERİLERİNDEN DERİN ÖĞRENME TABANLI
MEL-SPEKTROGRAM ANALİZİ İLE KALP ÜFÜRÜMÜ TESPİTİ

YÜKSEK LİSANS

Asmaa Abduh Mohammed Saeed ALTAIB

ŞUBAT-2026
GÜMÜŞHANE



T.C.

**GÜMÜŞHANE ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

YAPAY ZEKA VE AKILLI SİSTEMLER ANA BİLİM DALI

**FONOKARDİYOGRAM VERİLERİNDEN DERİN ÖĞRENME TABANLI
MEL-SPEKTROGRAM ANALİZİ İLE KALP ÜFÜRÜMÜ TESPİTİ**

**HEART MUMRMUR DETECTION WITH DEEP LEARNING BASED
MEL SPECTROGRAM ANALYSIS FROM PHONOCARDIOGRAM DATA**

YÜKSEK LİSANS

Asmaa Abduh Mohammed Saeed ALTAIB

ŞUBAT-2026

GÜMÜŞHANE



T.C.

**GÜMÜŞHANE ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

YAPAY ZEKA VE AKILLI SİSTEMLER ANA BİLİM DALI

**FONOKARDİYOGRAM VERİLERİNDEN DERİN ÖĞRENME TABANLI
MEL-SPEKTROGRAM ANALİZİ İLE KALP ÜFÜRÜMÜ TESPİTİ**

**HEART MUMRMUR DETECTION WITH DEEP LEARNING BASED
MEL SPECTROGRAM ANALYSIS FROM PHONOCARDIOGRAM DATA**

YÜKSEK LİSANS

Asmaa Abduh Mohammed Saeed ALTAIB

Danışman: Dr. Öğr. Üyesi Hayati TÜRE

**ŞUBAT-2026
GÜMÜŞHANE**

BİLİMSEL ETİĞE UYGUNLUK BEYANI

Yüksek Lisans Tezi olarak hazırlamış olduğum “**Fonokardiyogram Verilerinden Derin Öğrenme Tabanlı Mel-spektrogram Analizi ile Kalp Üfürümü Tespiti**” isimli bu tezimin, tamamen kendi çalışmam olduğunu, her alıntıya kaynak gösterdiğimi, alıntı yaptığım tüm çalışmaları kaynakçada belirttiğimi ve Gümüşhane Üniversitesi’nin lisanslı kullanıcısı olduğu intihal yazılım programı ile Lisansüstü Eğitim Enstitüsü’nün belirlediği kıstaslara uygun olarak raporladığımı taahhüt ederim. Tezimin kâğıt ve elektronik kopyalarının Gümüşhane Üniversitesi Lisansüstü Eğitim Enstitüsü arşivinde saklanmasına izin verdiğimi onaylarım.

Lisansüstü Eğitim ve Öğretim Yönetmeliği’nin ilgili maddeleri uyarınca gereğinin yapılmasını arz ederim.

04/02/2026

.....
Asmaa Abduh Mohammed Saeed ALTAIB

TEŞEKKÜR

Bu tez çalışmasının yürütülmesi ve metnin hazırlanması sürecinde , değerli bilgi ve deneyimleriyle bana rehberlik eden, her aşamada desteğini esirgemeyen danışmanım Dr. Öğr. Üyesi Hayati TÜRE'ye katkılarından dolayı içtenlikle teşekkür ederim.

Asmaa Abduh Mohammed Saeed ALTAIB

GÜMÜŞHANE – 2026



ÖZET

Kalp oskültasyonu, kalp seslerinin değerlendirilmesi ve kapak hastalıklarının belirlenmesinde kullanılan temel klinik yöntemlerden biridir. Ancak bu yöntemin doğruluğu büyük ölçüde klinisyenin deneyimine bağlıdır ve bu durum tanısal sonuçlarda değişkenliğe yol açabilmektedir.

Bu çalışma, söz konusu sınırlamaları azaltmak amacıyla kalp seslerinin sınıflandırılmasına yönelik derin öğrenme tabanlı otomatik bir sistem önermektedir. Bu kapsamda fonokardiyogram (FKG) sinyalleri Mel-spektrogram temsiline dönüştürülerek zaman-frekans alanındaki ayırt edici özellikler elde edilmiştir.

Elde edilen özellikler, kalp seslerini dört sınıfa ayırmak için kullanılmıştır: normal kalp sesleri, üfürümlü anormal sesler, normal sonuçlu üfürümler ve üfürümsüz anormal sesler. Sınıflandırma sürecinde Konvolüsyonel Sinir Ağları (CNN) kullanılmış ve ResNet50 ile ResNet152 gibi transfer öğrenme tabanlı mimarilerden yararlanılmıştır.

Veri kümesindeki sınıf dengesizliğini azaltmak için hedefli veri artırma yöntemleri uygulanmıştır. Deneysel sonuçlar, ResNet152 modelinin yaklaşık %84.6 doğruluk oranı ile en iyi performansı gösterdiğini ortaya koymuştur.

Elde edilen bulgular, önerilen yaklaşımın yüksek ayırt edicilik sağladığını ve klinik karar destek sistemlerinde kullanılabileceğini göstermektedir.

Anahtar Kelimeler: Derin öğrenme, Kalp sesi sınıflandırma, Mel-spektrogram, ResNet, Üfürüm algılama

SUMMARY

Heart auscultation is a fundamental clinical method used to evaluate heart sounds and detect valve diseases. However, its accuracy largely depends on the clinician's experience, which may lead to variability in diagnostic outcomes.

This study proposes a deep learning-based automatic system for classifying heart sounds to reduce these limitations. Phonocardiogram (PCG) signals are first transformed into Mel-spectrogram representations to extract discriminative time-frequency features.

These features are used to classify heart sounds into four categories: normal, murmur-related abnormal, normal outcome murmurs, and abnormal sounds without murmurs. Convolutional Neural Networks (CNNs) are employed, utilizing transfer learning architectures such as ResNet50 and ResNet152.

To address class imbalance, targeted data augmentation techniques are applied. Experimental results show that the ResNet152-based model achieves the highest performance with approximately 84.6% accuracy.

The findings demonstrate that the proposed approach provides strong discriminative capability and has significant potential for use in clinical decision support systems.

Keywords: Deep learning, Heart sound classification, Mel-spectrogram, ResNet, Murmur detection

İÇİNDEKİLER

KABUL VE ONAY	III
BİLİMSEL ETİĞE UYGUNLUK BEYANI.....	IV
TEŞEKKÜR.....	V
ÖZET.....	VI
SUMMARY	VII
İÇİNDEKİLER	VIII
TABLolar DİZİNİ	XIII
ŞEKİLLER DİZİNİ.....	XIV
SİMGELER VE KISALTMALAR DİZİNİ	XV
1. GİRİŞ	1
1.1. Kalp Sesi Analizinin Önemi.....	1
1.2. Fonokardiyogram (FKG) Tabanlı Kalp Üfürümü Analizi ve Otomatik Tespit Yaklaşımları.....	2
1.3. Araştırmanın Motivasyonu	4
1.3.1. Klinik ve Halk Sağlığı Motivasyonu.....	4
1.3.2. Teknolojik ve Metodolojik Motivasyon.....	4
1.3.3. Pratik Motivasyon ve Uygulama Potansiyeli	6
1.4. Araştırma Problemi ve Mevcut Yaklaşımların Sınırlamaları.....	7
1.5. Literatürdeki Araştırma Boşlukları ve Tezin Konumlandırılması	8
1.5.1. Pediatrik Kalp Seslerinde Sınırlı Odaklanma ve Yetersiz Üfürüm Açıklamaları...9	
1.5.2. Özellik Temsillerinin Parçalı Kullanımı	9
1.5.3. Pediatrik Üfürüm Tespitinde Hibrit CNN–LSTM Mimarilerinin Sınırlı İncelenmesi	10
1.5.4. Ön İşleme, Veri Artırma ve Veri Bölme Stratejilerinin Yetersiz Raporlanması ..11	
1.5.5. Klinik Alaka Düzeyi ve Uygulama Kısıtlamalarına İlişkin Sınırlı Tartışmalar....11	
1.5.6. Tezin Literatüre Katkısı ve Konumlandırılması	12
1.6. Tezin Genel Amacı	12
1.7. Araştırmanın Amaçları ve Özel Hedefleri	13
1.7.1. Veri Setinin Anlaşılması ve Hazırlanması	13
1.7.2. Özellik Çıkarma ve Seçme Süreci.....	14
1.7.3. Temel (Baseline) Modellerin Geliştirilmesi ve Değerlendirilmesi	15

1.7.4. Derin Öğrenme Modellerinin Tasarımı ve Karşılaştırılması	15
1.7.5. Hibrit CNN–LSTM Çerçevesinin Geliştirilmesi	16
1.8. Araştırma Soruları.....	18
1.9. Çalışmanın Önemi.....	20
1.9.1. Klinik Önemi.....	20
1.9.2. Metodolojik Önemi	21
1.9.3. Toplumsal ve Sağlıkta Eşitlik Açısından Önemi	21
1.10. Tezin Katkıları	22
1.10.1. Pediatri Odaklı Kapsamlı FKG Analiz Süreci	22
1.10.2. Kalp Üfürümü Tespiti için Zengin ve Hibrit Özellik Temsili	23
1.10.3. Geleneksel ve Derin Öğrenme Modellerinin Sistemik Karşılaştırılması	24
1.10.4. Hibrit CNN–LSTM Çerçevesinin Tasarımı ve Değerlendirilmesi	24
1.10.5. Klinik Odaklı Değerlendirme ve Yorumlama.....	25
1.11. Kapsam ve Sınırlamalar	25
1.11.1. Kapsam.....	26
1.11.2. Sınırlamalar	27
1.12. Tezin yapısı	27
2. LİTERATÜR TARAMASI.....	30
2.1. Giriş.....	30
2.2. Kardiyovasküler Fizyoloji ve Kalp Seslerinin Akustik Temeli	31
2.3. Kalp Üfürümlerinin Akustik Kökeni ve Klinik Önemi.....	33
2.4. Fonokardiyografi (FKG) ve Çok Noktalı Oskültasyon.....	35
2.5. Fonokardiyografi (FKG) Analizi için Sayısal Sinyal İşleme (DSP) Teknikleri	37
2.6. Özellik Temsili Olarak Mel-spektrogramlar	39
2.7. Klasik Makine Öğrenimi Yaklaşımları	41
2.8. Otomatik Kalp Üfürümü Tespiti için Derin Öğrenme	42
2.9. 2022–2024 Dönemi En Gelişmiş Üfürüm Tespit Sistemleri	43
2.10. Literatür Taramasının Özeti	45
3. METODOLOJİ	49
3.1. Araştırma Metodolojisine Genel Bakış.....	49
3.1.1. Veri Kümesinin Elde Edilmesi ve Seçimi.....	49
3.1.2. Veri Ön İşleme ve Veri Artırma.....	49
3.1.3. Hibrit Özellik Çıkarımı	49
3.1.4. Model Mimarisinin Tasarımı	50
3.1.5. Eğitim Süreci ve Hiperparametre Ayarlaması	50

3.1.6. Değerlendirme ve Karşılaştırma	51
3.2. Veri Kümesi Tanımı.....	52
3.2.1. Pediatrik Kalp Sesi Veri Kümesinin Kaynağı.....	52
3.2.2. Sınıf Dağılımı.....	52
3.2.3. Etik Hususlar.....	53
3.3. Veri Ön İşleme İş Akışı.....	54
3.3.1. Gürültüyle Başa Çıkma ve Sinyal Bütünlüğü	54
3.3.2. Kalp Sesi Segmentasyonu (S1 , S2 , Sistol , Diyastol)	55
3.3.3. Yeniden Örnekleme ve Uzunluk Normalizasyonu	56
3.3.4. Veri Artırma	57
3.4. Özellik Çıkarma Yöntemleri	58
3.4.1. Mel--Spektrogram Üretimi.....	58
3.4.2. Gelişmiş Mel-spektrogram Özellikleri.....	59
3.4.3. Geleneksel Ses Özellikleri	61
3.4.4. Hibrit Özellik Seçiminin Gerekçesi	62
3.5. Model Mimarisinin Tasarımı	63
3.5.1. Mel-spektrogramlar İçin CNN Mimarisi	64
3.5.2. Zamansal Bağımlılıkların Modellenmesi için LSTM Modülü.....	65
3.5.3. Hibrit CNN–LSTM Birleştirme Stratejisi.....	67
3.5.4. Mel ve Geleneksel Özellikler için Çoklu Girdili Model.....	68
3.5.5. Düzenleştirme, Aktivasyon Fonksiyonları ve Optimizasyon	69
3.6. Eğitim Süreci.....	69
3.6.1. Veri Bölme Stratejisi.....	70
3.6.2. Hiperparametre Ayarlaması	71
3.6.3. Kayıp Fonksiyonları ve Optimizasyon Algoritmaları.....	72
3.7. Değerlendirme Ölçütleri.....	73
3.8. Deneysel Kurulum	74
3.8.1. Donanım Ortamı.....	74
3.8.2. Yazılım Ortamı.....	74
3.9. Karşılaştırma için Temel (Baseline) Modeller.....	75
3.9.1. Geleneksel Makine Öğrenimi Sınıflandırıcıları	76
3.9.2. Temel Derin Öğrenme Modelleri.....	76
3.9.3. Gelişmiş Transfer Öğrenme Mimarileri.....	77
3.10. İstatistiksel Analiz.....	78
3.11. Doğrulama Yöntemi.....	79

3.11.1. Dahili Doğrulama.....	79
3.11.2. Harici Doğrulama.....	79
3.12. Bölüm Özeti	80
4. SONUÇLAR	82
4.1. Sonuçlar Bölümüne Giriş.....	82
4.2. Veri Kümesi Özeti ve Deneysel Kurulum Sonuçları	83
4.2.1. Ön İşleme Öncesi ve Sonrası veri Kümesi Bileşimi	83
4.2.2. Ön İşleme, Segmentasyon ve Veri Artırma	84
4.2.3. Veri Artırma ve Bölme Sonrası Nihai Veri Kümesi	85
4.2.4. Deneysel Kurulumun Genel Görünümü.....	86
4.3. Hibrit CNN–LSTM Modelinin Performansı.....	87
4.3.1. Genel Performans Ölçütleri.....	87
4.3.2. Doğruluk, Kesinlik, Duyarlılık ve F1-Skorunun Yorumlanması.....	88
4.3.3. Özgüllük ve Duyarlılık Açısından Değerlendirme	88
4.3.4. Karmaşıklık (Confusion) Matrisi	89
4.3.5. ROC Eğrisi ve AUC.....	90
4.4. Eğitim ve Doğrulama Eğrileri	91
4.4.1. CNN–LSTM Modelinin Öğrenme Dinamikleri.....	91
4.4.2. Aşırı Öğrenme ve Yetersiz Öğrenmenin Değerlendirilmesi.....	91
4.5. Özellik Katkısı Analizi.....	93
4.5.1. Mel-spektrogram Tabanlı Özelliklerin Rolü.....	93
4.5.2. Geleneksel Ses Özelliklerinin Katkısı.....	93
4.5.3. Hibrit Temsil ve Ayırt Edici Güç.....	94
4.6. Temel (Baseline) Modellerle Karşılaştırma	95
4.6.1. Geleneksel Makine Öğrenimi Sınıflandırıcıları	96
4.6.2. Temel Gözlemler.....	96
4.6.3. Gelişmiş Derin Öğrenme Mimarileri	97
4.6.4. İstatistiksel Değerlendirme ve Sonuçların Yorumlanması.....	99
4.7. Hata Analizi	100
4.8. Bulguların Tartışılması.....	101
4.8.1. Modelin Öğrendikleri.....	101
4.8.2. Hibrit Mimarinin Performansı Neden İyileştirdiği	102
4.8.3. Önceki Literatürle İlişki (2020–2025)	103
4.9. Sonuçlar Bölümünün Özeti	104
5. TARTIŞMA	108

5.1. Giriş.....	108
5.2. Temel Bulguların Yorumlanması.....	108
5.2.1. Geri Çağırma ve Duyarlılığın Klinik Önemi	107
5.2.2. Kritik Bir Performans Ölçütü Olarak F1-Skoru.....	107
5.3. Hibrit CNN–LSTM Mimarisi Neden Daha İyi Performans Gösterdi.....	108
5.3.1. CNN ile Spektral Öğrenmenin Rolü	108
5.3.2. LSTM İle Zamansal Modelleme	109
5.3.3. CNN ve LSTM Bütünleşmesinin Etkisi.....	109
5.4. Önceki Çalışmalarla Karşılaştırma (2020–2025).....	110
5.4.1. CirCor ve Physionet Literatürü ile Uyum	110
5.4.2. Çok Görevli Yüksek Doğruluklu CNN Modelleriyle Karşılaştırma.....	111
5.5. Önerilen Yöntemin Güçlü Yönleri.....	112
5.6. Çalışmanın Sınırlılıkları	113
5.7. Klinik Uygulamalar İçin Çıkarımlar	114
5.8. Gelecek Çalışmalar İçin Öneriler.....	115
5.8.1. Gerçek Hastane Verileri İle Doğrulama.....	115
5.9. Özet	116
6. SONUÇ VE GELECEK ÇALIŞMALAR.....	118
6.1. Giriş.....	118
6.2. Araştırmanın Özeti	118
6.3. Başlıca Bulguların Özeti	119
6.3.1. Spektro-Zamansal Modellemenin Önemi	119
6.3.2. Geleneksel Makine Öğreniminin İçgörü Sağlaması Ancak Optimal Performans Sunmaması.....	120
6.3.3. Sağlam Ön İşlemenin Model Genellemesine Katkısı	121
6.4. Bilimsel ve Klinik Katkılar	121
6.5. Çalışmanın Sınırlılıkları	122
6.6. Gelecek Çalışmalar için Öneriler	123
6.7. Nihai Sonuç	123
KAYNAKÇA.....	125
ÖZGEÇMİŞ	129

TABLÖLAR DİZİNİ

Tablo 1. FKG analizinde kullanılan özellik temsillerinin karşılaştırılması	40
Tablo 2. Literatürde otomatik kalp üfürümü tespiti için kullanılan yaklaşımların özellik temsilleri, modelleme teknikleri, güçlü yönleri ve temel sınırlamaları açısından karşılaştırmalı özet	46
Tablo 3. Ön işleme öncesi FKG veri kümesinin özeti (hasta düzeyi) (Değerler CirCor DigiScope veri kümesinden türetilmiş ve uygulanan filtreleme kurallarına dayanmaktadır.)	84
Tablo 4. Veri artırma sonrasında segment düzeyindeki sınıf dağılımının özeti	85
Tablo 5. Test kümesinde hibrit CNN–LSTM performansı (üfürüm var / yok)	88
Tablo 6. Eğitim ve test performansları açısından aşırı öğrenme karşılaştırması	92
Tablo 7. Geleneksel makine öğrenimi modelleri ile CNN–LSTM’nin karşılaştırılması (F1-skoru)	96
Tablo 8. Kalp üfürümü tespitinde gelişmiş derin öğrenme mimarilerinin performans karşılaştırması (test kümesi)	98
Tablo 9. CNN–LSTM modelinin önceki kalp üfürümü tespit çalışmalarıyla karşılaştırılması	103

ŞEKİLLER DİZİNİ

Şekil 1. Kardiyak oskültasyon sırasında duyulan temel kalp seslerinin (S1 ve S2) ve patolojik üfürümlerin şematik gösterimi.....	2
Şekil 2. Kalp sesi sinyallerinin derin öğrenme tabanlı sınıflandırma için kullanılan Mel-spektrogram zaman–frekans temsiline dönüştürülmesinin örneği.....	5
Şekil 3. Mel-spektrogram girdileri üzerinde çalışan önerilen hibrit CNN–LSTM mimarisinin genel yapısı	17
Şekil 4. Kalp döngüsü sırasında birinci (S1) ve ikinci (S2) kalp seslerinin oluşumu.....	32
Şekil 5. Normal ve patolojik kalp üfürümlerinin Mel-spektrogram zaman–frekans temsillerinin karşılaştırılması.....	35
Şekil 6. Hibrit CNN–LSTM üfürüm tespit modeli için ROC eğrisi ($AUC = 0,879$)....	90
Şekil 7. Kalp üfürümü tespiti için kullanılan gelişmiş modellerin karmaşıklık (confusion) matrisleri.	99

SİMGELER VE KISALTMALAR DİZİNİ

Adam	: Adaptive Moment Estimation (Uyarlamalı Moment Tahmini)
ANOVA	: Analysis of Variance (Varyans Analizi)
AUC	: Area Under the Curve (Eğri Altındaki Alan)
BER	: Band Energy Ratio (Bant Enerji Oranı)
CHD	: Congenital Heart Disease (Konjenital Kalp Hastalığı)
CinC	: Computers in Cardiology (Kardiyolojide Bilgisayarlar)
CNN	: Convolutional Neural Network (Evrişimli Sinir Ağı)
Cohen's κ	: Cohen's Kappa Coefficient (Cohen'in Kappa Katsayısı)
CPU	: Central Processing Unit (Merkezi İşlem Birimi)
CUDA	: Compute Unified Device Architecture (Birleşik Aygıt Hesaplama Mimarisi)
CVD	: Cardiovascular Diseases (Kardiyovasküler Hastalıklar)
CWT	: Continuous Wavelet Transform (Sürekli Dalgacık Dönüşümü)
D	: Accuracy (Doğruluk)
DL	: Deep Learning (Derin Öğrenme)
DSP	: Digital Signal Processing (Dijital Sinyal İşleme)
ECG	: Electrocardiogram (Elektrokardiyogram)
EEG	: Electroencephalogram (Elektroensefalogram)
EfficientNet	: Efficient Neural Network Architecture (Verimli Sinir Ağı Mimarisi)
F1	: F1 Score (F1 Skoru)
FFT	: Fast Fourier Transform (Hızlı Fourier Dönüşümü)
FKG	: Phonocardiogram (Fonokardiyogram)
FN	: False Negative (Yanlış Negatif)
FP	: False Positive (Yanlış Pozitif)
GhostNet	: Lightweight Convolutional Neural Network (Hafifletilmiş Evrişimli Sinir Ağı)
GPU	: Graphics Processing Unit (Grafik İşlem Birimi)
Grad-CAM	: Gradient-weighted Class Activation Mapping (Gradyan Ağırlıklı Sınıf Aktivasyon Haritalaması)
Hz	: Hertz (Hertz)
IoT	: Internet of Things (Nesnelerin İnterneti)
kHz	: Kilohertz (Kilohertz)
KKH	: Congenital Heart Disease (Konjenital Kalp Hastalığı)

KNN	: K-Nearest Neighbors (K-En Yakın Komşu)
LDA	: Linear Discriminant Analysis (Doğrusal Diskriminant Analizi)
LIME	: Local Interpretable Model-Agnostic Explanations (Yerel Yorumlanabilir Model-Agnostik Açıklamalar)
LR	: Learning Rate (Öğrenme Oranı)
LSTM	: Long Short-Term Memory (Uzun Kısa Süreli Bellek)
MB	: Megabyte (Megabayt)
MFCC	: Mel-Frequency Cepstral Coefficients (Mel-Frekans Kepstral Katsayıları)
ML	: Machine Learning (Makine Öğrenimi)
MobileNet	: Mobile-Oriented Neural Network Architecture (Mobil Odaklı Sinir Ağı)
MV	: Mitral Valve (Mitral Kapak Alanı)
PSD	: Power Spectral Density (Güç Spektral Yoğunluğu)
PV	: Pulmonary Valve (Pulmoner Kapak Alanı)
QDA	: Quadratic Discriminant Analysis (Kuadratik Diskriminant Analizi)
ReLU	: Rectified Linear Unit (Doğrultulmuş Doğrusal Birim)
ResNet152	: Residual Network (152 Layers) (Artık Ağ – 152 Katmanlı)
ResNet50	: Residual Network (50 Layers) (Artık Ağ – 50 Katmanlı)
RF	: Random Forest (Rastgele Orman)
RFE	: Recursive Feature Elimination (Özyinelemeli Özellik Eleme)
RMS	: Root Mean Square (Karekök Ortalama Kare)
RNN	: Recurrent Neural Network (Tekrarlayan Sinir Ağı)
ROC	: Receiver Operating Characteristic Curve (Alıcı İşletim Karakteristiği Eğrisi)
S1	: First Heart Sound (Birinci Kalp Sesi)
S2	: Second Heart Sound (İkinci Kalp Sesi)
S3	: Third Heart Sound (Üçüncü Kalp Sesi)
S4	: Fourth Heart Sound (Dördüncü Kalp Sesi)
SHAP	: Shapley Additive Explanations (Shapley Katkı Değerleri)
SOTA	: State of the Art (En İleri Düzey)
STFT	: Short-Time Fourier Transform (Kısa Zamanlı Fourier Dönüşümü)
SVM	: Support Vector Machines (Destek Vektör Makineleri)
TN	: True Negative (Gerçek Negatif)
TP	: True Positive (Gerçek Pozitif)
TV	: Tricuspid Valve (Triküspit Kapak Alanı)
WAV	: Waveform Audio File Format (Dalga Biçimi Ses Dosyası Formatı)
ZCR	: Zero Crossing Rate (Sıfır Geçiş Oranı)

1. GİRİŞ

1.1. Kalp Sesi Analizinin Önemi

Kardiyovasküler hastalıklar (KVH'ler), küresel sağlık açısından önemli bir yük oluşturmaktadır ve her yıl yaklaşık 17,9 milyon kişinin ölümüne neden olarak dünya genelinde önde gelen ölüm nedeni olmaya devam etmektedir (Dünya Sağlık Örgütü [WHO], 2021). Bu hastalıkların etkisi, tanı araçlarına, uzman sağlık profesyonellerine ve erken tarama sistemlerine erişimin genellikle sınırlı olduğu düşük ve orta gelirli ülkelerde daha belirgin hâle gelmektedir.

KVH'ler arasında konjenital kalp hastalığı (KKH), pediatrik popülasyon açısından kritik bir sağlık sorunudur. Küresel ölçekte canlı doğumların yaklaşık %1'ini etkilemekte ve çocukluk çağı morbiditesi ile mortalitesine önemli ölçüde katkıda bulunmaktadır (Amerikan Kalp Derneği [AHA], 2023). Kalp anomalilerinin erken teşhisi büyük önem taşımaktadır; çünkü gecikmiş tanı kalp yetmezliği, gelişimsel bozukluklar ve yaşam süresinin kısalması gibi ciddi uzun vadeli komplikasyonlara yol açabilmektedir (AHA, 2023).

Geleneksel olarak akustik stetoskop kullanılarak gerçekleştirilen kardiyak oskültasyon, kalp seslerini değerlendirmek ve üfürümleri tespit etmek amacıyla kullanılan temel klinik yöntemlerden biridir. Birinci kalp sesi (S1) ve ikinci kalp sesi (S2), fizyolojik kapak kapanışlarıyla ilişkilidir. Buna karşılık üçüncü kalp sesi (S3), dördüncü kalp sesi (S4) ve üfürümler gibi ek akustik bulgular, altta yatan patolojik durumların göstergesi olabilmektedir.

Klinik açıdan önemli bir yöntem olmasına rağmen oskültasyon, büyük ölçüde klinisyenin eğitimi ve deneyimine bağlı olan öznel bir değerlendirme sürecidir. Bu durum, deneyimli hekimler arasında dahi gözlemciler arası farklılıklara yol açabilmektedir (Reyna vd., 2023). Bu sınırlamalar, özellikle birinci basamak sağlık hizmetlerinde ve kaynakların kısıtlı olduğu bölgelerde daha belirgin hâle gelmektedir.

Dijital stetoskopların ve fonokardiyogramın (FKG) geliştirilmesi, oskültasyon sürecinin bilgisayar ortamında kaydedilip analiz edilebilen ölçülebilir bir biyomedikal sinyale dönüştürülmesini mümkün kılmıştır. FKG, kalp seslerinin zaman-frekans analizinin gerçekleştirilmesine, nicel özelliklerin çıkarılmasına ve kayıtların uzun süreli olarak depolanmasına olanak sağlamaktadır.

Bunun yanı sıra, CirCor DigiScope pediatrik veri kümesi gibi büyük ölçekli veri tabanlarının erişilebilir hâle gelmesi, otomatik kalp üfürümü tespiti amacıyla yapay zekâ tabanlı tanı sistemlerinin geliştirilmesini önemli ölçüde kolaylaştırmıştır (Oliveira vd., 2022).



Şekil 1. Kardiyak oskültasyon sırasında duyulan temel kalp seslerinin (S1 ve S2) ve patolojik üfürümlerin şematik gösterimi

1.2. Fonokardiyogram (FKG) Tabanlı Kalp Üfürümü Analizi ve Otomatik Tespit Yaklaşımları

FKG kayıtlarından kalp üfürümlerinin otomatik olarak tespit edilmesi, biyomedikal sinyal işleme ve yapay zekâ alanlarında yoğun ilgi gören araştırma konularından biri hâline gelmiştir. Türbülanslı veya anormal kan akışının akustik yansıması olan kalp üfürümleri, özellikle pediatrik popülasyonda konjenital kalp hastalıklarının erken göstergeleri arasında yer almaktadır (Reyna vd., 2023). Bununla birlikte geleneksel oskültasyon yöntemi, klinisyen deneyimine bağımlılık, çevresel gürültü ve işitsel yorumlamanın öznel doğası gibi çeşitli sınırlamalar içermektedir. Bu nedenle hem klinik ortamlarda hem de kaynakların sınırlı olduğu bölgelerde güvenilir tarama sağlayabilecek otomatik ve nesnel sistemlerin geliştirilmesine yönelik ilgi giderek artmaktadır (Singh ve Agarwal, 2024).

Son on yılda dijital stetoskoplar ve yüksek kaliteli FKG toplama sistemleri önemli ölçüde gelişmiş; iyileştirilmiş bant genişliği ve azaltılmış ortam gürültüsü sayesinde daha tutarlı ve yüksek kaliteli kalp sesi kayıtlarının elde edilmesi mümkün hâle gelmiştir. Bu teknolojik gelişmeler, araştırmacıların kalp sesi sinyallerine gömülü tanınal örüntüleri ortaya çıkarmak amacıyla el yapımı özellik mühendisliğinden derin öğrenme tabanlı uçtan uca modellere kadar uzanan daha gelişmiş hesaplama tekniklerini incelemelerine olanak sağlamıştır (Araujo vd., 2022).

FKG sinyallerinin doğası gereği durağan olmayan ve hastalar arasında önemli farklılıklar gösteren biyomedikal sinyaller olması nedeniyle, literatürde normal ve patolojik kalp fonksiyonlarıyla ilişkili akustik özellikleri daha iyi yakalayabilmek için zaman alanı, frekans alanı ve zaman–frekans temsillerinin birlikte kullanılmasının önemi vurgulanmaktadır (Chen vd., 2023).

Son yıllarda George B. Moody PhysioNet/Kardiyolojide Bilgisayarlar (CinC) Challenge 2022 veri seti ve CirCor DigiScope koleksiyonu gibi büyük ve özenle hazırlanmış veri kümelerinin erişilebilir hâle gelmesiyle bu alanda önemli bir gelişme yaşanmıştır. Bu veri kümeleri, tekrarlanabilir karşılaştırmalı değerlendirmelerin yapılmasına olanak sağlamış ve araştırmacıların farklı sınıflandırma yaklaşımlarının performansını standart koşullar altında değerlendirmesine imkân tanımıştır (Oliveira vd., 2022; Reyna vd., 2023).

Bununla birlikte evrişimli sinir ağları (Konvolüsyonel Sinir Ağları (CNN) – CNN), Uzun Kısa Süreli Bellek ağları (Long Short-Term Memory – LSTM) gibi tekrarlayan mimariler ve hibrit CNN–LSTM yapıları dâhil olmak üzere derin öğrenme tabanlı modeller, FKG sinyallerinde bulunan spektral ve zamansal bağımlılıkları birlikte öğrenerek önemli performans iyileştirmeleri sağlamıştır (Netto ve Abraham, 2021; Zhao vd., 2023).

Daha önceki hesaplamalı yaklaşımlar ise ağırlıklı olarak geleneksel sinyal işleme ve makine öğrenimi yöntemlerine odaklanmıştır. Bu yöntemlerde, klasik sınıflandırıcıların eğitilmesi için spektral enerji, sıfır geçiş oranı, zarf varyasyonları ve Mel Frekans Cepstral Katsayıları (Mel-Frequency Cepstral Coefficients – MFCC) gibi elle tasarlanmış özellikler kullanılmaktadır.

Son yıllarda derin öğrenme yaklaşımları, ham veya minimum düzeyde işlenmiş sinyallerden doğrudan temsil öğrenimini mümkün kılmıştır. Evrişimli sinir ağları (CNN’ler), spektrogram tabanlı kalp sesi sınıflandırması görevlerinde güçlü performans sergilerken; LSTM ağları gibi tekrarlayan mimariler kalp döngüleri arasındaki zamansal bağımlılıkların modellenmesinde etkili olmaktadır. Uzamsal ve zamansal özellik öğrenimini birleştiren hibrit CNN–LSTM modelleri ise biyomedikal ses analizi görevlerinde daha yüksek sağlamlık ve genellenebilirlik göstermektedir (Nie vd., 2024).

Bu bağlamda, bu tez çalışması CirCor DigiScope veri kümesini kullanarak pediatrik kalp üfürümlerinin otomatik tespitini geliştirmeyi amaçlamaktadır. Bu amaç doğrultusunda Mel-spektrogram temsilleri ile geleneksel ses özelliklerini birleştiren hibrit bir CNN–LSTM çerçevesi önerilmektedir. Önerilen yaklaşımın, sınıflandırma doğruluğunu artırması ve çocuklarda konjenital kalp hastalıklarının erken taranmasına

yönelik güvenilir bilgisayar destekli tanı sistemlerinin geliştirilmesine katkı sağlaması hedeflenmektedir.

1.3. Araştırmanın Motivasyonu

Bu araştırmanın motivasyonu, klinik gereksinimler, sağlık hizmetlerine erişimde eşitlik ve teknolojik olanakların kesişiminden kaynaklanmaktadır.

1.3.1. Klinik ve Halk Sağlığı Motivasyonu

Pediyatrik kalp hastalıkları, özellikle uzman kardioloji hizmetlerinin bulunmadığı bölgelerde önemli bir halk sağlığı sorunu olmaya devam etmektedir. Kaynakların sınırlı olduğu birçok sağlık ortamında, genel pratisyen hekimler tarafından gerçekleştirilen oskültasyon, çoğu zaman birincil ve bazen de tek tarama yöntemi olarak kullanılmaktadır. Bununla birlikte araştırmalar, uzman olmayan klinisyenlerin patolojik üfürümlerin yaklaşık %30–50'sini gözden kaçırabileceğini göstermektedir (AHA, 2023). Tanıda meydana gelen bu tür gecikmeler, geri dönüşü olmayan kalp hasarı, artan morbidite ve sağlık hizmetleri üzerindeki yükün artması gibi ciddi sonuçlara yol açabilmektedir.

Pediyatrik tanı sürecini daha da karmaşık hâle getiren bir diğer faktör ise çocuklarda masum üfürümlerin yüksek görülme oranıdır. Çocukların yaklaşık %70'inde görülebilen bu masum üfürümler, gereksiz sevklerin artmasına, ebeveynlerde kaygı oluşmasına ve sınırlı sayıdaki ekokardiyografi hizmetleri üzerinde ek bir yük oluşmasına neden olabilmektedir (Singh ve Agarwal, 2024). Bu bağlamda objektif ve otomatik bir üfürüm tespit sistemi, triyaj doğruluğunu artırarak gereksiz sevkleri azaltabilir ve yüksek riskli çocukların zamanında değerlendirilmesine katkı sağlayabilir.

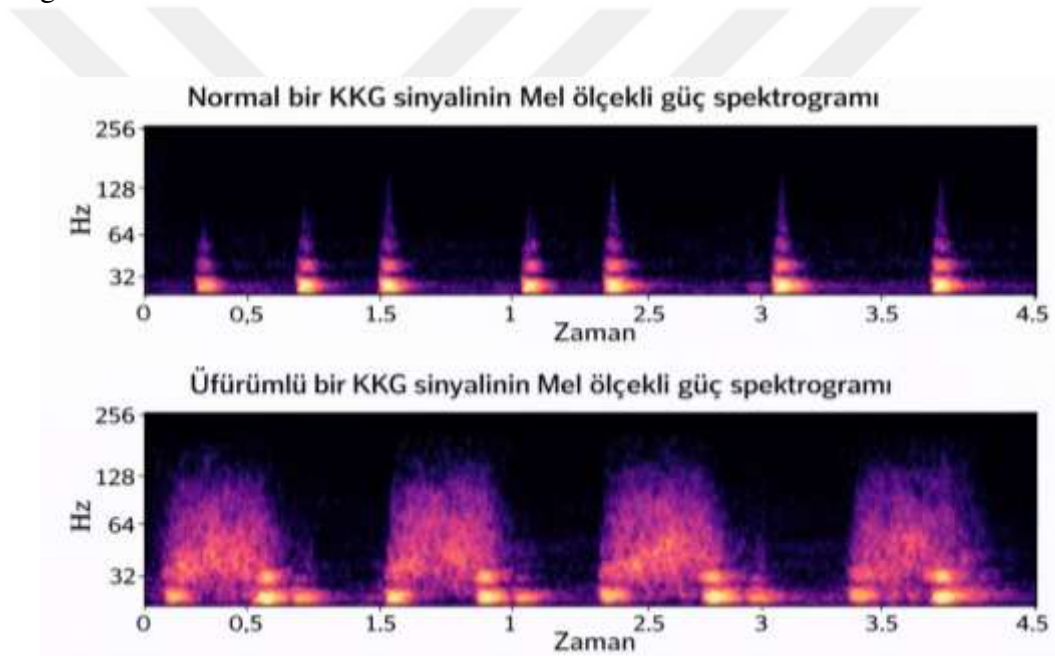
1.3.2. Teknolojik ve Metodolojik Motivasyon

Teknolojik açıdan kalp sesi analizi çeşitli zorluklar içermektedir. FKG sinyalleri durağan olmayan sinyaller olup zamanla değişen spektral içerik, geçici olaylar ve örtüşen bileşenler barındırmaktadır. Ayrıca kayıtlar çoğu zaman çevresel gürültü, hasta hareketlerinden kaynaklanan artefaktlar ve sensör yerleşimi ile cihaz özelliklerindeki değişkenliklerden etkilenmektedir. Bunun yanı sıra üfürümler, zamanlamalarına (sistolik, diyastolik veya sürekli), kökenlerine (kapak lezyonları veya şant lezyonları) ve hemodinamik şiddetlerine bağlı olarak farklı zaman–frekans özellikleri sergilemektedir.

Geleneksel özellik mühendisliği teknikleri, önemli bilgiler sağlayabilmesine rağmen kalp sesi sinyallerinde bulunan örüntülerin doğasında yer alan karmaşıklığı her zaman tam olarak yakalayamayabilir. Buna karşılık derin öğrenme modelleri, özellikle

Evrişimli Sinir Ağları (Konvolüsyonel Sinir Ağları (CNN) – CNN) ve Uzun Kısa Süreli Bellek ağları (Long Short-Term Memory – LSTM), Mel-spektrogram gibi zaman–frekans temsillerinden doğrudan zengin ve hiyerarşik özellikler öğrenebilme yeteneğine sahiptir.

Şekil 2’de gösterildiği gibi Mel-spektrogram temsili, FKG sinyallerini iki boyutlu zaman–frekans görüntülerine dönüştürmektedir. Bu dönüşüm, derin öğrenme modellerinin kalp seslerinin hem zamansal hem de spektral özelliklerini daha etkili biçimde öğrenmesine olanak sağlamaktadır. Bununla birlikte bu tür modeller manuel özellik çıkarımı gereksinimini büyük ölçüde ortadan kaldırırsa da, özellikle sınırlı veya dengesiz veri kümeleriyle çalışırken aşırı uyumun (overfitting) azaltılması için dikkatli bir mimari tasarım, uygun düzenleme stratejileri ve etkili veri ön işleme adımları gerekmektedir.



Şekil 2. Kalp sesi sinyallerinin derin öğrenme tabanlı sınıflandırma için kullanılan Mel-spektrogram zaman–frekans temsiline dönüştürülmesinin örneği

CirCor DigiScope veri kümesi, pediatrik popülasyonda kalp üfürümü tespitine ilişkin zorlukların incelenmesi için önemli bir araştırma ortamı sunmaktadır (Oliveira vd., 2022). Bu veri kümesi, aort, pulmoner, triküspit ve mitral olmak üzere dört farklı oskültasyon bölgesinden elde edilen çoklu fonokardiyogram (FKG) kayıtlarının yanı sıra ayrıntılı üfürüm etiketleri ve demografik bilgiler içermektedir. Bu çok yönlü yapı, farklı derin öğrenme mimarilerinin ve çeşitli girdi temsillerinin sistematik ve karşılaştırmalı biçimde değerlendirilmesine olanak tanımaktadır.

Bununla birlikte veri kümesi, klinik tarama senaryolarında sıkça karşılaşılan belirgin bir sınıf dengesizliği problemini de barındırmaktadır. Normal kayıtların üfürüm

içeren kayıtlara kıyasla daha fazla olması, modelleme sürecinde yanlılık (bias) oluşmasına ve sınıflandırma performansının dengesiz şekilde etkilenmesine neden olabilmektedir (He ve Garcia, 2009). Bu durum, aşağıdaki dengesizlik oranı ile nicel olarak ifade edilebilir (Eşitlik 1):

$$IR = \frac{N_{majority}}{N_{minority}} \quad (\text{Eşitlik 1})$$

Bu eşitlikte $N_{majority}$ çoğunluk sınıfındaki örnek sayısını, $N_{minority}$ ise azınlık sınıfındaki örnek sayısını temsil etmektedir.

Bu tez çalışmasının temel amacı, söz konusu zorlukları ele alabilecek metodolojik açıdan titiz ve yeniden üretilebilir bir iş hattı tasarlamaktır. Bu kapsamda, insan işitsel algısıyla uyumlu zaman–frekans yapısını temsil eden Mel-spektrogram temsilleri ile zaman, frekans, kepsral ve istatistiksel alanlardan elde edilen geleneksel ses özellikleri birlikte değerlendirilmiştir (Logan, 2000).

Ayrıca hem geleneksel makine öğrenimi sınıflandırıcıları hem de çağdaş derin öğrenme mimarileri kapsamlı biçimde incelenmiş ve karşılaştırılmıştır. Bu süreç sonucunda, kalp üfürümlerini daha doğru ve daha sağlam biçimde tespit edebilmek amacıyla, spektral örüntüleri öğrenebilen evrimsel sinir ağları ile kardiyak döngüler boyunca zamansal bağımlılıkları modelleyebilen uzun-kısa süreli bellek (LSTM) ağlarını birleştiren hibrit bir CNN–LSTM mimarisi önerilmiştir (Hochreiter ve Schmidhuber, 1997).

1.3.3. Pratik Motivasyon ve Uygulama Potansiyeli

Araştırmanın yenilikçi yönlerinin yanı sıra, önerilen çerçevenin pratikte uygulanabilir olması da bu çalışmanın temel motivasyon kaynaklarından birini oluşturmaktadır. Uzun vadeli bir bakış açısıyla, geliştirilen bu tür modellerin birinci basamak sağlık hizmeti sunan kliniklerde, okullarda veya sağlık hizmetlerine erişimin kısıtlı olduğu uzak bölgelerde tarama amacıyla kullanılan dijital stetoskoplar ve mobil cihazlara entegre edilmesi hedeflenmektedir (Topol, 2019).

Bunun yanı sıra, oskültasyon kayıtlarının yapay zekâ destekli sistemler aracılığıyla uzaktan değerlendirilmesine olanak tanıyan teletıp platformları da önemli bir uygulama alanı sunmaktadır (WHO, 2021). Ayrıca bu çerçevenin, uzman hekimlere yönlendirme yapılmadan önce genel pratisyen hekimlere destekleyici bir ikinci görüş sağlayan klinik karar destek sistemleri kapsamında kullanılması; hem klinik iş yükünün azaltılması hem

de hastaların daha erken ve doğru şekilde yönlendirilmesi açısından önemli katkılar sağlayabilir (Sutton vd., 2020).

Bu vizyonun hayata geçirilebilmesi için geliştirilecek modellerin yalnızca yüksek doğruluk sunması yeterli değildir; aynı zamanda hesaplama açısından verimli, gerçek dünya koşullarındaki değişkenliklere karşı dayanıklı ve klinisyenler ile düzenleyici kurumlar tarafından güvenilir ve yorumlanabilir olarak değerlendirilebilecek nitelikte olması gerekmektedir (Doshi-Velez ve Kim, 2017). Bu tez çalışması, klinik açıdan anlamlı özellikleri derin öğrenme sürecine entegre ederek ve klinik önceliklerle uyumlu değerlendirme ölçütleri kullanarak model performansını sistematik biçimde analiz etmekte ve bu hedeflere katkı sağlamayı amaçlamaktadır.

1.4. Araştırma Problemi ve Mevcut Yaklaşımların Sınırlamaları

Dijital oskültasyon ve makine öğrenimi alanlarında önemli ilerlemeler kaydedilmiş olmasına rağmen, pediatrik kalp üfürümlerinin tespiti konusunda araştırma prototipleri ile klinik uygulamada kullanılabilecek sistemler arasında hâlâ belirgin bir boşluk bulunmaktadır (Clifford vd., 2016).

Mevcut klinik uygulamada pediatrik kalp üfürümlerinin değerlendirilmesi büyük ölçüde klinisyenlerin farklı uzmanlık düzeylerinde gerçekleştirdiği öznel oskültasyona dayanmaktadır. Üfürümün patolojik olabileceğinden şüphelenildiğinde veya klinik belirsizlik devam ettiğinde hastalar genellikle ekokardiyografik incelemeye yönlendirilmektedir. Bununla birlikte bu yaklaşım çeşitli önemli sınırlılıklar içermektedir. Klinik değerlendirmelerde gözlemciler arası değişkenlik yaygın olup, farklı hekimler üfürümün varlığı, zamanlaması veya klinik önemi konusunda farklı yorumlar yapabilmektedir (Mangione vd., 1993).

Buna ek olarak birçok bölgede pediatrik kardiyolog sayısının sınırlı olması tanı süreçlerinin gecikmesine yol açabilmektedir. Ekokardiyografiye yapılan aşırı yönlendirmeler maliyetleri ve bekleme sürelerini artırırken, yetersiz yönlendirmeler ise ciddi kardiyak patolojilerin gözden kaçma riskini beraberinde getirmektedir (McCrindle vd., 2007).

Teknik açıdan değerlendirildiğinde mevcut otomatik üfürüm tespit yöntemleri de çeşitli zorluklarla karşı karşıyadır. Literatürdeki birçok yaklaşım, üfürümün zamanlaması, yoğunluğu ve oskültasyon konumu gibi ayrıntılı tanımlayıcı bilgileri tam olarak kullanmadan problemi çoğunlukla normal ve anormal sınıflar arasında ikili bir sınıflandırma problemi olarak ele almaktadır (Potes vd., 2016).

Geleneksel makine öğrenimi modelleri ise büyük ölçüde elle tasarlanmış özelliklere dayanmakta ve bu özellikler farklı veri kümeleri ile kayıt koşulları arasında yeterli genelleme sağlayamayabilmektedir. Öte yandan bildirilen derin öğrenme yaklaşımlarının önemli bir kısmı yalnızca tek bir temsil biçimini, örneğin Mel-spektrogramlar veya ham dalga formlarını işlemekte ve geleneksel ses özelliklerinde yer alan tamamlayıcı bilgileri göz ardı edebilmektedir (Rubin vd., 2017).

Özellikle pediatrik popülasyonlar söz konusu olduğunda, aynı veri kümesi üzerinde tutarlı bir ön işleme ve değerlendirme protokolü kullanarak geleneksel makine öğrenimi yöntemleri ile derin öğrenme yaklaşımlarını sistematik biçimde karşılaştıran çalışmaların sayısı oldukça sınırlıdır (Reyna vd., 2023).

Bu bağlamda bu tezde ele alınan temel araştırma problemi; Mel-spektrogram temsilleri ile geleneksel ses özelliklerini bütünleştiren ve FKG kayıtlarından pediatrik kalp üfürümlerini yüksek doğruluk, sağlamlık ve potansiyel klinik uygulanabilirlik düzeyiyle otomatik olarak tespit edebilen hibrit bir derin öğrenme çerçevesinin nasıl tasarlanabileceği, uygulanabileceği ve değerlendirilebileceği sorusu etrafında şekillenmektedir. Bu temel problem aynı zamanda optimal sinyal ön işleme stratejilerinin belirlenmesi, etkili özellik çıkarımı ve seçimi, uygun model mimarisinin tasarlanması ve önerilen yaklaşımın geleneksel temel modellerle titiz bir biçimde karşılaştırılması gibi bir dizi alt araştırma sorusunu da kapsamaktadır.

1.5. Literatürdeki Araştırma Boşlukları ve Tezin Konumlandırılması

Giderek artan sayıda çalışma, klasik sinyal işleme yaklaşımlarını makine öğrenimi ve daha yakın dönemde derin öğrenme yöntemleri ile birleştirerek fonokardiyogram (FKG) sinyallerine dayalı otomatik kalp sesi sınıflandırmasını incelemiştir (Clifford vd., 2016; Springer vd., 2016). Bu çalışmalar, özellikle son yıllarda dijital oskültasyon teknolojilerindeki gelişmeler ve büyük ölçekli veri kümelerinin erişilebilir hâle gelmesi sayesinde önemli ilerlemeler kaydetmiştir (Reyna vd., 2023).

Bununla birlikte mevcut literatür incelendiğinde, özellikle pediatrik kalp üfürümlerinin otomatik tespitine yönelik çalışmalarda hem veri temsilleri hem de metodolojik yaklaşımlar açısından çeşitli sınırlılıkların ve araştırma boşluklarının hâlen varlığını sürdürdüğü görülmektedir (Potes vd., 2016; Rubin vd., 2017).

Aşağıdaki alt bölümlerde literatürde öne çıkan bu boşluklar sistematik biçimde ele alınmakta ve bu tez çalışmasının söz konusu boşlukları hangi yönlerden gidermeyi amaçladığı ortaya konulmaktadır.

1.5.1. Pediatrik Kalp Seslerinde Sınırlı Odaklanma ve Yetersiz Üfürüm Açıklamaları

Literatürde yer alan birçok çalışma, otomatik kalp sesi sınıflandırması için çoğunlukla yetişkin veya karma popülasyonlardan elde edilmiş veri kümelerine dayanmaktadır (Clifford vd., 2016; Potes vd., 2016). Bu çalışmalarda kullanılan etiketleme yaklaşımları genellikle sınırlı ayrıntı içermekte ve çoğu durumda kalp sesleri yalnızca “normal” ve “anormal” gibi ikili sınıflar altında ele alınmaktadır.

Ancak pediatrik popülasyonda kalp üfürümleri, zamanlamaları (sistolik, diyastolik veya sürekli), yoğunlukları ve anatomik kökenleri açısından daha karmaşık ve değişken özellikler sergileyebilmektedir (McCrindle vd., 2007). Bu durum, pediatrik vakalar için daha ayrıntılı ve açıklayıcı üfürüm etiketlerinin kullanılmasını gerekli kılmaktadır.

Bu bağlamda CirCor DigiScope veri kümesi, birden fazla oskültasyon bölgesinden elde edilen pediatrik FKG kayıtlarının yanı sıra ayrıntılı üfürüm açıklamaları sunması bakımından literatürde önemli bir konuma sahiptir (Oliveira vd., 2022). Söz konusu veri kümesi, PhysioNet / George B. Moody Challenge kapsamında ve çeşitli araştırmalarda kullanılmış olmasına rağmen, bu veri kümesini temel alarak pediatrik üfürüm tespitine odaklanan ve ön işleme adımları ile model karşılaştırmalarını ayrıntılı biçimde belgeleyen tez düzeyinde sistematik çalışmaların sayısı hâlen sınırlıdır (Reyna vd., 2023).

Ayrıca pediatrik kalp sesleri, yetişkin bireylere kıyasla daha yüksek kalp atım hızları, daha kısa kardiyak döngüler ve gelişimsel anatomik farklılıklar nedeniyle belirgin biçimde değişkenlik göstermektedir. Bu özellikler, yetişkin veri kümeleri üzerinde eğitilmiş modellerin pediatrik vakalara doğrudan uygulanmasını metodolojik açıdan sorunlu hâle getirmektedir (Liu vd., 2016). Dolayısıyla pediatrik kalp seslerine özgü zengin üfürüm açıklamalarını içeren veri kümeleri üzerinde gerçekleştirilen çalışmalar yalnızca sınıflandırma doğruluğu açısından değil, aynı zamanda klinik güvenilirlik ve genellenebilirlik bakımından da kritik bir öneme sahiptir.

1.5.2. Özellik Temsillerinin Parçalı Kullanımı

Mevcut literatürde kalp sesi sınıflandırmasına yönelik yöntemlerin önemli bir bölümünde özellik temsillerinin parçalı bir yaklaşımla ele alındığı görülmektedir (Potes vd., 2016; Rubin vd., 2017). Bazı çalışmalarda enerji, sıfır geçiş oranı (Zero Crossing Rate – ZCR), spektral merkez ve Mel frekans kepsral katsayıları (Mel-Frequency Cepstral Coefficients – MFCC) gibi el ile tasarlanmış özellikler çıkarılarak geleneksel makine öğrenimi sınıflandırıcılarına girdi olarak sunulmaktadır (Logan, 2000). Diğer çalışmalarda ise spektrogramlar, Mel-spektrogramlar veya sabit Q dönüşümü gibi

zaman–frekans temsilleri, görüntü benzeri yapılar hâlinde evrişimli sinir ağlarına aktarılmaktadır (LeCun vd., 2015).

Bu iki yaklaşımın her biri belirli avantajlar sunmasına rağmen tek başına kullanıldıklarında kalp sesi sinyallerinde bulunan tamamlayıcı bilgilerin bütüncül biçimde temsil edilmesini sınırlayabilmektedir. Geleneksel ses özellikleri, sinyalin belirli fiziksel ve istatistiksel özelliklerini açık biçimde yansıtabilmesine karşın karmaşık ve doğrusal olmayan örüntüleri yakalamada yetersiz kalabilmektedir. Buna karşılık derin öğrenme temelli temsiller yüksek boyutlu ve soyut özellikler öğrenebilse de klinik açıdan yorumlanabilirlik bakımından bazı zorluklar barındırmaktadır (Doshi-Velez ve Kim, 2017).

Bu bağlamda, farklı özellik uzaylarının birlikte kullanılması aşağıdaki gibi bir özellik füzyonu yaklaşımı ile ifade edilebilir (Eşitlik 2):

$$F_{combined} = [F_{time}, F_{frequency}, F_{cepstral}, F_{statistical}] \quad (\text{Eşitlik 2})$$

Bu eşitlikte F_{time} , $F_{frequency}$, $F_{cepstral}$ ve $F_{statistical}$ sırasıyla zaman, frekans, kepsral ve istatistiksel alanlardan elde edilen özellik vektörlerini temsil etmektedir.

Bu nedenle Mel-spektrogram temsilleri ile zaman, frekans, kepsral ve istatistiksel alanları kapsayan geleneksel ses özelliklerinin tek bir bütünleşik çerçevede birlikte ele alınması hem sınıflandırma performansının artırılması hem de model çıktılarının klinik açıdan daha anlamlı hâle getirilmesi bakımından önemli bir potansiyel sunmaktadır. Bununla birlikte literatürde bu tür bütünleşik yaklaşımları sistematik biçimde inceleyen çalışmaların sayısı hâlen oldukça sınırlıdır.

1.5.3. Pediatrik Üfürüm Tespitinde Hibrit CNN–LSTM Mimarilerinin Sınırlı İncelenmesi

Derin öğrenme temelli çalışmalar, evrişimli sinir ağlarının (CNN) yerel spektral örüntüleri yakalamada, uzun kısa süreli bellek (LSTM) ağlarının ise FKG sinyallerindeki zamansal bağımlılıkları modellemede etkili olduğunu göstermiştir (LeCun vd., 2015; Hochreiter ve Schmidhuber, 1997). Bu bulgular, kalp sesi analizi alanında CNN ve LSTM tabanlı yaklaşımların yaygın biçimde kullanılmasına yol açmıştır.

Bununla birlikte, bu iki mimarinin avantajlarını açık biçimde birleştiren hibrit CNN–LSTM yaklaşımlarının özellikle pediatrik veri setleri üzerinde sistematik olarak değerlendirildiği çalışmaların sayısı görece sınırlıdır (Rubin vd., 2017).

Mevcut literatürde, Mel-spektrogramlardan CNN tabanlı uzamsal özellik çıkarımını LSTM tabanlı zamansal modelleme ile bütünleştiren çalışmalar da sınırlı sayıdadır. Ayrıca bu hibrit modellerin aynı pediatrik veri seti üzerinde geleneksel makine öğrenimi yöntemleri ve alternatif derin öğrenme mimarileriyle (örneğin yalnızca CNN veya yalnızca LSTM tabanlı modeller ya da transfer öğrenimi yaklaşımlarıyla) kapsamlı biçimde karşılaştırıldığı analizler nadiren raporlanmaktadır. Bu durum, hibrit mimarilerin potansiyel avantajlarının ve sınırlılıklarının daha net biçimde ortaya konulmasını zorlaştırmaktadır.

Bu bağlamda, CNN ve LSTM bileşenlerinin ardışık olarak kullanıldığı hibrit bir model aşağıdaki şekilde ifade edilebilir (Eşitlik 3):

$$h_t = LSTM(CNN(X_t)) \quad (\text{Eşitlik 3})$$

Bu eşitlikte X_t giriş zaman–frekans temsillerini (örneğin Mel-spektrogram), $CNN(\cdot)$ uzamsal özellik çıkarımını ve $LSTM(\cdot)$ ise zamansal bağımlılıkların modellenmesini temsil etmektedir.

1.5.4. Ön İşleme, Veri Artırma ve Veri Bölme Stratejilerinin Yetersiz Raporlanması

Yayınlanan çalışmaların önemli bir bölümünde, ön işleme ve deneysel protokollere ilişkin kritik ayrıntıların yeterince raporlanmadığı görülmektedir (Stodden vd., 2016). Özellikle birden fazla oskültasyon bölgesinden elde edilen kayıtların nasıl işlendiği ve birleştirildiği, değişken uzunluktaki FKG sinyallerinin hangi yöntemlerle standartlaştırıldığı ve üfürüm pozitif vakalarda veri artırma işlemlerinin nasıl gerçekleştirildiği çoğu zaman açık biçimde belirtilmemektedir (Clifford vd., 2016).

Benzer şekilde, eğitim, doğrulama ve test veri kümelerinin hasta düzeyinde veri sızıntısını önleyecek biçimde nasıl oluşturulduğuna ilişkin ayrıntıların eksikliği çalışmaların tekrarlanabilirliğini sınırlamakta ve farklı modeller arasında adil karşılaştırmalar yapılmasını zorlaştırmaktadır (Varoquaux vd., 2017). Bu tür metodolojik belirsizlikler, bildirilen performans sonuçlarının klinik ve bilimsel açıdan güvenilirliğini olumsuz yönde etkileyebilmektedir.

1.5.5. Klinik Alaka Düzeyi ve Uygulama Kısıtlamalarına İlişkin Sınırlı Tartışmalar

Literatürde birçok çalışma yüksek sınıflandırma doğrulukları raporlamasına rağmen, önerilen modellerin klinik iş akışlarına nasıl entegre edilebileceğine ilişkin

tartışmalar genellikle sınırlı kalmaktadır (Topol, 2019). Model karmaşıklığının ve çıkarım sürelerinin taşınabilir cihazlar üzerindeki uygulanabilirliği, enerji tüketimi ve gerçek zamanlı kullanım senaryolarına olan etkileri çoğu zaman ayrıntılı biçimde ele alınmamaktadır (Lane vd., 2015).

Ayrıca model çıktılarının klinik muhakeme süreçleriyle nasıl uyumlu biçimde yorumlanabileceği ve yanlış pozitif (FP) veya yanlış negatif (FN) sınıflandırmaların pediatrik hastalar üzerindeki olası klinik sonuçları da literatürde yeterince tartışılmamıştır (Kelly vd., 2019). Bu eksiklikler, araştırma prototipleri ile klinik olarak kullanılabilir sistemler arasındaki boşluğun devam etmesine katkıda bulunmaktadır.

1.5.6. Tezin Literatüre Katkısı ve Konumlandırılması

Bu tez, yukarıda belirtilen araştırma boşluklarını gidermeyi amaçlamaktadır. Bu kapsamda, CirCor DigiScope veri seti kullanılarak pediatrik kalp üfürümlerinin otomatik tespitine açık biçimde odaklanılmakta; Mel-spektrogram temsilleri, zaman, frekans, kepsral ve istatistiksel alanları kapsayan toplam 128 geleneksel ses özelliği ile bütünleştirilmektedir (Oliveira vd., 2022).

Ayrıca, tutarlı ve tekrarlanabilir bir deneysel protokol çerçevesinde çeşitli geleneksel makine öğrenimi yöntemleri ile derin öğrenme mimarileri sistematik biçimde karşılaştırılmaktadır. Hasta düzeyinde veri bölünmesini içeren ayrıntılı bir ön işleme ve veri artırma süreci sunularak metodolojik şeffaflık sağlanmaktadır (Varoquaux vd., 2017).

Bununla birlikte, önerilen hibrit CNN–LSTM çerçevesinin klinik etkileri ve uygulama potansiyeli kapsamlı biçimde tartışılmaktadır. Bu yönleriyle tez, pediatrik kalp üfürümlerinin otomatik tespitine yönelik güvenilir, genellenebilir ve klinik açıdan anlamlı bilgisayar destekli tanı sistemlerinin geliştirilmesine katkı sunmayı amaçlamaktadır.

1.6. Tezin Genel Amacı

Bu çalışmanın amacı, fonokardiyogram (FKG) kayıtlarından pediatrik kalp üfürümlerinin otomatik olarak tespit edilmesi için Mel-spektrogram temsilleri ile geleneksel ses özelliklerini entegre eden hibrit bir CNN–LSTM çerçevesi tasarlamak, uygulamak ve değerlendirmektir. Ayrıca önerilen modelin performansı, geleneksel makine öğrenimi yöntemleri ve alternatif derin öğrenme modelleri ile sistematik olarak karşılaştırılmaktadır (Goodfellow vd., 2016).

Bu amaç, pediatrik uygulanabilirlik ve sağlam bir deneysel tasarım vurgusuyla hem metodolojik yeniliği hem de pratik değerlendirmeyi kapsamaktadır. Model

performansının nesnel olarak değerlendirilebilmesi için sınıflandırma doğruluğu aşağıdaki şekilde tanımlanmaktadır (Eşitlik 4):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (\text{Eşitlik 4})$$

Bu eşitlikte TP doğru pozitif, TN doğru negatif, FP yanlış pozitif ve FN yanlış negatif sınıflandırmaları ifade etmektedir.

1.7. Araştırmanın Amaçları ve Özel Hedefleri

Bu tezin temel amacı, fonokardiyogram (FKG) kayıtlarından pediatrik kalp üfürümlerinin otomatik olarak tespit edilmesine yönelik metodolojik açıdan titiz, karşılaştırmalı ve klinik açıdan anlamlı bir derin öğrenme çerçevesi geliştirmektir. Bu genel amaca ulaşmak için çalışma, veri hazırlama sürecinden model tasarımına ve performans değerlendirme aşamasına kadar uzanan çok aşamalı ve sistematik bir yaklaşım benimsemektedir (Goodfellow vd., 2016).

Bu kapsamda tez, aşağıda ayrıntılı olarak açıklanan özel hedefleri sistematik biçimde ele almaktadır.

1.7.1. Veri Setinin Anlaşılması ve Hazırlanması

Çalışmanın ilk aşamasında, CirCor DigiScope pediatrik FKG veri setinin kapsamlı biçimde anlaşılması ve analiz edilmesi hedeflenmektedir (Oliveira vd., 2022). Bu doğrultuda veri setinde yer alan üfürüm etiketlerinin dağılımı, farklı oskültasyon bölgelerinden elde edilen kayıtların özellikleri, kayıt sürelerindeki değişkenlik ve sınıf dengesizliği gibi temel istatistiksel ve yapısal özellikler ayrıntılı biçimde incelenmektedir. Bu analiz, pediatrik kalp seslerinin doğasına özgü zorlukların daha iyi anlaşılmasına ve sonraki metodolojik adımların bilinçli biçimde tasarlanmasına olanak sağlamaktadır.

Ayrıca FKG sinyallerinin güvenilir ve tekrarlanabilir biçimde analiz edilebilmesi amacıyla standart bir ön işleme boru hattı tasarlanmakta ve uygulanmaktadır. Bu boru hattı; sinyal kırpma veya segmentasyon, yeniden örnekleme, genlik normalizasyonu ve çevresel gürültü ile artefaktların azaltılmasına yönelik işlemleri içermektedir (Springer vd., 2016).

Bu bağlamda sinyalin normalize edilmesi aşağıdaki şekilde ifade edilebilir (Eşitlik 5):

$$x_{norm}(t) = \frac{x(t) - \mu}{\sigma} \quad (\text{Eşitlik 5})$$

Bu eşitlikte $x(t)$ ham FKG sinyalini, μ sinyalin ortalamasını ve σ standart sapmasını temsil etmektedir.

Böylece farklı kayıt koşullarından kaynaklanan varyasyonların en aza indirilmesi ve modellerin genelleme yeteneğinin artırılması hedeflenmektedir.

1.7.2. Özellik Çıkarma ve Seçme Süreci

Çalışmanın ikinci aşamasında, FKG kayıtlarından hem derin öğrenme hem de geleneksel makine öğrenimi yaklaşımlarına uygun temsil biçimlerinin elde edilmesi hedeflenmektedir. Bu amaçla pediatrik kalp seslerinin spektral özelliklerini yansıtacak şekilde seçilmiş parametreler kullanılarak Mel-spektrogram temsilleri hesaplanmaktadır (Logan, 2000). Örnekleme hızı, pencere boyutu, atlama uzunluğu, mel bant sayısı ve analiz edilen frekans aralığı gibi parametreler, pediatrik kalp seslerinin karakteristik yapısı dikkate alınarak belirlenmektedir.

Mel-spektrogram dönüşümü aşağıdaki şekilde ifade edilebilir (Eşitlik 6):

$$S_{mel}(t, f) = Mel(|STFT(x(t))|^2) \quad (\text{Eşitlik 6})$$

Bu eşitlikte $x(t)$ giriş sinyalini, $STFT(\cdot)$ kısa zamanlı Fourier dönüşümünü ve $Mel(\cdot)$ mel ölçeğine dönüşümü temsil etmektedir.

Buna ek olarak zaman, frekans, kepstral ve istatistiksel alanları kapsayan kapsamlı bir geleneksel ses özellikleri kümesi çıkarılmaktadır. Bu özellikler arasında genlik zarfı ölçümleri, toplam ve karekök ortalama kare (Root Mean Square – RMS) enerji değerleri, sıfır geçiş oranı (ZCR), çarpıklık ve basıklık gibi istatistiksel tanımlayıcılar, spektral merkez ve bant genişliği ölçümleri ile Mel frekans kepstral katsayıları (MFCC) yer almaktadır (Rabiner ve Schafer, 2011).

Çıkarılan bu özelliklerin sayısını azaltmak ve en ayırt edici bileşenleri belirlemek amacıyla çeşitli özellik seçimi yöntemleri uygulanmaktadır. Bu yöntemler arasında, ceza terimi aracılığıyla önemsiz değişkenleri sıfıra yaklaştıran Lasso tabanlı lojistik regresyon, grup ortalamaları arasındaki farkların istatistiksel anlamlılığını değerlendirmeye dayanan Varyans Analizi (ANOVA) kapsamında hesaplanan F-istatistiği ve yinelemeli biçimde en zayıf değişkenleri eleyerek model performansını optimize eden özyinelemeli özellik eleme teknikleri yer almaktadır.

ANOVA tabanlı F-istatistiği aşağıdaki şekilde tanımlanabilir (Eşitlik 7):

$$F = \frac{\text{Gruplar arası varyans}}{\text{Gruplar içi varyans}} \quad (\text{Eşitlik 7})$$

1.7.3. Temel (Baseline) Modellerin Geliştirilmesi ve Değerlendirilmesi

Bir sonraki aşamada, seçilen geleneksel özellikler kullanılarak çeşitli klasik makine öğrenimi modellerinin geliştirilmesi ve değerlendirilmesi amaçlanmaktadır. Bu kapsamda lojistik regresyon, destek vektör makineleri (SVM), K-en yakın komşu (KNN) algoritması, karar ağaçları, AdaBoost, Doğrusal Diskriminant Analizi (LDA) ve Kuadratik Diskriminant Analizi (QDA) gibi yöntemler eğitilmekte ve uygun hiperparametre ayarlamaları yapılmaktadır (Hastie vd., 2009).

Geliştirilen bu modeller; doğruluk, hassasiyet, geri çağırma ve F1-skoru gibi ölçütlerin yanı sıra sınıflandırma performansını eşik değerlerinden bağımsız olarak değerlendirmeye imkân tanıyan Alıcı İşletim Karakteristiği (ROC) eğrisi altında kalan alan (AUC) ve karışıklık matrisi gibi klinik açıdan anlamlı performans ölçütleri kullanılarak değerlendirilmektedir (Fawcett, 2006).

Hassasiyet (Precision) ve geri çağırma (Recall) aşağıdaki şekilde tanımlanmaktadır (Eşitlik 8 ve 9):

$$Precision = \frac{TP}{TP+FP} \quad (\text{Eşitlik 8})$$

$$Recall = \frac{TP}{TP+FN} \quad (\text{Eşitlik 9})$$

F1-skoru ise hassasiyet ve geri çağırmanın harmonik ortalaması olarak aşağıdaki şekilde hesaplanmaktadır (Eşitlik 10):

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (\text{Eşitlik 10})$$

Elde edilen sonuçlar, derin öğrenme tabanlı yaklaşımlarla yapılacak karşılaştırmalar için temel bir referans performans düzeyi oluşturmaktadır.

1.7.4. Derin Öğrenme Modellerinin Tasarımı ve Karşılaştırılması

Tezin ilerleyen aşamalarında, Mel-spektrogram temsilleri üzerinde çalışan çeşitli derin öğrenme mimarilerinin tasarlanması ve karşılaştırılması hedeflenmektedir. Bu bağlamda, öncelikle bağımsız Konvolüsyonel Sinir Ağları (CNN) ve LSTM tabanlı mimariler uygulanmaktadır (LeCun vd., 2015; Hochreiter ve Schmidhuber, 1997).

Ardından DenseNet, ResNet, MobileNet, GhostNet ve ağın derinlik, genişlik ve çözünürlük boyutlarını dengeli biçimde ölçekleyen EfficientNet gibi ileri düzey

mimariler değerlendirilmektedir (He vd., 2016; Tan ve Le, 2019). Bu mimariler, özellikle transfer öğrenme yaklaşımları ile birlikte kullanılarak pediatrik üfürüm tespitindeki performansları tutarlı bir eğitim ve değerlendirme düzeni altında karşılaştırılmaktadır.

Model eğitimi sırasında sınıflandırma problemleri için yaygın olarak kullanılan çapraz entropi kaybı aşağıdaki şekilde tanımlanmaktadır (Eşitlik 11):

$$\mathcal{L} = - \sum_{i=1}^N y_i \log(\hat{y}_i) \quad (\text{Eşitlik 11})$$

Bu eşitlikte y_i gerçek etiketleri, $\hat{y}_i$ ise model tarafından tahmin edilen olasılık değerlerini temsil etmektedir.

1.7.5. Hibrit CNN–LSTM Çerçevesinin Geliştirilmesi

Çalışmanın temel metodolojik katkılarından biri, Mel-spektrogramlardan uzamsal örüntülerin çıkarılması amacıyla çok katmanlı filtreleme işlemleri gerçekleştiren Konvolüsyonel Sinir Ağları (CNN) ile kalp döngüsü boyunca zamansal bağımlılıkları modelleyebilen ve geçmiş bilgiyi hafıza hücreleri aracılığıyla koruyabilen Uzun Kısa Süreli Bellek (Long Short-Term Memory – LSTM) ağlarının bir araya getirildiği hibrit bir mimarinin tasarlanmasıdır (LeCun vd., 2015; Hochreiter ve Schmidhuber, 1997). Bu hibrit CNN–LSTM mimarisi, kalp sesi sinyallerinin hem spektral hem de zamansal özelliklerini eş zamanlı olarak öğrenmeyi amaçlamaktadır.

Gerektiğinde, geleneksel ses özellikleri ile derin öğrenme temelli özellik temsillerinin geç füzyon yaklaşımıyla birleştirilmesi de incelenmektedir. Ayrıca eğitim sürecinde, gradyanların birinci ve ikinci moment tahminlerine dayalı olarak parametre güncellemesi yapan uyarlamalı bir optimizasyon yöntemi olan Adam (Kingma ve Ba, 2015) ile öğrenme oranını hata varyansına göre dinamik biçimde ayarlayan RMSProp (Tieleman ve Hinton, 2012) gibi optimizasyon algoritmaları kullanılmaktadır.

Bunun yanında erken durdurma (early stopping), öğrenme hızı planlama (learning rate scheduling) ve model kontrol noktaları (checkpointing) gibi düzenleme teknikleri uygulanarak modelin aşırı öğrenme eğilimi kontrol altına alınmaktadır (Goodfellow vd., 2016).

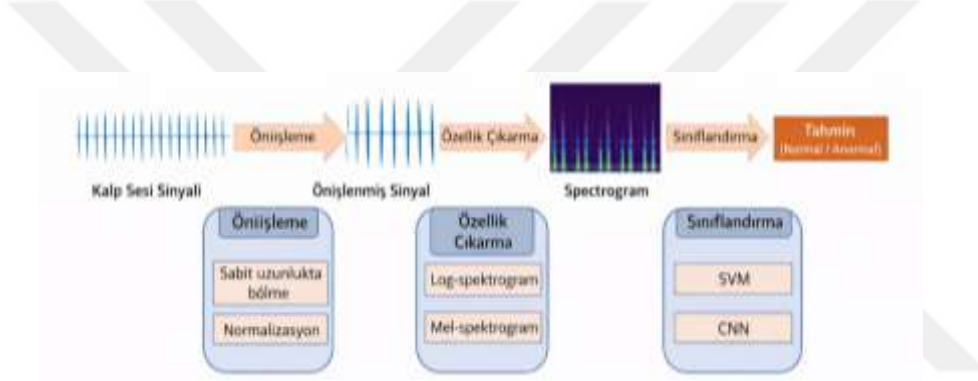
Önerilen hibrit modelin nihai sınıflandırma çıktısı, aşağıdaki softmax fonksiyonu ile elde edilmektedir (Eşitlik 12):

$$P(y = k \mid x) = \frac{e^{z_k}}{\sum_{j=1}^K e^{z_j}} \quad (\text{Eşitlik 12})$$

Bu eşitlikte z_k modelin k . sınıfı için ürettiği çıktı skorunu, K toplam sınıf sayısını ve $P(y = k | x)$ ise giriş sinyali için ilgili sınıfa ait olasılığı ifade etmektedir.

Şekil 3'te gösterildiği üzere önerilen çerçeve, farklı temsil biçimlerini bütünlleştirerek yalnızca sınıflandırma performansını değil, aynı zamanda modelin sağlamlığını ve genelleme yeteneğini de artırmayı hedeflemektedir.

Bu tezin önemli hedeflerinden biri, geliştirilen yöntemin yalnızca sınıflandırma doğruluğu açısından değil, aynı zamanda klinik açıdan anlamlı senaryolarda güvenilirlik ve genellenebilirlik bakımından da kapsamlı biçimde değerlendirilmesidir. Bu doğrultuda önerilen hibrit CNN–LSTM çerçevesinin performansı, hem geleneksel makine öğrenimi modelleri hem de farklı derin öğrenme mimarileriyle karşılaştırmalı olarak analiz edilmektedir.



Şekil 3. Mel-spektrogram girdileri üzerinde çalışan önerilen hibrit CNN–LSTM mimarisinin genel yapısı

Karşılaştırmalı değerlendirmelerin adil ve tekrarlanabilir olması amacıyla eğitim, doğrulama ve test veri bölünmeleri tüm modellerde tutarlı biçimde uygulanmakta; hasta düzeyinde ayırım sağlanarak veri sızıntısı riski en aza indirilmektedir. Böylece elde edilen performans sonuçlarının model mimarisinden kaynaklandığı daha güvenilir biçimde yorumlanabilmektedir.

Model performansı yalnızca genel ölçütlerle değerlendirilmemekte, aynı zamanda farklı üfürüm kategorilerindeki davranışın ayrıntılı olarak analiz edilmesine de odaklanılmaktadır. Bu bağlamda üfürüm varlığı, üfürüm yokluğu ve belirsiz durumlar gibi sınıflar arasındaki performans incelenmekte; sınıf dengesizliğinin modellere etkisi değerlendirilerek sağlamlık analizi gerçekleştirilmektedir. Ayrıca ROC eğrileri ve AUC değerleri üzerinden farklı karar eşiklerinde modelin duyarlılık ve özgüllük dengesi yorumlanmaktadır.

Bu yaklaşım, özellikle tarama amaçlı kullanımda yüksek duyarlılık gerektiren senaryolar ile yanlış pozitiflerin azaltılmasının kritik olduğu yüksek özgüllük odaklı senaryoların ayrı ayrı değerlendirilmesine olanak tanımaktadır.

Tezin bir diğer önemli yönü ise önerilen yöntemin klinik ve pratik uygulanabilirliğinin ele alınmasıdır. Bu kapsamda geliştirilen çerçevenin birinci basamak tarama uygulamalarına, teletıp tabanlı uzaktan değerlendirme sistemlerine ve klinik karar destek sistemlerine entegrasyon potansiyeli değerlendirilmektedir. Özellikle kaynakların sınırlı olduğu ortamlarda hesaplama maliyetleri, cihaz kısıtları, düzenleyici gereklilikler ve kullanım kolaylığı gibi faktörlerin etkisi analiz edilerek gerçek dünya kullanımına ilişkin kritik hususlar ortaya konulmaktadır.

Son olarak elde edilen bulgular doğrultusunda gelecekteki araştırma yönleri sistematik biçimde ele alınmaktadır. Bu kapsamda, model çıktılarının klinik açıdan güvenilir ve yorumlanabilir biçimde açıklanmasını sağlayacak açıklanabilir yapay zekâ yaklaşımlarının entegrasyonu önemli bir geliştirme alanı olarak öne çıkmaktadır. Bunun yanı sıra yalnızca üfürüm tespitine odaklanmak yerine, üfürümle ilişkili ek klinik çıktıları eş zamanlı olarak öğrenmeyi hedefleyen çoklu görev öğrenimi stratejileri de dikkate değer bir araştırma yönü sunmaktadır.

Gelecek çalışmalarda demografik bilgiler veya EKG gibi tamamlayıcı fizyolojik verilerin FKG sinyalleriyle birlikte kullanıldığı çoklu modlu füzyon yaklaşımlarının tanısal doğruluğu artırabileceği öngörülmektedir. Ayrıca önerilen modellerin taşınabilir ve uç (edge) cihazlarda etkin biçimde çalışabilmesi için model sıkıştırma, hesaplama verimliliğini artırma ve düşük güç tüketimine yönelik optimizasyon tekniklerinin araştırılması gerekmektedir.

Bu yönleriyle tez, yalnızca bir sınıflandırma modeli önermekle sınırlı kalmamakta; aynı zamanda önerilen yaklaşımın klinik bağlamda nasıl değerlendirilebileceğini, hangi koşullar altında uygulanabilir olduğunu ve hangi yöntemlerle geliştirilebileceğini bütüncül ve eleştirel bir bakış açısıyla ortaya koymaktadır.

1.8. Araştırma Soruları

Bu tezde belirlenen amaçlar ve oluşturulan metodolojik çerçeve doğrultusunda, çalışma boyunca yanıtlanması hedeflenen temel araştırma soruları sistematik biçimde tanımlanmaktadır. Bu araştırma soruları, geliştirilen yöntemlerin performansını, karşılaştırmalı üstünlüğünü ve klinik uygulanabilirliğini bütüncül bir bakış açısıyla değerlendirmeyi amaçlamakta; aynı zamanda tez kapsamında yürütülen deneysel analizlerin yapısını, kapsamını ve akışını yönlendiren kavramsal bir çerçeve sunmaktadır.

Bu bağlamda tez, öncelikle Mel-spektrogram temsillerinin zaman, frekans, kepsral ve istatistiksel alanlardan türetilmiş kapsamlı bir geleneksel ses özellikleri kümesiyle bütünleştirilmesi durumunda, pediatrik FKG kayıtlarında kalp üfürümü bulunan ve bulunmayan örneklerin ayırt edilmesinde ne ölçüde etkili olduğunu incelemeyi amaçlamaktadır (Logan, 2000). Bu yaklaşım, algısal olarak anlamlı zaman–frekans temsillerinin el ile tasarlanmış akustik tanımlayıcılarla birlikte kullanılmasının ayırt edici güce katkısını ortaya koymayı hedeflemektedir.

Buna ek olarak, seçilen geleneksel ses özellikleri üzerinde eğitilen klasik makine öğrenimi sınıflandırıcılarının — lojistik regresyon, destek vektör makineleri (SVM), KNN algoritması, karar ağaçları ve AdaBoost gibi yöntemler — doğruluk, F1-skoru ve ROC–AUC gibi yaygın performans ölçütleri açısından, Mel-spektrogram temsilleri kullanılarak eğitilen derin öğrenme modelleriyle karşılaştırıldığında nasıl bir performans sergilediği ayrıntılı biçimde değerlendirilmektedir (Hastie vd., 2009; Fawcett, 2006). Bu karşılaştırma, el ile tasarlanmış özelliklere dayalı yaklaşımların derin öğrenme temelli yöntemlere kıyasla güçlü ve zayıf yönlerini ortaya koymayı amaçlamaktadır.

Tezde ayrıca, FKG sinyallerinin hem uzamsal (zaman–frekans) hem de zamansal yapısını birlikte modelleyen hibrit bir CNN–LSTM çerçevesinin, CirCor DigiScope pediatrik veri seti üzerinde bağımsız CNN, bağımsız LSTM ve transfer öğrenme tabanlı mimarilere kıyasla daha üstün bir performans sunup sunmadığı araştırılmaktadır (Oliveira vd., 2022). Bu soru, kalp üfürümlerinin doğru biçimde tespit edilmesi için spektral örüntülerin yanı sıra kardiyak döngüler boyunca ortaya çıkan zamansal bağımlılıkların modellenmesinin gerekliliğini değerlendirmeyi amaçlamaktadır.

Bununla birlikte, sabit uzunluklu sinyal segmentasyonu, üfürüm pozitif örnekler için uygulanan veri artırma stratejileri ve çoklu oskültasyon bölgelerinden elde edilen FKG kayıtlarının birlikte işlenmesi gibi ön işleme tercihlerinin, üfürüm tespit modellerinin performansı ve genellenebilirliği üzerindeki etkisi de ayrıntılı biçimde ele alınmaktadır (Springer vd., 2016). Bu analiz, model başarısının yalnızca mimari tasarıma değil, aynı zamanda veri hazırlama sürecindeki metodolojik tercihlere de güçlü biçimde bağlı olduğunu göstermeyi amaçlamaktadır.

Son olarak, önerilen hibrit çerçevenin özellikle kaynakların sınırlı olduğu gerçek dünya klinik ortamlarında otomatik pediatrik üfürüm taraması için ne ölçüde pratik olarak uygulanabilir olduğu ve klinik açıdan anlamlı bir karar destek temeli sunup sunmadığı değerlendirilmektedir (Topol, 2019). Bu bağlamda geliştirilen sistemin yalnızca sayısal performans ölçütleri açısından değil, aynı zamanda klinik kullanım senaryoları açısından da potansiyel katkıları tartışılmaktadır.

Bu araştırma soruları, tez kapsamında yürütülen ampirik çalışmaları yapılandırmakta ve izleyen bölümlerde sunulan deneysel sonuçların, karşılaştırmalı analizlerin ve tartışmaların tutarlı ve sistematik bir çerçeve içinde ele alınmasını sağlamaktadır.

1.9. Çalışmanın Önemi

Bu araştırmanın önemi; klinik uygulamalar, metodolojik katkılar ve daha geniş toplumsal etkiler olmak üzere birden fazla düzeyde değerlendirilebilir. Çalışma yalnızca teknik bir sınıflandırma problemini ele almakla kalmamakta, aynı zamanda pediatrik kardiyovasküler hastalıkların erken tanısı ve sağlık hizmetlerine erişimin iyileştirilmesi gibi kritik konulara da katkı sunmayı amaçlamaktadır (WHO, 2021).

Bu bağlamda, geliştirilen yaklaşımın özellikle erken tarama süreçlerinde kullanılması, gereksiz sevklerin azaltılması ve sınırlı uzman kaynaklarının daha etkin kullanılmasına katkı sağlayabileceği öngörülmektedir. Ayrıca, otomatik ve güvenilir karar destek sistemlerinin yaygınlaşması, düşük ve orta gelirli bölgelerde sağlık hizmetlerine erişimdeki eşitsizliklerin azaltılmasına da destek olabilir.

Metodolojik açıdan ise bu çalışma, farklı veri temsillerinin bütünleştirilmesi, hibrit derin öğrenme mimarilerinin sistematik biçimde değerlendirilmesi ve tekrarlanabilir deneysel protokollerin oluşturulması bakımından literatüre katkı sunmaktadır. Bu yönleriyle tez, yalnızca mevcut yöntemlerin performansını artırmayı değil, aynı zamanda gelecekte geliştirilecek çalışmalar için sağlam bir metodolojik temel oluşturmayı hedeflemektedir.

1.9.1. Klinik Önemi

Klinik açıdan değerlendirildiğinde, önerilen hibrit CNN–LSTM çerçevesi özellikle pediatrik kardiyoloji uzmanlarına erişimin sınırlı olduğu ortamlarda geleneksel oskültasyona nesnel ve standartlaştırılmış bir destek aracı olarak kullanılma potansiyeline sahiptir (Topol, 2019). Bu tür ortamlarda yapılan oskültasyon değerlendirmeleri büyük ölçüde klinisyenin deneyimine bağlı olduğundan, otomatik ve tutarlı bir karar destek mekanizmasının tanı sürecine önemli katkılar sağlaması beklenmektedir.

Önerilen yaklaşımın, patolojik üfürümlere sahip çocukların daha erken ve doğru biçimde belirlenmesine olanak tanıyarak gecikmiş tanıya bağlı komplikasyon risklerini azaltabileceği öngörülmektedir (McCrindle vd., 2007). Bunun yanı sıra masum üfürümlerin güvenilir biçimde ayırt edilebilmesi, gereksiz sevklerin ve ekokardiyografik

incelemelerin azaltılmasına katkı sağlayarak hem aileler üzerindeki psikolojik yükü hem de sağlık sistemleri üzerindeki maliyet baskısını hafifletebilir.

Ayrıca geliştirilen çerçevenin tıp öğrencileri ve genç klinisyenler için eğitsel bir araç olarak kullanılması da mümkündür. Algoritmik değerlendirmeler ve açıklamalı örnekler aracılığıyla oskültasyon bulgularının sistematik biçimde analiz edilmesi, klinik eğitimin desteklenmesine ve dinleme becerilerinin daha tutarlı biçimde geliştirilmesine katkı sağlayabilir (Kelly vd., 2019).

1.9.2. Metodolojik Önemi

Metodolojik açıdan değerlendirildiğinde, bu tez pediatrik FKG sinyallerinin analizi için titiz, şeffaf ve tekrarlanabilir bir çalışma çerçevesi sunmaktadır. Ön işleme, segmentasyon, veri artırma ve veri bölme stratejilerinin açık biçimde belgelenmesi, literatürde sıklıkla karşılaşılan metodolojik belirsizliklerin azaltılmasına katkı sağlamaktadır (Varoquaux vd., 2017). Bu yaklaşım, elde edilen sonuçların güvenilirliğini artırmakta ve farklı çalışmalar arasında daha adil karşılaştırmalar yapılmasını kolaylaştırmaktadır.

Çalışma ayrıca Mel-spektrogram temsillerinin geleneksel ses özellikleriyle bütünleşik bir çerçevede ele alınmasının sağladığı tamamlayıcı değeri ortaya koymaktadır. Bu entegrasyon, hem derin öğrenme tabanlı modellerin yüksek temsil gücünden hem de geleneksel özelliklerin yorumlanabilirliğinden yararlanmayı mümkün kılmaktadır (Logan, 2000). Bununla birlikte klasik makine öğrenimi yöntemlerinden modern derin ve hibrit mimarilere kadar geniş bir model yelpazesinin aynı veri seti üzerinde ve tutarlı değerlendirme ölçütleri kullanılarak karşılaştırılması, literatürde nadir görülen kapsamlı bir analiz sunmaktadır.

Önerilen hibrit CNN–LSTM mimarisi için ele alınan tasarım ilkeleri yalnızca kalp sesi analiziyle sınırlı kalmayıp, solunum sesi analizi, uyku evrelerinin sınıflandırılması ve diğer biyomedikal ses işleme uygulamalarına da genellenebilir nitelik taşımaktadır (LeCun vd., 2015).

1.9.3. Toplumsal ve Sağlıkta Eşitlik Açısından Önemi

Bu araştırmanın toplumsal önemi, sağlık hizmetlerine eşit erişimin desteklenmesi ve dijital sağlık çözümlerinin yaygınlaştırılması bağlamında değerlendirilebilir. Birinci basamak sağlık kliniklerinde, toplum sağlığı merkezlerinde ve uzak bölgelerde kullanılabilecek otomatik tanı destek araçlarının geliştirilmesi, yüksek kaliteli sağlık hizmetlerine erişimdeki eşitsizliklerin azaltılmasına katkı sağlayabilir (WHO, 2021).

Ayrıca pandemi koşulları, sağlık personeli yetersizliği ve yaşlanan nüfus gibi küresel zorluklar göz önünde bulundurulduğunda teletıp ve dijital sağlık uygulamalarının önemi giderek artmaktadır (Topol, 2019). Bu tez kapsamında geliştirilen yaklaşım, uzaktan değerlendirme ve tarama senaryolarında kullanılabilecek yapısıyla bu tür girişimleri destekleyici bir rol üstlenmektedir. Özellikle ileri görüntüleme olanaklarının sınırlı olduğu sağlık sistemlerinde, kaynakların daha etkin ve maliyet açısından verimli biçimde kullanılmasına katkı sağlaması beklenmektedir.

Pediyatrik popülasyona odaklanan ve ölçeklenebilir teknolojilere entegre edilebilecek bir çözüm sunan bu çalışma, uzun vadede çocuklarda kardiyovasküler hastalıklara bağlı morbidite ve mortalitenin azaltılmasına yönelik küresel çabalara katkı sunmayı amaçlamaktadır.

1.10. Tezin Katkıları

Bu tez, pediyatrik kalp üfürümlerinin otomatik tespitine yönelik literatüre hem metodolojik hem de uygulamaya dönük çok boyutlu katkılar sunmaktadır. Çalışma yalnızca yeni bir model önermekle sınırlı kalmamakta; veri hazırlama sürecinden performans değerlendirmesine ve klinik yorumlamaya kadar uzanan bütüncül ve sistematik bir araştırma çerçevesi ortaya koymaktadır.

Bu kapsamda tez, veri ön işleme, özellik çıkarımı, model tasarımı ve değerlendirme aşamalarını kapsayan entegre bir metodolojik yaklaşım sunarak, literatürde sıklıkla ayrı ayrı ele alınan bileşenleri tek bir tutarlı yapı altında birleştirmektedir.

Ayrıca çalışma, geleneksel makine öğrenimi yöntemleri ile modern derin öğrenme ve hibrit mimarilerin aynı veri seti ve ortak değerlendirme ölçütleri altında karşılaştırılmasını sağlayarak, farklı yaklaşımların güçlü ve zayıf yönlerinin sistematik biçimde ortaya konulmasına olanak tanımaktadır.

Bu bağlamda tezin başlıca katkıları aşağıda özetlenmektedir.

1.10.1. Pediyatri Odaklı Kapsamlı FKG Analiz Süreci

Tezin temel katkılarından biri, CirCor DigiScope pediyatrik FKG veri setine özgü olarak tasarlanmış kapsamlı, şeffaf ve tekrarlanabilir bir analiz sürecinin geliştirilmesidir. Bu süreç, hasta seçimi ilkelerinden başlayarak çoklu oskültasyon bölgelerinden elde edilen sinyallerin bütüncül biçimde ele alınmasına kadar uzanmaktadır (Oliveira vd., 2022).

Sabit uzunluklu segmentasyon yaklaşımı (örneğin 5 saniyelik pencereler), yeniden örnekleme ve genlik normalizasyonu gibi ön işleme adımları sistematik biçimde

uygulanarak farklı kayıt koşullarından kaynaklanan varyasyonların etkisi azaltılmaktadır. Bu yapı, sinyallerin tutarlı biçimde temsil edilmesini sağlayarak model performansının daha güvenilir biçimde değerlendirilmesine katkıda bulunmaktadır.

Özellikle üfürüm pozitif vakalar için hedefe yönelik veri artırma stratejilerinin kullanılması, sınıf dengesizliğinin etkilerini hafifletmeyi ve modellerin nadir görülen patolojik örüntülere karşı duyarlılığını artırmayı amaçlamaktadır. Ayrıca eğitim, doğrulama ve test veri kümeleri arasında hasta düzeyinde ayırım sağlanması, veri sızıntısı riskini en aza indirerek elde edilen sonuçların genellenebilirliği konusunda daha gerçekçi değerlendirmeler yapılmasına olanak tanımaktadır (Varoquaux vd., 2017).

Bu yönüyle çalışma, pediatrik FKG analizine yönelik tekrarlanabilir ve metodolojik açıdan sağlam bir referans iş akışı sunmaktadır.

1.10.2. Kalp Üfürümü Tespiti için Zengin ve Hibrit Özellik Temsili

Bu tez, pediatrik kalp üfürümlerinin ayırt edilmesi için zengin ve çok boyutlu bir özellik temsili yaklaşımı sunmaktadır. Bu kapsamda Mel-spektrogram temsilleri, pediatrik kalp seslerinin spektral özelliklerini en iyi şekilde yansıtacak biçimde özenle ayarlanmış parametreler kullanılarak hesaplanmaktadır (Logan, 2000). Bununla birlikte çalışma yalnızca derin öğrenme temsillerine odaklanmamakta; zaman, frekans, kepsral ve istatistiksel alanları kapsayan geniş bir geleneksel ses özellikleri kümesi de sistematik biçimde çıkarılmaktadır.

Bu kapsamlı özellik seti; genlik zarfı ölçümleri, toplam ve RMS enerji değerleri, ZCR, çarpıklık ve basıklık gibi istatistiksel tanımlayıcılar, tepe frekansı ve başlangıçla ilişkili özellikler, bant enerji oranları, otokorelasyon tabanlı ölçümler, spektral merkez ve bant genişliği ile Mel frekans kepsral katsayılarını (MFCC) içermektedir.

Ayrıca Lasso tabanlı yöntemler, ANOVA F-değeri ve özyinelemeli özellik eleme gibi özellik seçimi teknikleri uygulanarak geleneksel sınıflandırıcılar için en ayırt edici özelliklerin belirlenmesi sağlanmaktadır (Hastie vd., 2009). Bu yaklaşım, gereksiz veya düşük katkı sağlayan değişkenlerin elenmesine olanak tanırken, model performansını artırmakta ve aynı zamanda elde edilen sonuçların yorumlanabilirliğini güçlendirmektedir.

Bu yönüyle önerilen çerçeve, derin öğrenme temsillerinin yüksek ayırt edici gücü ile geleneksel özelliklerin açıklanabilirliğini bir araya getirerek dengeli ve kapsamlı bir özellik temsili sunmaktadır.

1.10.3. Geleneksel ve Derin Öğrenme Modellerinin Sistemik Karşılaştırılması

Tezin önemli katkılarından biri, geleneksel makine öğrenimi yöntemleri ile modern derin öğrenme yaklaşımlarının aynı pediatrik veri seti üzerinde ve tutarlı bir deneysel protokol çerçevesinde sistematik biçimde karşılaştırılmasıdır. Bu doğrultuda lojistik regresyon, destek vektör makineleri (SVM), K-en yakın komşu (KNN) algoritması, karar ağaçları, AdaBoost ile doğrusal ve kuadratik diskriminant analizleri gibi klasik yöntemler uygulanmakta ve uygun hiperparametre ayarlamaları gerçekleştirilmektedir (Hastie vd., 2009).

Buna ek olarak, Mel-spektrogram girdileri üzerinde çalışan temel CNN ve LSTM mimarileri ile DenseNet, EfficientNet, ResNet, MobileNet, GhostNet ve dikkat tabanlı CNN mimarileri gibi gelişmiş ve transfer öğrenme temelli modeller değerlendirilmektedir (He vd., 2016; Tan ve Le, 2019).

Tüm bu modeller; doğruluk, hassasiyet, geri çağırma, F1-skoru, ROC–AUC ve karışıklık matrisi gibi klinik açıdan anlamlı performans ölçütleri kullanılarak karşılaştırılmakta; böylece farklı yaklaşımlar arasındaki performans–karışıklık ödünleşimleri ayrıntılı biçimde analiz edilmektedir (Fawcett, 2006).

Bu kapsamlı karşılaştırma, yalnızca en yüksek doğruluğu sağlayan modeli belirlemekle kalmamakta; aynı zamanda farklı yöntemlerin güçlü ve zayıf yönlerini ortaya koyarak belirli klinik kullanım senaryoları için en uygun yaklaşımın seçilmesine olanak tanımaktadır.

1.10.4. Hibrit CNN–LSTM Çerçevesinin Tasarımı ve Değerlendirilmesi

Bu çalışmanın özgün katkılarından biri, pediatrik kalp üfürümü tespitine özel olarak tasarlanmış hibrit bir CNN–LSTM mimarisinin geliştirilmesidir. Önerilen mimaride evrişimli katmanlar, Mel-spektrogram temsillerinden uzamsal ve spektral örüntüleri öğrenirken; LSTM katmanları kalp döngüleri boyunca ortaya çıkan zamansal bağımlılıkları modellemektedir (LeCun vd., 2015; Hochreiter ve Schmidhuber, 1997). Bu sayede FKG sinyallerinin hem zaman–frekans hem de zamansal yapısı bütüncül biçimde değerlendirilebilmektedir.

Ayrıca, uygun aşamalarda geleneksel ses özellik vektörlerinin derin öğrenme temelli temsillerle geç füzyon (late fusion) yaklaşımıyla birleştirilmesi incelenerek tamamlayıcı bilgilerin etkin biçimde kullanılması ve sınıflandırma sağlamlığının artırılması hedeflenmektedir. Bu entegrasyon, farklı temsil türlerinin güçlü yönlerini bir araya getirerek modelin genelleme yeteneğini desteklemektedir.

Deneyisel sonuçlar, önerilen hibrit CNN–LSTM çerçevesinin aynı pediatrik veri seti üzerinde değerlendirilen geleneksel makine öğrenimi yöntemlerine ve çeşitli derin transfer öğrenimi mimarilerine kıyasla daha yüksek F1-skoru ve ROC–AUC değerleri elde ettiğini göstermektedir. Bu bulgular, hem spektral hem de zamansal özelliklerin birlikte modellenmesinin pediatrik kalp üfürümü tespitinde kritik bir avantaj sağladığını ortaya koymaktadır.

1.10.5. Klinik Odaklı Değerlendirme ve Yorumlama

Tezin bir diğer önemli katkısı, model performansının yalnızca genel doğruluk değerleri üzerinden değil, aynı zamanda klinik kullanım senaryoları bağlamında yorumlanmasıdır. Bu kapsamda, sınıf dengesizliği koşullarında duyarlılık ve özgüllük arasındaki dengenin önemi vurgulanmakta; tarama amaçlı kullanımda yanlış negatif sonuçların (FN) olası klinik etkileri özellikle dikkate alınmaktadır (Fawcett, 2006).

Ayrıca model çıktılarının ve güven puanlarının klinik karar verme süreçlerinde nasıl kullanılabileceği de sistematik biçimde ele alınmaktadır. Bu bağlamda geliştirilen yaklaşımın, birinci basamak sağlık hizmetlerinde triyaj aracı olarak veya teletıp ortamlarında ikinci görüş mekanizması olarak nasıl konumlandırılabileceği tartışılmaktadır (Topol, 2019).

Bunun yanında hesaplama karmaşıklığı, dijital stetoskoplar veya mobil uygulamalarla entegrasyon potansiyeli, model yorumlanabilirliği ve klinisyen güvenini artırmaya yönelik hususlar da kapsamlı biçimde değerlendirilmiştir. Bu değerlendirme, önerilen yöntemin yalnızca deneysel ortamda değil, aynı zamanda gerçek dünya koşullarında uygulanabilirliğini de ortaya koymaktadır.

Bu katkılar bütüncül olarak değerlendirildiğinde, tez pediatrik kalp üfürümü tespiti alanındaki mevcut çalışmaları ileriye taşımakta ve yapay zekâ destekli dijital oskültasyon sistemleri için sağlam, şeffaf ve genellenebilir bir metodolojik temel sunmaktadır. Bu yönüyle çalışma, hem akademik araştırmalar hem de gerçek dünya klinik uygulamaları açısından değerli bir referans niteliği taşımaktadır.

1.11. Kapsam ve Sınırlamalar

Bu bölümde çalışmanın kapsamı açık ve sistematik bir biçimde tanımlanmakta ve elde edilen bulguların hangi sınırlar içinde değerlendirilmesi gerektiği eleştirel bir bakış açısıyla tartışılmaktadır. Araştırma kapsamında yapılan metodolojik tercihler, kullanılan veri kaynakları ve problem tanımı doğrultusunda belirli varsayımlar yapılmış ve bazı sınırlamalar kabul edilmiştir.

Bu sınırlamaların açık ve şeffaf biçimde ifade edilmesi, tezin katkılarının daha gerçekçi bir çerçevede yorumlanmasına olanak tanımakta ve elde edilen bulguların genellenebilirliğine ilişkin sınırların net biçimde ortaya konulmasını sağlamaktadır.

Ayrıca bu yaklaşım, gelecekte gerçekleştirilecek çalışmalar için geliştirilmesi gereken yönlerin belirlenmesine katkı sunarak, literatürde daha sağlam ve karşılaştırılabilir araştırmaların yapılmasına zemin hazırlamaktadır.

1.11.1. Kapsam

Bu çalışmanın nüfus odak noktası, CirCor DigiScope veri setinde yer alan pediatrik hastalarla sınırlıdır. İncelenen örnekler doğumdan 21 yaşına kadar olan bireyleri kapsamaktadır. Bu nedenle elde edilen bulgular, yetişkin popülasyona veya veri setinde temsil edilmeyen farklı demografik gruplara doğrudan genellenebilir nitelikte değildir. Pediatrik kalp seslerinin fizyolojik ve akustik özelliklerinin yetişkinlere kıyasla anlamlı biçimde farklılık göstermesi, bu sınırlamanın metodolojik açıdan bilinçli bir tercih olduğunu ortaya koymaktadır.

Çalışmada ele alınan sinyal modalitesi, standart oskültasyon bölgelerinden elde edilen FKG kayıtları ile sınırlandırılmıştır. Elektrokardiyografi (EKG) sinyalleri, ekokardiyografik görüntüler veya klinik laboratuvar verileri gibi ek modaliteler analiz kapsamına dahil edilmemiştir. Bu tercih, çalışmanın odak noktasını dijital oskültasyon ve ses temelli tanı yaklaşımları üzerinde yoğunlaştırmayı amaçlamaktadır. Çoklu modal veri entegrasyonu, bu tezde önerilen çerçevenin ötesinde olup gelecekteki çalışmalar için doğal bir genişleme alanı sunmaktadır.

Görev tanımı açısından bakıldığında, çalışmanın birincil amacı pediatrik FKG kayıtlarında kalp üfürümlerinin varlığını veya yokluğunu tespit etmektir. Uygun olduğu durumlarda “bilinmeyen” kategorisi de değerlendirmeye dahil edilmiştir. Bununla birlikte üfürüm yoğunluğunun ayrıntılı derecelendirilmesi, spesifik etiyolojik tanıların belirlenmesi veya uzun dönem klinik sonuçların tahmin edilmesi bu çalışmanın kapsamı dışında bırakılmıştır. Bu sınırlama, otomatik tarama ve erken uyarı sistemlerine odaklanan bir yaklaşımın benimsenmesiyle uyumludur.

Tüm deneysel analizler CirCor DigiScope veri seti üzerinde gerçekleştirilmiştir. Veri seti, ayrıntılı üfürüm açıklamaları ve çoklu oskültasyon bölgelerinden elde edilen kayıtlar içermesi bakımından önemli avantajlar sunmakla birlikte, belirli bir edinim ortamını, kayıt cihazlarını ve klinik bağlamı temsil etmektedir. Bu durum, elde edilen bulguların söz konusu bağlam çerçevesinde yorumlanmasını gerekli kılmaktadır (Oliveira vd., 2022).

1.11.2. Sınırlamalar

Bu çalışmada elde edilen bulgular, kullanılan veri seti ve benimsenen metodolojik tercihlere bağlı olarak bazı sınırlamalar içermektedir. Öncelikle, CirCor DigiScope veri seti üzerinde eğitilen modellerin farklı hastanelerden, farklı kayıt cihazlarından veya farklı popülasyonlardan elde edilen FKG kayıtlarına doğrudan genellenebilirliği sınırlı olabilir. Bu tür senaryolarda modele özgü uyarlama süreçleri veya ek ince ayar (fine-tuning) işlemlerine ihtiyaç duyulması muhtemeldir (Oliveira vd., 2022).

Bir diğer önemli sınırlama, veri setinde gözlemlenen sınıf dengesizliğidir. Normal kayıtların üfürüm içeren kayıtlara kıyasla daha fazla olması, özellikle nadir görülen veya sınırdaki üfürüm vakalarının model tarafından doğru biçimde öğrenilmesini zorlaştırabilmektedir. Veri artırma stratejileri uygulanmış ve hasta düzeyinde dikkatli veri bölme yöntemleri kullanılmış olsa da, bu dengesizliğin model performansı üzerindeki etkileri tamamen ortadan kaldırılamamıştır.

Yorumlanabilirlik açısından değerlendirildiğinde, geleneksel ses özelliklerinin kullanımı ve bazı mimari tasarım tercihleri model çıktılarının anlaşılabilirliğini belirli ölçüde artırmaktadır. Bununla birlikte, hibrit CNN–LSTM mimarileri de dahil olmak üzere derin öğrenme tabanlı yaklaşımlar, daha basit doğrusal modellere kıyasla doğası gereği daha karmaşık yapılar sunmaktadır. Grad-CAM, dikkat haritaları veya özellik önem analizi gibi gelişmiş açıklanabilirlik yöntemleri bu tezin ana kapsamı dışında tutulmuş ve gelecekteki çalışmalar için önerilmiştir (Selvaraju vd., 2017).

Son olarak, gerçek dünya klinik uygulamalarına geçiş sürecinde dikkate alınması gereken düzenleyici onaylar, elektronik sağlık kayıtlarıyla entegrasyon, kullanıcı arayüzü tasarımı ve klinisyen eğitimi gibi faktörler bu tezde kavramsal düzeyde ele alınmakta, ancak deneysel olarak incelenmemektedir. Bu çalışma teknik ve metodolojik bir temel sunmakta olup, tam klinik uygulamaya geçiş için disiplinler arası iş birliği ve ek doğrulama çalışmalarının gerekli olduğu kabul edilmektedir (Topol, 2019).

Bu sınırlamaların açık ve sistematik biçimde ifade edilmesi, tezin katkılarının daha sağlıklı bir çerçevede değerlendirilmesine olanak tanımakta ve gelecekte yapılacak araştırmaların hangi yönleri odaklanabileceğini netleştirmektedir.

1.12. Tezin yapısı

Bu tez, kalp üfürümlerinin FKG kayıtlarından otomatik olarak tespit edilmesine yönelik literatürü, önerilen yöntemi ve elde edilen bulguları sistematik ve bütüncül bir çerçevede sunacak şekilde yapılandırılmıştır.

Bölüm 2 (Literatür İncelemesi), kalp sesi analizi ve otomatik üfürüm tespiti alanındaki önceki çalışmaları kapsamlı biçimde ele almaktadır. Bu bölüm, kalp sesleri ve üfürümlerle ilişkili kardiyovasküler fizyolojiye genel bir bakışla başlamakta ve fonokardiyografinin bir tanı aracı olarak rolünü tartışmaktadır. Devamında zaman alanı, frekans alanı ve zaman–frekans tabanlı yaklaşımlar ile klasik makine öğrenimi yöntemleri incelenmektedir. Son olarak CNN, Tekrarlayan Sinir Ağları (RNN), hibrit mimariler, dikkat mekanizmaları ve PhysioNet/CinC yarışmaları kapsamında geliştirilen güncel derin öğrenme tabanlı sistemler ele alınarak literatürdeki araştırma boşlukları vurgulanmaktadır.

Bölüm 3 (Malzemeler ve Yöntemler), bu tezde benimsenen metodolojiyi ayrıntılı biçimde açıklamaktadır. Bu bölümde CirCor DigiScope pediatrik kalp sesi veri setinin yapısı, anotasyonları, oskültasyon bölgeleri ve etiket dağılımları tanıtılmaktadır. Ardından FKG ön işleme adımları (yeniden örnekleme, segmentasyon, normalizasyon ve veri artırma) sunulmakta; Mel-spektrogramların ve geleneksel ses özelliklerinin çıkarımı ile özellik seçimi teknikleri açıklanmaktadır. Bölüm, önerilen hibrit CNN–LSTM çerçevesinin mimarisi, eğitim süreci ve değerlendirme ölçütlerinin tanıtılmasıyla tamamlanmaktadır.

Bölüm 4 (Deney Sonuçları), gerçekleştirilen deneylerin bulgularını sunmaktadır. Bölüm, geleneksel makine öğrenimi modellerinin temel performanslarıyla başlamakta; ardından CNN, LSTM ve transfer öğrenimi tabanlı mimarilerin Mel-spektrogram girdileri üzerindeki sonuçlarını raporlamaktadır. Bu bölümün temel odağı, hibrit CNN–LSTM modelinin doğruluk, hassasiyet, geri çağırma, F1-skoru ve ROC–AUC gibi nicel ölçütler ile karışıklık matrisleri üzerinden kapsamlı biçimde değerlendirilmesidir. Karşılaştırmalı analizlerde klinik açıdan anlamlı performans ödünleşimlerine özel vurgu yapılmaktadır.

Bölüm 5 (Tartışma), elde edilen nicel sonuçları araştırma soruları ve mevcut literatür bağlamında yorumlamaktadır. Bu bölümde sınıf dengesizliği koşullarında model davranışı, kayıtlar arası değişkenliğe karşı dayanıklılık ve olası klinik uygulama senaryoları ele alınmaktadır. Ayrıca Mel-spektrogramların ve geleneksel ses özelliklerinin model performansına katkıları ayrıntılı biçimde tartışılmaktadır.

Bölüm 6 (Sonuçlar ve Gelecek Çalışmalar), tezin temel katkılarını özetlemekte ve araştırma amaçları doğrultusunda elde edilen bulguları yeniden değerlendirmektedir. Bu bölümde ayrıca model açıklanabilirliği, çoklu görev ve çoklu modal yaklaşımlar, farklı veri kümelerine alan adaptasyonu ve klinik ile teletıp ortamlarında potansiyel uygulamalar gibi gelecekteki araştırma yönleri sunulmaktadır.

Bu yapı sayesinde tez, okuyucuyu klinik ve teknolojik arka plandan başlayarak (Bölüm 1), mevcut literatür ve açık araştırma sorunları üzerinden (Bölüm 2), önerilen yöntem ve deneysel değerlendirmelere (Bölüm 3 ve 4) yönlendirmekte; eleştirel tartışma ve gelecek perspektifi ile (Bölüm 5 ve 6) bütüncül bir sonuca ulaşmaktadır.



2. LİTERATÜR TARAMASI

2.1. Giriş

FKG kayıtlarından kalp üfürümlerinin otomatik olarak tespit edilmesi, son yıllarda biyomedikal sinyal işleme ve yapay zekâ alanlarında yoğun ilgi gören araştırma konularından biri hâline gelmiştir. Türbülanslı veya anormal kan akışının akustik bir yansıması olan kalp üfürümleri, özellikle pediatrik popülasyonda konjenital kalp hastalıklarının (CHD) erken göstergeleri arasında yer almakta ve kardiyovasküler morbiditenin önemli nedenlerinden biri olmaya devam etmektedir (Reyna vd., 2023; McCrindle vd., 2007).

Geleneksel oskültasyon klinik pratikte yaygın olarak kullanılmasına rağmen, klinisyenin deneyimine bağımlı olması, çevresel gürültüye duyarlılığı ve işitsel yorumlamanın öznel doğası nedeniyle çeşitli sınırlamalar içermektedir. Bu durum, özellikle kaynakların kısıtlı olduğu sağlık ortamlarında kalp üfürümlerinin güvenilir ve tutarlı biçimde taranmasını zorlaştırmaktadır. Bu bağlamda, objektif ölçütlere dayanan ve klinik karar sürecini destekleyebilen otomatik üfürüm tespit sistemlerine yönelik ilgi giderek artmaktadır (Thompson vd., 2019).

Son on yılda dijital stetoskop teknolojilerinde ve FKG veri toplama sistemlerinde kaydedilen ilerlemeler, daha geniş bant genişliği, yüksek örnekleme doğruluğu ve azaltılmış gürültü seviyeleri sayesinde daha kaliteli kalp sesi kayıtlarının elde edilmesini mümkün kılmıştır. Bu gelişmeler, araştırmacıların el yapımı özelliklere dayalı yaklaşımlardan derin öğrenme tabanlı uçtan uca modellere geçişini hızlandırmıştır (Araujo vd., 2022; LeCun vd., 2015).

FKG sinyalleri, durağan olmayan yapıları ve bireyler arası değişkenlikleri nedeniyle analiz edilmesi zor biyomedikal sinyaller arasında yer almaktadır. Bu nedenle literatürde, zaman alanı, frekans alanı ve zaman–frekans temsillerinin birlikte kullanılması önerilmektedir. Özellikle derin öğrenme tabanlı zaman–frekans analizlerinin, kalp seslerindeki ayırt edici örüntüleri daha etkin biçimde yakalayabildiği gösterilmiştir (Chen vd., 2023).

Bu alandaki gelişmeler, özellikle George B. Moody PhysioNet Challenge 2022 ve CirCor DigiScope gibi büyük ölçekli veri setlerinin erişilebilir hâle gelmesiyle hız kazanmıştır. Bu veri kümeleri, farklı algoritmaların standart koşullar altında

karşılaştırılmasına olanak sağlayarak literatürdeki ilerlemeyi önemli ölçüde hızlandırmıştır (Oliveira vd., 2022; Reyna vd., 2023).

Son yıllarda evrimsel sinir ağları (CNN), tekrarlayan sinir ağları (özellikle LSTM) ve hibrit CNN–LSTM mimarileri, FKG sinyallerindeki spektral ve zamansal bağımlılıkları birlikte modelleyebilme yetenekleri sayesinde dikkat çekici performans artışları göstermiştir (Yaseen vd., 2018; Kiranyaz vd., 2021).

Buna karşın hasta anatomisindeki bireysel farklılıklar, oskültasyon noktalarının çeşitliliği, arka plan gürültüsü ve kullanılan stetoskop donanımındaki değişkenlik gibi faktörler, genellenebilir otomatik üfürüm tespit sistemlerinin geliştirilmesini zorlaştıran temel problemler arasında yer almaktadır (Kelly vd., 2019).

2.2. Kardiyovasküler Fizyoloji ve Kalp Seslerinin Akustik Temeli

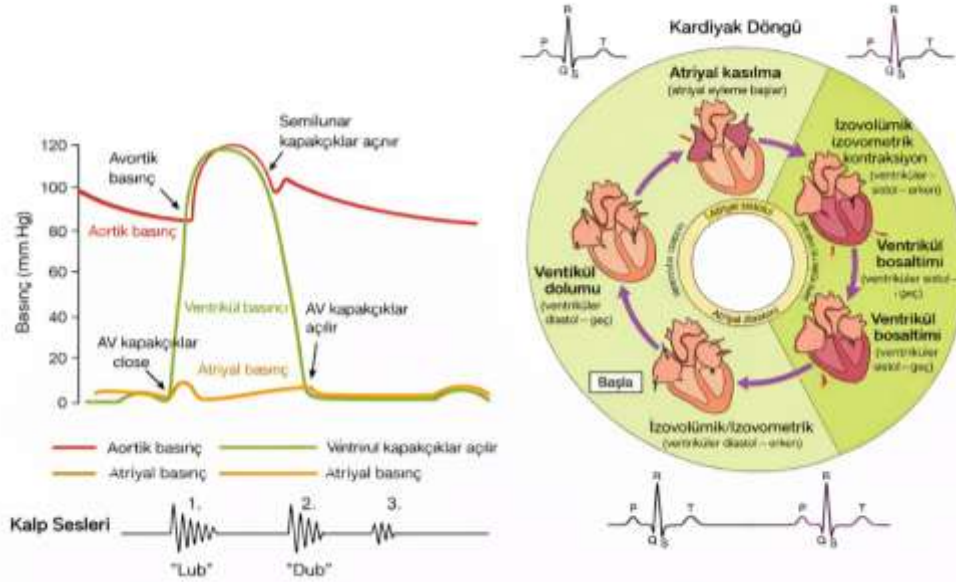
Kalp seslerinin fizyolojik ve akustik kökenlerinin anlaşılması, FKG kayıtlarının doğru biçimde yorumlanabilmesi ve normal kardiyak aktivitenin patolojik üfürümlerden ayırt edilebilmesi açısından temel bir gerekliliktir. Kalp sesleri, kalp döngüsü sırasında meydana gelen mekanik olayların bir sonucu olarak ortaya çıkmakta olup başlıca kapakçık hareketleri, kan akışındaki ani hız değişimleri ve kardiyak yapıların titreşimleri ile ilişkilidir (Springer vd., 2016; McGee, 2018).

Normal bir kalp döngüsünde en belirgin iki akustik olay birinci (S1) ve ikinci (S2) kalp sesleridir. Birinci kalp sesi (S1), sistolün başlangıcında mitral ve triküspit kapakların kapanmasıyla oluşur. Bu süreç, ventriküllerin izovolümetrik kasılması sırasında kapakçıkların gerilmesine ve çevre dokuların titreşmesine neden olarak görece düşük frekanslı (yaklaşık 20–150 Hz) ve daha uzun süreli bir akustik yapı üretir. S1'in genliği ve süresi; ventriküler kontraktile, kapakçık özellikleri ve atriyoventriküler basınç farkları gibi fizyolojik değişkenlerden etkilenmektedir (McGee, 2018; Leatham, 1991).

İkinci kalp sesi (S2) ise diyastolün başlangıcında, intraventriküler basıncın hızla düşmesini takiben aort ve pulmoner kapakların kapanmasıyla meydana gelir. S2, S1'e kıyasla daha kısa süreli ve daha yüksek frekans bileşenleri içeren keskin bir akustik olay olarak tanımlanır. Aort ve pulmoner kapakların kapanma zamanlamasındaki küçük farklılıklar, özellikle pediatrik popülasyonda fizyolojik olarak duyulabilen S2 ayrışmasına (splitting) yol açabilmektedir. Bu ayrışma çoğunlukla solunumla değişkenlik gösterir ve normal bir fizyolojik bulgu olarak kabul edilir (McCrindle vd., 2007).

Bu temel mekanik olaylar, FKG tabanlı analizlerde ayırt edici zaman–frekans imzaları üretmektedir. S1 ve S2 kalp sesleri, zaman–frekans düzleminde farklı süre ve spektral dağılımlar sergileyerek kalp döngüsünün sistolik ve diyastolik fazlarının açık

biçimde ayrılmasını mümkün kılmaktadır. S1'in daha geniş bantlı ve uzun süreli enerji dağılımına karşılık, S2 daha kısa süreli ve nispeten daha yüksek frekans bileşenleri içeren keskin bir yapı ile karakterizedir. Bu farklılıklar, otomatik segmentasyon algoritmaları ve özellik çıkarım yöntemleri için kritik referans noktaları sağlamaktadır (Springer vd., 2016).



Şekil 4. Kalp döngüsü sırasında birinci (S1) ve ikinci (S2) kalp seslerinin oluşumu

S1 ve S2 kalp seslerinin ötesinde, belirli fizyolojik ve patolojik durumlarda üçüncü (S3) ve dördüncü (S4) kalp sesleri de ortaya çıkabilmektedir. S3 kalp sesi, diyastolün erken fazında hızlı ventriküler dolum sırasında meydana gelir ve özellikle çocuklarda ve genç bireylerde fizyolojik bir bulgu olarak kabul edilebilir. Bununla birlikte, erişkinlerde S3 varlığı sıklıkla kalp yetmezliği ve volüm yüklenmesi gibi patolojik durumlarla ilişkilendirilmektedir (McGee, 2018).

S4 kalp sesi ise atriyal kontraksiyonun, uyumu azalmış veya sertleşmiş bir ventriküle karşı gerçekleşmesi sonucu oluşur ve genellikle patolojik bir bulgu olarak değerlendirilir. S4, özellikle ventriküler hipertrofi ve diyastolik disfonksiyon ile ilişkilendirilmektedir (McCrindle vd., 2007; McGee, 2018).

Bu ek kalp sesleri, özellikle gürültülü klinik ortamlarda, taşikardi veya aritmi içeren kayıtlarda ve üfürümlerin eşlik ettiği durumlarda FKG sinyallerinin yorumlanmasını daha karmaşık hâle getirebilmektedir. S3 ve S4'ün frekans içeriklerinin S1, S2 ve bazı üfürüm bileşenleri ile örtüşebilmesi, otomatik sınıflandırma sistemlerinde yanlış pozitif (false positive) ve yanlış negatif (false negative) sonuçların ortaya çıkma riskini artırmaktadır.

Bu durum, kalp sesi analizinde spektral ayrıştırma ve zaman–frekans temsillerinin önemini daha da belirgin hâle getirmektedir (Springer vd., 2016).

Dolayısıyla, kalp seslerinin fizyolojik temellerinin literatürde ayrıntılı biçimde ele alınması ve bu bilginin sinyal işleme ile modelleme aşamalarına entegre edilmesi, otomatik üfürüm tespit sistemlerinin doğruluğunu ve klinik güvenilirliğini artırmak açısından kritik bir gereklilik olarak değerlendirilmektedir.

2.3. Kalp Üfürümlerinin Akustik Kökeni ve Klinik Önemi

Kalp üfürümleri, kalp ve büyük damarlar içerisindeki kan akışının laminar rejimden saparak türbülanslı hâle gelmesi sonucu ortaya çıkan akustik fenomenlerdir. Bu türbülanslı akış çoğu zaman yapısal veya fonksiyonel kardiyak anormalliklerle ilişkilidir ve kapak darlıkları, kapak yetmezlikleri, septal defektler veya anormal vasküler şantlar gibi hemodinamik bozukluklara ilişkin önemli klinik ipuçları sunar. Bu nedenle kalp üfürümlerinin akustik özelliklerinin doğru biçimde tanımlanması ve yorumlanması, hem klinik değerlendirme hem de otomatik tanı sistemlerinin geliştirilmesi açısından kritik bir öneme sahiptir (McCrindle vd., 2007; McGee, 2018; Thompson vd., 2019).

Klinik pratikte kalp üfürümleri, kalp döngüsü içerisindeki zamanlamalarına göre sistolik, diyastolik veya sürekli üfürümler olarak sınıflandırılmaktadır. Bu sınıflandırma, üfürümün altında yatan patofizyolojik mekanizmaya ilişkin önemli bilgiler sağlamaktadır. Özellikle diyastolik üfürümler, aort yetmezliği veya mitral stenoz gibi ciddi patolojilerle daha güçlü biçimde ilişkilendirildiğinden klinik açıdan daha özgül kabul edilmektedir (McCrindle vd., 2007; McGee, 2018). Bununla birlikte yalnızca zamanlama değil, üfürümün şiddeti, süresi, tonal karakteri ve crescendo, decrescendo veya holosistolik gibi morfolojik desenleri de akustik profili belirleyen temel unsurlardır.

Akustik açıdan masum (fizyolojik) ve patolojik üfürümler arasında belirgin farklar bulunmaktadır. Masum üfürümler genellikle düşük yoğunluklu, dar bantlı ve çoğunlukla 200 Hz'in altında yoğunlaşan frekans bileşenleri sergilerken, patolojik üfürümler daha geniş bantlı spektral dağılımlar göstermekte ve frekans içeriği çoğu zaman daha yüksek bantlara uzanmaktadır. Bu durum, yüksek hızlı jet akımları ve düzensiz türbülans yapıları ile ilişkilidir (Araujo vd., 2022; Yaseen vd., 2018). Bu geniş bantlı ve düzensiz enerji dağılımları, patolojik üfürümlerin FKG sinyallerinde ayırt edici bir akustik imza oluşturmaya neden olmaktadır.

Bu akustik farklar zaman–frekans temsilleri kullanılarak daha açık ve nesnel bir biçimde ortaya konulabilmektedir. Özellikle Mel-spektrogramlar, frekans içeriğini insan işitme algısına daha yakın bir ölçekte sunarak üfürümlerin görsel olarak ayırt edilmesini

kolaylaştırmaktadır (Logan, 2000; LeCun vd., 2015). Normal ve patolojik üfürümlerin Mel-spektrogram temsillerinde, enerji dağılımının frekans ve zaman boyunca farklı biçimlerde yoğunlaştığı gözlemlenmektedir. Masum üfürümlerde enerji genellikle düşük frekans bantlarında daha düzenli bir yapı sergilerken, patolojik üfürümlerde daha geniş bantlı ve düzensiz bir spektral yapı ortaya çıkmaktadır.

Bu tür zaman–frekans temsilleri, otomatik üfürüm tespit sistemleri açısından kritik bir rol oynamaktadır. Özellikle evrişimli sinir ağları (CNN), bu temsiller üzerinden spektral ve zamansal örüntüleri öğrenerek patolojik üfürümlerin daha etkin biçimde sınıflandırılmasına olanak tanımaktadır (Rubin vd., 2016; Potes vd., 2016). Dolayısıyla bu görsel temsiller yalnızca betimleyici değil, aynı zamanda modern derin öğrenme tabanlı sistemlerin temelini oluşturan önemli bir ara temsil katmanıdır.

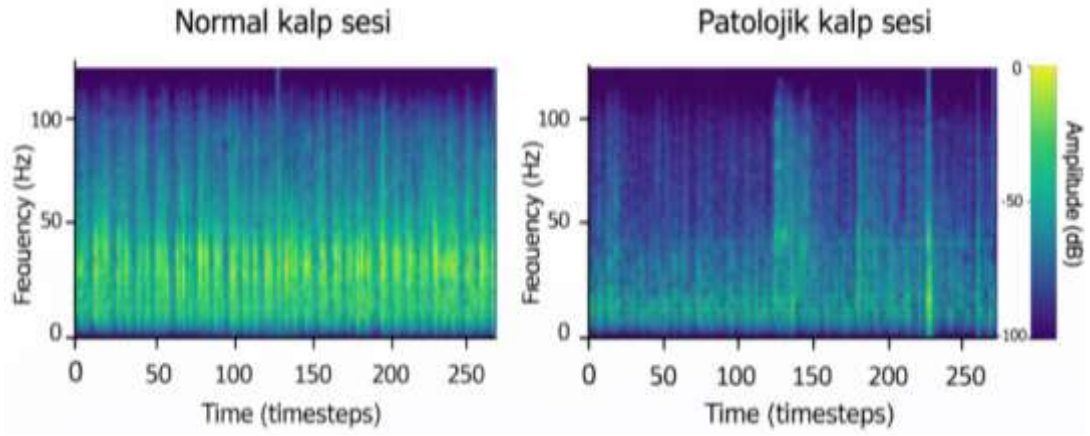
Literatürde sıklıkla vurgulanan bir diğer önemli husus, insan oskültasyon performansındaki belirgin değişkenliktir. Üfürüm tespitindeki klinik doğruluk, hekimin eğitimi, deneyimi ve değerlendirme ortamının akustik koşullarına büyük ölçüde bağlıdır. Yapılan çalışmalar, eğitilmiş klinisyenler arasında dahi gözlemciler arası uyumu ölçen Cohen'in kappa katsayısının çoğu zaman sınırlı seviyelerde kaldığını göstermektedir (McGee, 2018; Mangione vd., 1993; Reyna vd., 2023).

Bu katsayı aşağıdaki şekilde tanımlanmaktadır:

$$\kappa = \frac{P_o - P_e}{1 - P_e} \quad (\text{Eşitlik 13})$$

Burada P_o gözlenen uyum oranını, P_e ise rastlantısal beklenen uyumu ifade etmektedir.

Bu durum, geleneksel oskültasyonun öznel ve hataya açık doğasını yansıtmakta olup özellikle pediatrik hasta gruplarında ve birinci basamak sağlık hizmetlerinde daha tutarlı ve objektif tanı desteği sağlayabilecek otomatik üfürüm tespit sistemlerine duyulan ihtiyacı daha da belirgin hâle getirmektedir.



Şekil 5. Normal ve patolojik kalp üfürümlerinin Mel-spektrogram zaman–frekans temsillerinin karşılaştırılması.

Sonuç olarak, kalp üfürümlerinin akustik kökenlerinin ve klinik öneminin ayrıntılı biçimde anlaşılması, hem manuel oskültasyonun sınırlamalarını ortaya koymakta hem de FKG tabanlı otomatik analiz sistemlerinin geliştirilmesi için güçlü bir bilimsel gerekçe sunmaktadır. Klinik değerlendirmelerde gözlemciye bağlı değişkenlik ve sınırlı tekrarlanabilirlik, yapay zekâ destekli sistemlere olan ihtiyacı daha da belirgin hâle getirmektedir (Mangione vd., 1993; McGee, 2018; Thompson vd., 2019).

Masum ve patolojik üfürümler arasındaki spektral ve zamansal farkların sistematik biçimde tanımlanması, bu tez kapsamında önerilen hibrit derin öğrenme yaklaşımlarının temelini oluşturan ayırt edici özelliklerin neden ve nasıl seçildiğini açıklayan kritik bir kuramsal altyapı sağlamaktadır. Özellikle zaman–frekans temsilleri ve derin öğrenme tabanlı özellik öğrenme yaklaşımları, bu farklılıkların daha etkin biçimde modellenmesine olanak tanımaktadır (Araujo vd., 2022; LeCun vd., 2015; Rubin vd., 2016).

Bu bağlamda, fizyolojik bilgi ile sinyal işleme ve yapay zekâ temelli yöntemlerin bütünleştirilmesi, kalp üfürümlerinin daha güvenilir, tutarlı ve klinik açıdan anlamlı biçimde analiz edilmesini mümkün kılmaktadır. Literatürde, veri temelli yaklaşımlar ile klinik bilginin entegrasyonunun tanısal doğruluğu artırdığı ve karar destek sistemlerinin etkinliğini güçlendirdiği açıkça vurgulanmaktadır (Kelly vd., 2019; Reyna vd., 2023). Bu bütünleşik yaklaşım, otomatik kardiyak oskültasyon sistemlerinin geliştirilmesi için sağlam ve ölçeklenebilir bir temel sunmaktadır.

2.4. Fonokardiyografi (FKG) ve Çok Noktalı Oskültasyon

Fonokardiyografi (FKG), ilk analog uygulamalarından günümüzde kullanılan gelişmiş dijital sistemlere kadar önemli bir evrim geçirmiştir. Modern dijital stetoskoplar,

yüksek duyarlılığa sahip sensörler ve gelişmiş analog–dijital dönüştürücüler sayesinde ileri düzey sayısal sinyal işleme ve yapay zekâ tabanlı analizler için uygun, yüksek sadakatli kalp sesi kayıtları sunmaktadır. Güncel FKG sistemlerinde genellikle 1–4 kHz aralığında örnekleme hızları tercih edilmekte olup, bu aralık hem temel kalp seslerinin hem de daha yüksek frekans bileşenleri içeren üfürümlerin güvenilir biçimde temsil edilmesine olanak sağlamaktadır (Oliveira vd., 2022; Springer vd., 2016).

Modern fonokardiyografi sistemlerinde sinyal kalitesini artırmaya yönelik birçok gürültü azaltma tekniği hem donanım hem de yazılım düzeyinde uygulanmaktadır. Analog ön yükselticiler, bant sınırlayıcı filtreler ve uyarlanabilir gürültü azaltma algoritmaları; solunum sesleri, ortam gürültüsü ve mekanik artefaktların etkisini azaltarak daha temiz FKG sinyallerinin elde edilmesini sağlamaktadır. Bu tür ön işlemler, özellikle düşük genlikli üfürüm bileşenlerinin korunması ve sonraki analizlerin kararlılığı açısından kritik öneme sahiptir (Springer vd., 2016; Schmidt vd., 2010).

Tipik bir FKG dalga biçimi, kalp döngüsüne karşılık gelen farklı akustik bileşenlerden oluşmaktadır. Bunlar arasında S1 ve S2 gibi geçici ve yüksek enerjili olaylar, diyastolik fazda gözlenen düşük frekanslı bileşenler ve zamansal ile spektral açıdan değişkenlik gösteren üfürüm segmentleri yer almaktadır. FKG sinyallerinin durağan olmaması, yani zaman içinde frekans içeriğinin değişkenlik göstermesi, analiz sürecini karmaşık hâle getirmektedir. Bu karmaşıklık; göğüs anatomisi, sensör yerleşimi, solunum hareketleri ve hasta özgü fizyolojik farklılıklar gibi faktörlerle daha da artmaktadır (Liu vd., 2016; Nie vd., 2024).

Klinik oskültasyon pratiği geleneksel olarak dört ana torasik dinleme bölgesine dayanmaktadır: aortik, pulmonik, triküspit ve mitral alanlar. Bu bölgelerin her biri, farklı kardiyak yapıların akustik yansımalarını öne çıkaran özgün bir akustik pencere sunmaktadır. Üfürümlerin bu bölgelerdeki dağılımı, yalnızca varlığını değil aynı zamanda olası etiyolojisini değerlendirmede de önemli klinik ipuçları sağlamaktadır (McCrindle vd., 2007; McGee, 2018).

Son yıllarda CirCor DigiScope gibi veri setleriyle önem kazanan çok noktalı oskültasyon yaklaşımı, mekânsal bilginin sayısal analiz süreçlerine entegre edilmesini mümkün kılmıştır. Birden fazla oskültasyon bölgesinden elde edilen FKG kayıtları, üfürümlerin anatomik dağılımına bağlı olarak daha güvenilir biçimde yakalanmasını sağlamaktadır (Oliveira vd., 2022).

Bu yaklaşım, tek noktadan alınan kayıtlarda ortaya çıkabilen bölgesel akustik zayıflama, sensör yerleşim hataları veya lokal artefaktlardan kaynaklanan yanlış sınıflandırma riskini önemli ölçüde azaltmaktadır. Ayrıca çok kanallı veri yapıları, derin

öğrenme modellerinin yalnızca zamansal ve spektral özellikleri değil aynı zamanda mekânsal örüntüleri de öğrenmesine olanak tanımaktadır (Yaseen vd., 2018; Potes vd., 2016).

George B. Moody PhysioNet Challenge kapsamında sunulan büyük ölçekli ve standartlaştırılmış veri setleri ile birlikte değerlendirildiğinde, çok konumlu FKG verileri patolojik ve normal akustik örüntülerin daha genellenebilir temsillerinin öğrenilmesini desteklemiştir. Bu gelişme, farklı hasta popülasyonları ve kayıt koşulları arasında sınırlı genelleme performansı sorununu önemli ölçüde azaltmıştır (Reyna vd., 2023; Kelly vd., 2019).

2.5. Fonokardiyografi (FKG) Analizi için Sayısal Sinyal İşleme (DSP) Teknikleri

Sayısal sinyal işleme (Digital Signal Processing – DSP) teknikleri, otomatik kalp sesi analizi ve üfürüm tespiti sistemlerinin temelini oluşturmaktadır. FKG sinyalleri, hem geçici (S1, S2 gibi) hem de sürekli (üfürüm) bileşenler içermesi nedeniyle durağan olmayan biyomedikal sinyaller sınıfına girmektedir. Bu durum, ham sinyallerden doğrudan ayırt edici özelliklerin çıkarılmasını zorlaştırmakta ve güvenilir analiz için çok aşamalı ön işleme ve özellik çıkarma süreçlerini gerekli kılmaktadır (Springer vd., 2016; Liu vd., 2016).

Erken dönem FKG analiz sistemlerinde DSP yaklaşımlarının odağı, öncelikle sinyal kalitesinin iyileştirilmesine yönelmiştir. Solunum gürültüsü, hareket artefaktları ve baz çizgisi kaymaları gibi bileşenlerin bastırılması amacıyla bant geçiren filtreleme yaygın olarak kullanılmıştır. Literatürde kalp seslerinin temel enerjisinin çoğunlukla 20–400 Hz aralığında yoğunlaştığı bildirilmiş ve bu frekans bandına odaklanan filtreleme stratejileri geliştirilmiştir (Araujo vd., 2022; Schmidt vd., 2010).

Daha gelişmiş DSP yaklaşımlarında ise dalgacık dönüşümü (wavelet transform), uyarlamalı filtreleme ve sinyal normalizasyonu gibi teknikler kullanılmaktadır. Özellikle dalgacık tabanlı yöntemler, hem zaman hem frekans çözünürlüğü sağlayarak FKG sinyallerindeki kısa süreli ve düşük genlikli üfürüm bileşenlerinin daha etkin biçimde ortaya çıkarılmasına olanak tanımaktadır (Addison, 2005; Nigam ve Priemer, 2005). Bu tekniklerin birlikte kullanımı, sinyal-gürültü oranını artırarak analiz performansını önemli ölçüde iyileştirmektedir.

Zaman alanı özellikleri, FKG analizinin en erken kullanılan tanımlayıcıları arasında yer almaktadır. Bu özellikler arasında toplam enerji, RMS enerji, genlik zarfı (envelope), sıfır geçiş oranı (ZCR) ve sistolik–diyastolik oranlar bulunmaktadır. Bu tür özellikler,

kalp döngüsünün zamansal yapısını doğrudan yansıtarak kardiyak mekaniklerin analizine katkı sağlamaktadır (Chen vd., 2023; Liang vd., 1997).

Bununla birlikte, zaman alanı özellikleri çoğunlukla sinyalin genel genlik ve süre karakteristiklerini temsil etmekte olup, patolojik üfürümlerde gözlenen karmaşık spektral yapıları yeterince yakalayamamaktadır. Bu sınırlılık, analiz odağının frekans alanı ve zaman–frekans temsillerine kaymasına neden olmuştur. Özellikle Fourier dönüşümü ve kısa zamanlı Fourier dönüşümü (STFT) gibi yöntemler, sinyalin frekans içeriğini analiz etmek için yaygın olarak kullanılmaktadır (Rabiner ve Schafer, 2011).

Frekans alanı analizinde Hızlı Fourier Dönüşümü (Fast Fourier Transform – FFT) ve Güç Spektral Yoğunluğu (Power Spectral Density – PSD) kestirimi gibi yöntemler ön plana çıkmaktadır. FFT aşağıdaki şekilde tanımlanmaktadır:

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi kn/N} \quad (\text{Eşitlik 14})$$

Bu dönüşüm sayesinde spektral merkez, bant genişliği ve enerji dağılımı gibi özellikler nicel olarak incelenebilmektedir. Literatürde patolojik üfürümlerin, masum üfürümlere kıyasla daha geniş spektral yayılım ve daha düzensiz güç dağılımı sergilediği çeşitli çalışmalarla gösterilmiştir (Araujo vd., 2022; Yaseen vd., 2018). Ancak FFT tabanlı yaklaşımlar, sinyalin tamamını tek bir zaman penceresi içerisinde ele aldığından kardiyak döngü boyunca ortaya çıkan zamansal değişimleri koruyamamaktadır (Rabiner ve Schafer, 2011).

FKG sinyallerinin zamana bağlı olarak değişkenlik göstermesi, zaman–frekans analizini kaçınılmaz hâle getirmiştir. Kısa Zamanlı Fourier Dönüşümü (Short-Time Fourier Transform – STFT), dalgacık dönüşümleri ve benzeri yöntemler, sinyalin frekans içeriğini belirli zaman pencereleri içerisinde analiz ederek lokal yapıları ortaya çıkarmaktadır. STFT sabit çözünürlüklü ve yorumlanabilir bir spektrogram temsili sunarken, Sürekli Dalgacık Dönüşümü (Continuous Wavelet Transform – CWT) zaman ve frekans çözünürlüğü arasında daha esnek bir denge sağlayarak ani ve geçici olayların analizinde avantaj sunmaktadır (Addison, 2005; Nigam ve Priemer, 2005).

Bu zaman–frekans temsilleri, daha sonraki yıllarda derin öğrenme tabanlı FKG analiz sistemlerinde yaygın olarak kullanılan Mel-spektrogram gibi türev gösterimlerin kavramsal temelini oluşturmuştur. Özellikle mel ölçeği, insan işitme algısına daha yakın bir frekans çözünürlüğü sunarak üfürümlerin algısal olarak anlamlı bileşenlerinin vurgulanmasına olanak tanımaktadır (Logan, 2000; Stevens vd., 1937). Bu nedenle modern otomatik üfürüm tespit sistemlerinde kullanılan spektral temsiller, klasik DSP

tekniklerinin doğal bir uzantısı olarak değerlendirilmektedir (LeCun vd., 2015; Rubin vd., 2016).

2.6. Özellik Temsili Olarak Mel-spektrogramlar

Mel spektrogramlar, FKG analizi için en güçlü ve en yaygın kullanılan zaman–frekans temsillerinden biri hâline gelmiştir. Kısa Zamanlı Fourier Dönüşümü’nden (STFT) türetilen ve insan işitme sisteminin doğrusal olmayan algısal özellikleriyle uyumlu mel frekans ölçeğine haritalanan bu gösterim, özellikle üfürümlerin yoğunlaştığı düşük frekans bantlarını belirginleştirmektedir (Logan, 2000; Stevens vd., 1937). Bu özellik, Mel spektrogramların klinik açıdan anlamlı akustik bileşenleri ön plana çıkarırken yüksek frekanslı gürültü etkilerini görece bastırmasını sağlamaktadır. Ayrıca bu temsiller, evrişimli sinir ağları (CNN) için uygun, kompakt ve yüksek ifade gücüne sahip iki boyutlu bir giriş formatı sunmaktadır (LeCun vd., 2015; Rubin vd., 2016).

Literatürde yer alan çok sayıda çalışma, Mel spektrogram tabanlı derin öğrenme modellerinin, veri setinin karmaşıklığına ve kullanılan mimariye bağlı olarak, ham dalga biçimi (raw waveform) tabanlı yaklaşımlara kıyasla daha üstün performans sergilediğini göstermektedir. Özellikle kalp sesi sınıflandırma problemlerinde, Mel spektrogramların spektral ve zamansal örüntüleri daha ayırt edici biçimde temsil edebildiği ve sınıflandırma doğruluğunu anlamlı ölçüde artırdığı rapor edilmiştir (Araujo vd., 2022; Yaseen vd., 2018; Potes vd., 2016).

Bu performans artışı, Mel spektrogramların üfürümlere özgü harmonik yapıları ve türbülans kaynaklı düzensiz spektral bileşenleri vurgularken, tanısal açıdan daha az anlamlı olan yüksek frekanslı gürültü bileşenlerini bastırabilme yeteneğinden kaynaklanmaktadır. Bu durum, özellikle gürültüye duyarlı klinik kayıtlar üzerinde model dayanıklılığını artıran önemli bir faktör olarak değerlendirilmektedir (Liu vd., 2016).

Mel spektrogramların etkinliği, pencere uzunluğu, adım boyutu (hop size), FFT çözünürlüğü ve mel filtre bankası sayısı gibi parametrelere doğrudan bağlıdır. Daha kısa pencere uzunlukları zamansal çözünürlüğü artırırken frekans çözünürlüğünü sınırlamakta; daha uzun pencere uzunlukları ise frekans çözünürlüğünü artırırken zamansal hassasiyeti azaltmaktadır. Benzer şekilde mel bant sayısının artırılması spektral ayrıntıyı zenginleştirirken hesaplama maliyetini de artırmaktadır (Rabiner ve Schafer, 2011; Araujo vd., 2022).

Dolayısıyla Mel spektrogram parametrelerinin dikkatli biçimde optimize edilmesi, hem ayırt edici hem de genellenebilir özellik temsillerinin elde edilmesi açısından kritik

bir tasarım adımı olarak öne çıkmaktadır. Mel frekans ölçeği aşağıdaki Eşitlik 15 ile tanımlanmaktadır:

$$m = 2595 \log_{10} \left(1 + \frac{f}{700} \right) \quad (\text{Eşitlik 15})$$

Bu dönüşüm, frekans eksenini insan işitme sisteminin algısal duyarlılığına uygun biçimde doğrusal olmayan bir ölçekte yeniden haritalayarak düşük frekans bantlarında daha yüksek çözünürlük sağlamaktadır. Bu özellik, pediatrik kalp üfürümlerinin çoğunlukla yoğunlaştığı düşük frekans bölgelerinin daha belirgin biçimde temsil edilmesine katkıda bulunmaktadır.

Bu avantajları daha somut biçimde ortaya koymak amacıyla Tablo 1, Mel spektrogramların kalp sesi analizinde yaygın olarak kullanılan diğer özellik temsilleriyle karşılaştırmalı bir özetini sunmaktadır.

Bu nedenle Mel-spektrogramlar, hem zaman hem de frekans alanındaki bilgileri dengeli bir şekilde temsil edebilme yetenekleri sayesinde öne çıkmaktadır. Yorumlanabilirliklerinin yüksek olması, araştırmacıların ve klinisyenlerin model çıktılarının arkasındaki örüntüleri daha iyi anlamalarına olanak tanırken, aynı zamanda hesaplama maliyetlerinin görece düşük olması bu yöntemi pratik açıdan da avantajlı kılmaktadır. Bu özellikleri sayesinde Mel-spektrogramlar, gerek klasik makine öğrenimi algoritmalarında gerekse derin öğrenme tabanlı üfürüm sınıflandırma yaklaşımlarında yaygın olarak tercih edilmekte ve etkin bir temsil yöntemi olarak önemini korumaktadır.

Tablo 1. FKG analizinde kullanılan özellik temsillerinin karşılaştırılması

Temsil	Avantajlar	Sınırlamalar
Ham dalga formu	Ön işleme gerektirmez; sinyalin tamamını korur	Gürültüye duyarlıdır; düşük performans; modellenmesi zordur
Standart spektrogram	Yüksek spektral ayrıntı; yaygın olarak desteklenir	Algısal olarak hizalı değildir; yüksek frekanslara aşırı ağırlık verebilir
MFCC'ler	Kompakt özellikler; konuşma görevlerinde etkilidir	Yüksek frekanslara aşırı ağırlık verebilir; sınırlı temsil gücü
Mel-spektrogram	Dengeli temsil; üfürüm frekanslarını vurgular; dayanıklıdır	Yüksek bellek gereksinimi; parametre seçimine duyarlıdır
Sabit-Q Dönüşümü (CQT)	Harmonik sinyaller için uygundur; düşük frekans çözünürlüğü yüksektir	Hesaplama maliyeti yüksektir; FKG için karmaşıktır

2.7. Klasik Makine Öğrenimi Yaklaşımları

Derin öğrenme temelli yaklaşımların yaygınlaşmasından önce, FKG sinyallerinden otomatik kalp üfürümü tespiti alanında el ile tasarlanmış özelliklere dayanan klasik makine öğrenimi yöntemleri baskın bir rol oynamıştır. Bu yaklaşımlar, zaman alanı, frekans alanı ve kepsral alan özelliklerinden elde edilen istatistiksel tanımlayıcıların çeşitli sınıflandırıcılar aracılığıyla modellenmesine dayanmaktadır. Özellikle sınırlı veri kümeleri ve düşük hesaplama maliyetine sahip sistemler için klasik yöntemler uzun süre uygulanabilir ve güvenilir çözümler sunmuştur (Potes vd., 2016; Springer vd., 2016).

Destek Vektör Makineleri (Support Vector Machines – SVM), yüksek boyutlu özellik uzaylarında güçlü genelleme kabiliyeti ve çekirdek (kernel) fonksiyonları aracılığıyla doğrusal olmayan karar sınırlarını modelleyebilmesi nedeniyle literatürde en yaygın kullanılan sınıflandırıcılardan biri olmuştur (Cortes ve Vapnik, 1995). Özellikle Mel Frekans Kepsral Katsayıları (MFCC) ve spektral özelliklerle birleştirilen SVM tabanlı modellerin kalp sesi sınıflandırma görevlerinde yüksek doğruluk oranlarına ulaşabildiği çeşitli çalışmalarda gösterilmiştir (Rubin vd., 2016; Liu vd., 2016). Bununla birlikte, SVM performansının çekirdek seçimi ve hiperparametre ayarlarına duyarlı olması, farklı veri setleri arasında tutarlı performans elde edilmesini zorlaştırabilmektedir.

k-En Yakın Komşu (k-Nearest Neighbors – KNN) algoritması, basitliği ve sezgisel yapısı nedeniyle erken dönem FKG sınıflandırma çalışmalarında sıklıkla tercih edilmiştir. Küçük ve dengeli veri kümelerinde kabul edilebilir performans sunmasına rağmen, yöntemin gürültüye duyarlılığı ve yüksek boyutlu veri uzaylarında artan hesaplama maliyeti önemli dezavantajlar olarak öne çıkmaktadır (Cover ve Hart, 1967; Araujo vd., 2022). Ayrıca çıkarım aşamasında tüm veri kümesinin kullanılması gerekliliği, gerçek zamanlı sistemler için sınırlayıcı bir faktör oluşturmaktadır.

Ağaç tabanlı yöntemler, özellikle Rastgele Ormanlar (Random Forest) ve gradyan artırma algoritmaları, klasik makine öğrenimi literatüründe önemli bir yer edinmiştir. Bu modeller, birden fazla karar ağacının birleşimiyle daha kararlı ve genellenebilir tahminler üretmekte ve özellik önem skorları aracılığıyla yorumlanabilirlik sağlamaktadır (Breiman, 2001; Chen ve Guestrin, 2016). FKG analizi bağlamında, enerji dağılımı, spektral merkez ve bant genişliği gibi özelliklerin sınıflandırma kararına katkısının açık biçimde incelenebilmesi, bu yöntemleri klinik açıdan değerli hâle getirmiştir.

Bununla birlikte, klasik makine öğrenimi yaklaşımlarının temel sınırlaması, performanslarının büyük ölçüde el ile tasarlanmış özelliklerin kalitesine bağımlı olmasıdır. En uygun özellik kümesinin veri toplama koşullarına, hasta özelliklerine ve

patoloji türüne bağlı olarak değişmesi, bu yöntemlerin genellenebilirliğini sınırlamaktadır. Ayrıca FKG sinyallerinde bulunan karmaşık ve örtüşen akustik örüntülerin manuel özellik mühendisliği ile tam olarak temsil edilememesi, bu yaklaşımların performansını sınırlayan önemli bir faktör olarak ortaya çıkmıştır. Bu nedenlerle, özellik çıkarımı ve sınıflandırmayı uçtan uca öğrenebilen derin öğrenme temelli yaklaşımlara yönelim hız kazanmıştır (LeCun vd., 2015; Araujo vd., 2022).

2.8. Otomatik Kalp Üfürümü Tespiti için Derin Öğrenme

Derin öğrenmenin FKG analizi alanına dâhil edilmesi, otomatik kalp üfürümü tespiti yaklaşımlarında köklü bir paradigma değişimini temsil etmektedir. El ile tasarlanmış özelliklere dayanan klasik makine öğrenimi yöntemlerinin aksine, derin öğrenme mimarileri ham veya minimal düzeyde işlenmiş sinyallerden doğrudan ayırt edici temsilleri öğrenebilme yeteneğine sahiptir. Bu uçtan uca öğrenme yaklaşımı, özellikle kalp üfürümlerinin zamansal ve spektral değişkenliğinin yüksek olması nedeniyle büyük önem taşımaktadır (LeCun vd., 2015; Reyna vd., 2023).

Evrişimli Sinir Ağları (CNN), FKG tabanlı sınıflandırma görevlerinde başarı elde eden ilk derin öğrenme mimarileri arasında yer almıştır. CNN'ler özellikle Mel spektrogram gibi iki boyutlu zaman–frekans temsilleri üzerinde çalışmaya oldukça elverişlidir. Bu mimarilerde erken evrişim katmanları lokal akustik olayları yakalarken, daha derin katmanlar üfürümlere özgü karmaşık spektral örüntüleri öğrenmektedir (Rubin vd., 2016; Potes vd., 2016). Derin CNN mimarilerinin (özellikle ResNet) sunduğu artık bağlantılar (skip connections), daha derin ağların eğitilmesini mümkün kılarak genelleme performansını artırmıştır (He vd., 2016).

Bununla birlikte yalnızca CNN tabanlı modeller, uzun süreli zamansal bağımlılıkları modelleme konusunda sınırlı kalabilmektedir. FKG sinyallerinin ardışık ve periyodik doğası göz önüne alındığında, zamansal bağımlılıkların modellenmesi kritik öneme sahiptir. Bu amaçla Tekrarlayan Sinir Ağları (RNN) ve özellikle LSTM birimleri literatürde yaygın olarak kullanılmıştır. LSTM mimarileri, kaybolan gradyan problemini azaltarak uzun menzilli bağımlılıkları öğrenebilmekte ve kalp döngüsü boyunca değişen akustik örüntüleri modelleyebilmektedir (Hochreiter ve Schmidhuber, 1997).

Bu farklı güçlü ve zayıf yönler, hibrit CNN–LSTM mimarilerinin geliştirilmesini teşvik etmiştir. Bu mimarilerde CNN katmanları spektral özellikleri çıkarırken, LSTM katmanları bu özelliklerin zamansal evrimini modellemektedir. Literatürde hibrit CNN–LSTM yaklaşımlarının, tek başına CNN veya LSTM modellerine kıyasla daha yüksek ve

daha kararlı performans sağladığı tutarlı biçimde rapor edilmiştir (Yaseen vd., 2018; Araujo vd., 2022).

Son yıllarda dikkat (attention) mekanizmaları ve transformer tabanlı mimariler de bu alanda önemli bir araştırma yönü hâline gelmiştir. Dikkat mekanizmaları, modelin girişin farklı bölümlerine farklı ağırlıklar atamasını sağlayarak tanısallık açısından önemli zaman–frekans bölgelerini vurgulamaktadır (Vaswani vd., 2017). Transformer tabanlı modeller, yalnızca öz-dikkat mekanizmaları aracılığıyla küresel bağımlılıkları modelleyebilmekte ve özellikle uzun dizilerde üstün performans sunabilmektedir.

Bununla birlikte, bazı transformer tabanlı yaklaşımlar umut verici sonuçlar sunsa da, literatürde rapor edilen yüksek doğruluk değerlerinin veri seti büyüklüğü ve deneysel kurulumlara duyarlı olduğu görülmektedir. Bu nedenle bu tür sonuçların genellenebilirliği dikkatli biçimde değerlendirilmelidir.

Bir diğer önemli gelişme, çoklu modlu (multimodal) derin öğrenme yaklaşımlarının ortaya çıkmasıdır. Bu yaklaşımlar, ham dalga formları, Mel spektrogramlar ve MFCC gibi farklı temsil türlerini birlikte işleyerek daha zengin özellik uzayları oluşturmaktadır (Baltrusaitis vd., 2019). Bu tür mimariler, veri çeşitliliğine ve gürültüye karşı daha dayanıklı temsiller üretebilmektedir.

Etiketli FKG veri kümelerinin sınırlı olması, derin öğrenme modellerinin eğitimi açısından önemli bir zorluk oluşturmaktadır. Bu bağlamda öz-denetimli (self-supervised) ve yarı denetimli öğrenme yaklaşımları giderek daha fazla önem kazanmaktadır. Özellikle karşıt öğrenme (contrastive learning) yöntemleri, sınırlı etiketli veri ile güçlü temsiller öğrenmeyi mümkün kılmaktadır (Chen vd., 2020).

Genel olarak değerlendirildiğinde, literatür otomatik kalp üfürümü tespiti alanının; hibrit, dikkat tabanlı ve veri verimliliği yüksek derin öğrenme mimarilerine doğru evrildiğini açık biçimde ortaya koymaktadır. Bu eğilim, FKG sinyallerinin karmaşık doğasının tek bir model veya temsil ile tam olarak açıklanamayacağını göstermektedir.

2.9. 2022–2024 Dönemi En Gelişmiş Üfürüm Tespit Sistemleri

Son yıllarda otomatik kalp üfürümü tespiti alanındaki ilerlemeler, büyük ölçüde standartlaştırılmış kıyaslama veri setleri ve açık bilimsel yarışmalar tarafından yönlendirilmiştir. Bu bağlamda George B. Moody PhysioNet Challenge, hem metodolojik gelişmeleri hızlandırması hem de farklı yaklaşımların adil biçimde karşılaştırılmasına olanak tanınması açısından merkezi bir rol oynamıştır. Özellikle 2022 yılında düzenlenen yarışma, çoklu oskültasyon bölgelerinden elde edilen ve cihaz ile hasta koşulları bakımından yüksek derecede heterojenlik içeren geniş ölçekli bir pediatrik

FKG veri setini literatüre kazandırmıştır (Reyna vd., 2023). Bu veri seti, modellerin yalnızca belirli veri dağılımlarına değil, gerçek dünya klinik varyasyonlarına da genellenebilmesini hedefleyen bir değerlendirme çerçevesi sunmuştur.

Yarışmada en iyi performansı sergileyen sistemlerin ortak özellikleri incelendiğinde, Mel spektrogram tabanlı temsillerin baskın biçimde kullanıldığı görülmektedir. Bu temsiller genellikle evrişimli sinir ağları (CNN) ile birleştirilmiş; bazı çalışmalar ise dikkat (attention) mekanizmaları ekleyerek tanısal açıdan önemli zaman–frekans bölgelerine odaklanmayı başarmıştır (Potes vd., 2016; Rubin vd., 2016). Bununla birlikte hasta düzeyinde veri bölme stratejilerinin benimsenmesi, eğitim ve test kümeleri arasında veri sızıntısını önleyerek daha gerçekçi performans değerlendirmeleri yapılmasını sağlamıştır (Kaufman vd., 2012).

Performans değerlendirme ölçütleri de bu dönemde klinik ihtiyaçları daha iyi yansıtacak biçimde evrim geçirmiştir. Geleneksel doğruluk ölçütlerinin ötesinde, dengesiz veri kümeleri için daha uygun olan weighted accuracy ve F1-skoru gibi metrikler ön plana çıkmıştır. Bu metrikler, özellikle klinik açıdan daha kritik olan yanlış negatif hataları daha ağır biçimde cezalandırarak tanısal güvenilirliği artırmayı hedeflemektedir (Saito ve Rehmsmeier, 2015). PhysioNet Challenge 2022 sonuçları, heterojen veri setleri üzerinde yüksek performans elde etmenin zorluklarını açıkça ortaya koymuştur (Reyna vd., 2023).

PhysioNet yarışmaları dışında, son yıllarda yapılan çalışmalar da alanın yönünü belirlemektedir. Özellikle çoklu modlu ve dikkat tabanlı yaklaşımlar, FKG sinyallerinin karmaşık doğasını daha iyi modelleyebilmek için tercih edilmektedir (Baltrusaitis vd., 2019; Vaswani vd., 2017). Bu tür yaklaşımlar, farklı temsil biçimlerini ve küresel bağımlılıkları birlikte ele alarak daha güçlü ve genellenebilir modeller geliştirilmesine katkı sağlamaktadır.

Bu dönemde geliştirilen ileri düzey sistemlerin bir diğer ortak özelliği, kapsamlı veri artırma stratejilerinin kullanılmasıdır. Zaman esnetme, gürültü ekleme ve sinyal segmentasyonu gibi teknikler, sınırlı veri kümelerinde aşırı öğrenme riskini azaltarak model genellenebilirliğini artırmaktadır (Shorten ve Khoshgoftaar, 2019). Ayrıca saliency haritaları ve dikkat görselleştirmeleri gibi açıklanabilirlik yöntemleri, model kararlarının klinik açıdan daha anlaşılır hâle gelmesini sağlamaktadır (Selvaraju vd., 2017; Lundberg ve Lee, 2017).

Genel olarak değerlendirildiğinde, son yıllarda geliştirilen otomatik üfürüm tespit sistemleri belirli tasarım ilkeleri etrafında birleşmektedir. Bunlar arasında Mel spektrogram temsilleri, hibrit mimariler, dikkat mekanizmaları, veri artırma teknikleri ve

hasta düzeyinde değerlendirme stratejileri yer almaktadır. Bu eğilimler, bu tezde önerilen yaklaşımın metodolojik temelini doğrudan desteklemektedir. Özellikle optimize edilmiş ön işleme adımlarıyla desteklenen Mel spektrogram temsilleri ve zamansal bağımlılıkları modelleyebilen hibrit CNN–LSTM çerçevesi, literatürdeki güncel gelişmelerle uyumlu ve bilimsel olarak gerekçelendirilmiş bir çözüm sunmaktadır.

2.10. Literatür Taramasının Özeti

İncelenen literatür, otomatik kalp üfürümü tespitinin gelişim sürecinde geleneksel oskültasyondan dijital sinyal işleme (DSP) temelli yaklaşımlara, ardından klasik makine öğrenimi yöntemlerine ve son olarak derin öğrenme tabanlı mimarilere doğru tutarlı ve aşamalı bir evrim yaşandığını açık biçimde ortaya koymaktadır (LeCun vd., 2015; Reyna vd., 2023). Bu evrim sürecinin her aşaması, FKG tabanlı tanı sistemlerinin günümüzde ulaştığı olgunluk düzeyine erişmesinde kritik katkılar sağlamıştır.

İlk aşamada, kardiyovasküler fizyoloji ve kalp üfürümlerinin akustik kökenlerine ilişkin çalışmalar, FKG sinyallerinin biyolojik temelini ortaya koyarak modellerin öğrenmesi gereken temel spektral ve zamansal ipuçlarını belirginleştirmiştir (McCrindle vd., 2007; McGee, 2018). Fonokardiyografi, özellikle çoklu oskültasyon bölgelerinden elde edilen kayıtlar aracılığıyla kardiyak mekaniklerin zengin, invaziv olmayan ve tekrarlanabilir bir temsili sunarak bu alanda temel bir veri kaynağı hâline gelmiştir (Oliveira vd., 2022).

Bunu izleyen süreçte DSP temelli yaklaşımlar, üfürüm imzalarının zaman ve frekans boyunca nasıl değiştiğini nicel olarak ortaya koymuş ve zaman–frekans temsillerinin gerekliliğini açık biçimde göstermiştir (Addison, 2005; Rabiner ve Schafer, 2011). Bu bağlamda Mel spektrogram temsilleri, insan işitme algısıyla uyumlu yapısı ve gürültüye karşı dayanıklılığı sayesinde hem klasik hem de derin öğrenme tabanlı modeller için güçlü ve dengeli bir özellik uzayı sunmuştur (Logan, 2000; Araujo vd., 2022).

Klasik makine öğrenimi yaklaşımları, otomatik üfürüm tespitinin teknik olarak uygulanabilirliğini göstermiş olsa da, el ile tasarlanmış özelliklere olan bağımlılıkları ve farklı veri setleri arasında sınırlı genelleme yetenekleri önemli bir kısıt olarak kalmıştır (Guyon ve Elisseeff, 2003). Bu sınırlamalar, özellik çıkarımı ve temsili doğrudan veriden öğrenebilen derin öğrenme yaklaşımlarına yönelimi hızlandırmıştır (LeCun vd., 2015).

Son yıllarda hibrit CNN–LSTM mimarileri ve attention/transformer tabanlı modeller, hem spektral hem de zamansal bağımlılıkları birlikte modelleyebilme yetenekleri sayesinde literatürde en ileri düzeyi (state-of-the-art) temsil etmektedir (Vaswani vd., 2017; Yaseen vd., 2018). Bu mimariler yalnızca yüksek sınıflandırma

performansı sunmakla kalmayıp, aynı zamanda genellenebilirlik ve veri çeşitliliğine karşı dayanıklılık açısından da önceki yaklaşımlara kıyasla belirgin avantajlar sağlamaktadır (Kelly vd., 2019).

Bu metodolojik evrim ve farklı yaklaşımlar arasındaki temel farklar Tablo 2’de özetlenmiştir. Özellikle Mel spektrogram temsillerinin hibrit derin sinir ağı mimarileriyle entegrasyonu, kalp üfürümlerinde gözlenen yüksek değişkenliğin ortaya koyduğu tanısız zorlukları ele almak için güçlü ve tutarlı bir çerçeve sunmaktadır. Bu bütüncül değerlendirme, Bölüm 3’te sunulan metodolojinin bilimsel gerekçesini doğrudan desteklemekte ve önerilen yaklaşımın literatürdeki konumunu açık biçimde ortaya koymaktadır.

Tablo 2. Literatürde otomatik kalp üfürümü tespiti için kullanılan yaklaşımların özellik temsilleri, modelleme teknikleri, güçlü yönleri ve temel sınırlamaları açısından karşılaştırmalı özet

Yaklaşım / Dönem	Temel Temsil (Girdi)	Kullanılan Teknikler	Güçlü Yönler	Temel Sınırlamalar	Örnek Çalışmalar
Geleneksel oskültasyon	İnsan işitmesi	Klinik dinleme	Hızlı, düşük maliyetli	Öznel, deneyime bağımlı, düşük tekrarlanabilirlik	–
DSP tabanlı analiz	Ham FKG	Bant geçiren filtre, FFT, STFT, dalgacık	Zaman–frekans içgörü, yorumlanabilir	Özellik seçimi manuel, sınırlı genelleme	Chen vd., 2023
Zaman alanı özellikleri	FKG dalga formu	Enerji, zarf, ZCR	Hesaplaması kolay	Yüksek frekanslı üfürümleri kaçırır	Araujo vd., 2022
Frekans alanı özellikleri	Spektrum (FFT/PSD)	Spektral merkez, bant genişliği	Türbülans tespiti	Zamansal bilgi kayıbı	Singh ve Agarwal, 2024
Klasik makine öğrenimi	El yapımı özellikler	SVM, kNN, RF, XGBoost	Küçük veri setlerinde güçlü	Özellik mühendisliğine bağımlı	Hosseinzadeh vd., 2024
Mel- spektrogram tabanlı DSP	Zaman– frekans (mel)	STFT + mel ölçeği	Algısal uyum, gürültü dayanımı	Parametre seçimine duyarlı	Lee vd., 2022
CNN tabanlı DL	Mel- spektrogram	CNN, ResNet	Güçlü spektral öğrenme	Uzun dönem bağımlılık zayıf	Lee vd., 2022
RNN / LSTM	Zaman dizisi	LSTM, Bi- LSTM	Zamansal modelleme	Spektral soyutlama zayıf	Netto ve Abraham, 2021
Hibrit CNN– LSTM	Mel- spektrogram + dizi	CNN + LSTM	Spektral + zamansal bütünlük	Mimari karmaşıklık	Netto ve Abraham, 2021
Attention / Transformer	Mel- spektrogram / multimodal	Self- attention, Transformer	Küresel bağlam, yorumlanabilirlik	Hesaplama maliyeti	Nie vd., 2024

Tablo 2. (Devamı)

Yaklaşım / Dönem	Temel Temsil (Girdi)	Kullanılan Teknikler	Güçlü Yönler	Temel Sınırlamalar	Örnek Çalışmalar
Multimodal DL	Waveform + spektrogram + MFCC	Çok kollu ağlar	Donanım ve ortam dayanımı	Senkronizasyon zorluğu	Zhao vd., 2023
SOTA (2022– 2024)	Mel + hibrit	CNN–LSTM + attention	Yüksek doğruluk, klinik uyum	Veri kıtlığı	Reyna vd., 2023

3. METODOLOJİ

3.1. Araştırma Metodolojisine Genel Bakış

Bu bölümde, Mel spektrogram temsilleri ile geleneksel ses özelliklerinin birlikte kullanıldığı ve FKG kayıtlarından pediatrik kalp üfürümlerinin otomatik olarak tespit edilmesini amaçlayan hibrit bir CNN–LSTM sisteminin tasarımı, eğitimi ve değerlendirilmesi için benimsenen metodolojik çerçeve ayrıntılı biçimde sunulmaktadır.

Önerilen metodoloji, veri kümesinin seçimi ve hazırlanması ile başlayan; sinyal ön işleme ve veri artırma adımlarıyla devam eden; ardından hibrit özellik çıkarımı, model mimarisinin tasarımı, eğitim süreci ve çoklu temel yaklaşımlarla gerçekleştirilen karşılaştırmalı nicel değerlendirme aşamalarını içeren bütüncül ve sistematik bir iş akışını takip etmektedir.

Bu yapı, hem geleneksel makine öğrenimi yöntemlerinin hem de modern derin öğrenme mimarilerinin güçlü yönlerini aynı deneysel çerçeve içerisinde değerlendirmeyi amaçlamakta ve önerilen yaklaşımın performansını kapsamlı biçimde analiz etmeye olanak tanımaktadır.

Genel iş akışı, aşağıda sunulan ana adımlar çerçevesinde özetlenmektedir:

3.1.1. Veri Kümesinin Elde Edilmesi ve Seçimi

Çalışmada kullanılan FKG kayıtları, George B. Moody PhysioNet/CinC 2022 Challenge kapsamında yayımlanan ve pediatrik popülasyona odaklanan geniş ölçekli bir kalp sesi veri tabanı olan CirCor DigiScope veri kümesinden elde edilmiştir.

Analize yalnızca kardiyak oskültasyonda standart olarak değerlendirilen dört ana dinleme bölgesinin tamamından elde edilmiş eksiksiz kayıtları bulunan hastalar dâhil edilmiştir. Bu bölgeler; aort kapak bölgesi (AV), mitral kapak bölgesi (MV), pulmoner kapak bölgesi (PV) ve triküspit kapak bölgesi (TV) olarak tanımlanmaktadır.

Bu seçim, kalp seslerinin farklı anatomik bölgelerde sergilediği mekânsal akustik varyasyonların modele dâhil edilmesini sağlayarak, tanısal doğruluğun artırılmasını ve hasta düzeyinde tutarlı bir değerlendirme yapılmasını amaçlamaktadır.

3.1.2. Veri Ön İşleme ve Veri Artırma

Ses kayıtları, dalga biçimi (Waveform Audio File Format – WAV) formatında olup, dört farklı oskültasyon bölgesinden elde edilen sinyalleri ilgili üfürüm etiketleriyle eşleştiren özel olarak geliştirilmiş bir veri yükleme modülü (DataLoaderML sınıfı) aracılığıyla sisteme aktarılmıştır.

Tüm sinyaller, model eğitimi sürecinde tutarlılığı sağlamak amacıyla 22.050 Hz örnekleme hızında ve 5 saniyelik sabit uzunlukta (110.250 örnek) olacak şekilde yeniden örneklenmiş ve normalize edilmiştir. Bu standardizasyon, değişken uzunluktaki kayıtların ortak bir temsil uzayında işlenmesine olanak tanımaktadır.

Üfürüm pozitif ve süresi 15 saniyeden uzun olan kayıtlar, sınıf dengesizliğini azaltmak amacıyla örtüşmeyen birden fazla segmente ayrılmıştır. Bu yaklaşım, az temsil edilen sınıfların veri dağılımını iyileştirerek modelin nadir görülen patolojik örüntülere karşı duyarlılığını artırmayı hedeflemektedir.

3.1.3. Hibrit Özellik Çıkarımı

Bu çalışmada, FKG sinyallerinden elde edilen bilginin mümkün olduğunca kapsamlı ve ayırt edici biçimde temsil edilebilmesi amacıyla iki tamamlayıcı özellik ailesi tanımlanmıştır.

İlk olarak, geleneksel tek boyutlu (1B) ses özellikleri kullanılarak her bir oskültasyon noktasına ait sinyallerden zaman alanı, frekans alanı, kepsral ve istatistiksel tanımlayıcılar çıkarılmıştır. Bu özellik çıkarım süreci, feature_extraction_ML.py dosyasında tanımlanan fonksiyonlar aracılığıyla gerçekleştirilmiş ve klasik makine öğrenmesi modelleri için yapılandırılmış ve ayırt edici bir özellik uzayı oluşturulmuştur.

Buna paralel olarak, kalp seslerinin zamansal ve spektral yapısını birlikte temsil edebilmek amacıyla Mel-spektrogram tabanlı iki boyutlu (2B) temsiller elde edilmiştir. Bu aşamada, mel_spectrogram_deep_learning.py dosyasında tanımlanan MelSpectrogramExtractor ve MelSpectrogramProcessor sınıfları kullanılarak FKG sinyalleri Mel-spektrogramlara dönüştürülmüş ve derin öğrenme mimarilerine uygun olacak şekilde ön işleme tabi tutulmuştur.

Bu iki özellik ailesi, el ile tasarlanmış özelliklerin sunduğu yorumlanabilir akustik tanımlayıcılar ile derin öğrenme tabanlı temsillerin yüksek soyutlama kapasitesini bir araya getirecek şekilde tamamlayıcı olarak yapılandırılmıştır. Bu sayede, üfürümlere özgü kısa süreli enerji yoğunlaşmaları, harmonik yapılar ve kardiyak döngüler boyunca değişen spektral örüntüler eş zamanlı olarak modellenebilmiştir.

3.1.4. Model Mimarisinin Tasarımı

Mel-spektrogram girdilerinden elde edilen uzamsal–spektral örüntüler ile kalp döngüsü boyunca ortaya çıkan zamansal dinamiklerin birlikte modellenenbilmesi amacıyla hibrit bir CNN–LSTM mimarisi geliştirilmiştir. Bu mimari, evrişimli katmanlar aracılığıyla zaman–frekans uzayındaki yerel örüntüleri öğrenirken, LSTM katmanları sayesinde bu örüntülerin zamansal evrimini modellemeyi hedeflemektedir.

Önerilen hibrit mimarinin katkısını daha açık biçimde ortaya koymak ve kapsamlı bir karşılaştırmalı analiz gerçekleştirmek amacıyla çalışmada alternatif model yapılandırmaları da uygulanmıştır. Bu kapsamda, yalnızca Mel-spektrogram girdileri kullanılarak eğitilen bağımsız bir evrişimsel sinir ağı (CNN) modeli değerlendirilmiş ve saf spektral temsiller üzerinden uzamsal örüntü öğrenmenin performansı analiz edilmiştir.

Buna ek olarak, sıralı özellik temsillerini kullanan bağımsız bir LSTM modeli uygulanarak zamansal bağımlılıkların tek başına modellenmesinin sınıflandırma performansına etkisi incelenmiştir. Bu model, özellikle kardiyak döngü boyunca ortaya çıkan ardışık yapıların temsil gücünü değerlendirmek amacıyla kullanılmıştır.

Ayrıca, Mel-spektrogram temsilleri ile el ile çıkarılmış geleneksel ses özelliklerini eş zamanlı olarak işleyen çok girdili bir CNN mimarisi tasarlanmış ve test edilmiştir. Bu yaklaşım, farklı özellik türlerinin erken veya orta düzeyde birleştirilmesinin (feature fusion) sınıflandırma performansı üzerindeki etkisini incelemeyi amaçlamaktadır.

Bu farklı model yapılandırmalarının birlikte değerlendirilmesi, önerilen hibrit CNN–LSTM mimarisinin hem spektral hem de zamansal bilgiyi bütüncül biçimde modellemesinin sağladığı performans kazanımını nicel olarak ortaya koymayı hedeflemektedir.

3.1.5. Eğitim Süreci ve Hiperparametre Ayarlaması

Bilgi sızıntısını önlemek amacıyla veri kümesi, hasta düzeyinde tabakalı (stratified) bir yaklaşım kullanılarak eğitim, doğrulama ve test alt kümelerine ayrılmıştır. Bu sayede aynı hastaya ait kayıtların yalnızca tek bir alt kümede yer alması sağlanmış ve modelin performansının gerçekçi biçimde değerlendirilmesi hedeflenmiştir.

Tüm modeller, Adam optimize edicisi kullanılarak mini-batch gradyan inişi yöntemiyle eğitilmiştir. Eğitim sürecinde aşırı öğrenmeyi (overfitting) önlemek ve modelin genelleme yeteneğini artırmak amacıyla erken durdurma (early stopping) ve öğrenme oranı zamanlaması (learning rate scheduling) tekniklerinden yararlanılmıştır.

Hiperparametreler, doğrulama kümesi üzerindeki performans metrikleri temel alınarak sistematik biçimde ayarlanmış ve her model için en uygun yapılandırma belirlenmiştir.

3.1.6. Değerlendirme ve Karşılaştırma

Bu çalışmada geliştirilen modellerin performansı, tıbbi yapay zekâ sistemlerinin değerlendirilmesinde yaygın olarak kabul edilen ve klinik açıdan anlamlı olan nicel ölçütler kullanılarak kapsamlı biçimde analiz edilmiştir. Bu kapsamda sınıflandırma başarımı; doğruluk (accuracy), kesinlik (precision), duyarlılık (recall/sensitivity), özgüllük (specificity), F1 skoru ve alıcı işletim karakteristiği (ROC) eğrisi altında kalan alan (AUC) ölçütleri üzerinden değerlendirilmiştir. Bu metrikler, özellikle dengesiz sınıf dağılımlarının bulunduğu tıbbi veri kümelerinde model performansının çok boyutlu olarak incelenmesine olanak tanımaktadır.

Nicel ölçütlere ek olarak, modellerin hata profillerini ve sınıflar arasındaki karışımı daha ayrıntılı biçimde analiz edebilmek amacıyla karmaşıklık matrisleri (confusion matrices) incelenmiştir. Bu analizler, yanlış pozitif (FP) ve yanlış negatif (FN) kararların dağılımını ortaya koyarak, modellerin tarama (screening) senaryolarında klinik açıdan taşıyabileceği risklerin değerlendirilmesine imkân sağlamıştır.

Ayrıca eğitim ve doğrulama öğrenme eğrileri kullanılarak modellerin yakınsama davranışı, aşırı uyum (overfitting) eğilimi ve genelleme kabiliyeti sistematik biçimde analiz edilmiştir.

Önerilen hibrit CNN–LSTM mimarisinin performans açısından görece üstünlüğünü ortaya koymak amacıyla model; el ile çıkarılmış özelliklere dayalı geleneksel makine öğrenimi sınıflandırıcıları ile literatürde yer alan gelişmiş transfer öğrenme tabanlı derin öğrenme mimarileriyle karşılaştırmalı olarak değerlendirilmiştir. Tüm karşılaştırmalar, hasta düzeyinde veri bölme stratejisi korunarak ve tüm modeller için aynı değerlendirme ölçütleri kullanılarak gerçekleştirilmiştir. Bu yaklaşım, farklı model türleri arasında adil, tutarlı ve tekrarlanabilir bir karşılaştırma yapılmasını sağlamaktadır.

Bu bölümde tanımlanan değerlendirme çerçevesi, Bölüm 4'te sunulan deneysel sonuçların doğru ve anlamlı biçimde yorumlanabilmesi için sağlam bir metodolojik temel oluşturmaktadır. Ayrıca önerilen hibrit mimarinin güçlü ve sınırlı yönlerinin yalnızca sayısal performans ölçütleri üzerinden değil, klinik bağlam ve potansiyel uygulama senaryoları dikkate alınarak bütüncül bir perspektifle değerlendirilmesine olanak tanımaktadır.

3.2. Veri Kümesi Tanımı

3.2.1. Pediatrik Kalp Sesi Veri Kümesinin Kaynağı

Bu çalışmada kullanılan tüm FKG kayıtları, PhysioNet platformunda barındırılan büyük ölçekli ve herkese açık bir veri tabanı olan CirCor DigiScope veri kümesinden (sürüm 1.0.x) elde edilmiştir. Söz konusu veri kümesi, Brezilya'nın kuzeydoğusunda yürütülen geniş kapsamlı pediatrik kardiyak tarama kampanyaları sırasında toplanmış olup, George B. Moody PhysioNet/Computing in Cardiology (CinC) Challenge 2022 için resmi veri kaynağı olarak kullanılmıştır (Reyna vd., 2023).

CirCor DigiScope veri kümesi, 0–21 yaş aralığındaki 1.568 pediatrik denekten elde edilen ve bir veya daha fazla oskültasyon bölgesinden kaydedilmiş toplam 5.272 kalp sesi kaydını içermektedir. Kayıtlar; aort (AV), pulmoner (PV), triküspit (TV) ve mitral (MV) kapak alanları olmak üzere dört standart oskültasyon bölgesinden elde edilmiştir. Bu çoklu konum yaklaşımı, kalp üfürümlerinin anatomik yayılımını ve bölgeye özgü akustik özelliklerini daha ayrıntılı biçimde yakalamaya olanak tanımaktadır.

Her bir FKG kaydı, klinik açıdan zengin ve ayrıntılı anotasyonlarla birlikte sunulmaktadır. Bu anotasyonlar; üfürüm varlığının *yok*, *var* veya *bilinmiyor* şeklinde sınıflandırılmasını; üfürümün sistolik, diyastolik veya sürekli olarak zamanlamasının belirlenmesini; üfürümün morfolojik özellikleri ile akustik karakteristiklerinin tanımlanmasını ve Levine derecesinin değerlendirilmesini kapsamaktadır. Ayrıca oskültasyonun gerçekleştirildiği anatomik konumlar açık biçimde belirtilmiş ve birinci (S1) ile ikinci (S2) kalp sesleri uzman doğrulamasıyla otomatik olarak bölümlendirilmiş şekilde sunulmuştur.

Bu ayrıntılı etiketleme yapısı, CirCor DigiScope veri kümesini yalnızca algoritmik sınıflandırma açısından değil, aynı zamanda klinik karar destek sistemlerinin geliştirilmesi açısından da son derece değerli bir kaynak hâline getirmektedir. Özellikle pediatrik popülasyona odaklanması ve çoklu oskültasyon bölgelerini içermesi, veri kümesini gerçek dünya tarama ve triyaj senaryoları için geliştirilen yapay zekâ sistemlerinin değerlendirilmesinde güçlü ve güvenilir bir referans veri kaynağı hâline getirmektedir.

3.2.2. Sınıf Dağılımı

Veri artırma uygulanmadan önce CirCor DigiScope veri kümesi, klinik tarama senaryolarında sıklıkla karşılaşılan tipik bir sınıf dengesizliği sergilemektedir. Özellikle pediatrik tarama çalışmalarında normal (üfürüm yok) kayıtların patolojik (üfürüm var)

vakalara kıyasla daha yüksek oranda bulunması beklenen bir durumdur. Bu durum, gerçek dünya klinik prevalansını yansıtmakla birlikte, otomatik sınıflandırma sistemlerinde model yanlılığı (bias) oluşma riskini de beraberinde getirmektedir.

Özellik çıkarımına yönelik olarak geliştirilen betikler aracılığıyla elde edilen özet istatistiklere dayalı analiz sonucunda, veri artırma ve uzunluk normalizasyonu işlemlerinin ardından segment düzeyindeki sınıf dağılımı belirlenmiştir. Buna göre, “üfürüm yok” sınıfına ait 695 segment, “üfürüm var” sınıfına ait 347 segment ve “üfürüm bilinmiyor” sınıfına ait 68 segment bulunmaktadır. Bu süreç sonunda, başlangıçta dört oskültasyon bölgesinden eksiksiz kayda sahip 574 pediatrik hastaya ait verilerden, her biri 5 saniye uzunluğunda olacak şekilde toplam 1.110 FKG segmenti oluşturulmuştur.

“Bilinmiyor” (unknown) etiketiyle işaretlenen segmentler, denetimli öğrenme tabanlı model eğitimi ve nicel performans değerlendirme aşamalarına dâhil edilmemiştir. Bununla birlikte bu segmentler, veri tutarlılığının doğrulanması ve sınırdaki yer alan vakaların nitel olarak incelenmesi açısından referans olarak değerlendirilmiştir.

Çekirdek deneylerde kullanılan nihai eğitim, doğrulama ve test yapılandırması, hasta düzeyinde tabakalı (stratified) bölme stratejisi kullanılarak oluşturulmuştur. Bu yaklaşım, aynı hastaya ait segmentlerin farklı alt kümelere dağılmasını engelleyerek bilgi sızıntısı (data leakage) riskini ortadan kaldırmayı amaçlamaktadır.

Bu kapsamda veri setinin alt kümelere dağılımı aşağıdaki gibidir:

- Eğitim kümesi: 510 segment (%47 üfürüm pozitif)
- Doğrulama kümesi: 124 segment (%27 üfürüm pozitif)
- Test kümesi: 124 segment (%24 üfürüm pozitif)

Bu bölümlendirme stratejisi, gerçek klinik sınıf dağılımını mümkün olduğunca korurken, geliştirilen modelin hem patolojik hem de normal vakalar üzerindeki genelleme kabiliyetinin güvenilir ve tarafsız bir biçimde değerlendirilmesine olanak tanımaktadır.

3.2.3. Etik Hususlar

Bu çalışmada kullanılan CirCor DigiScope veri kümesi, PhysioNet platformu üzerinden kamuya açık olarak yayımlanmadan önce, ilgili yerel etik kurullar tarafından onaylanan pediatrik tarama çalışmaları kapsamında toplanmıştır (Reyna vd., 2023). Veri kümesi, yayımlanma aşamasında tamamen anonimleştirilmiş olup doğrudan hasta tanımlayıcılarına (isimler, kesin doğum tarihleri, adresler veya iletişim bilgileri) yer verilmemektedir.

Bu tez kapsamında yalnızca veri kümesinin kamuya açık ve anonimleştirilmiş sürümü kullanılmıştır. Yeniden kimliklendirmeye yönelik herhangi bir girişimde

bulunulmamış; bireysel hasta bilgilerine ulaşma veya çıkarım yapma amacı güdülmemiştir. Tüm analizler, önceden kaydedilmiş FKG sinyalleri üzerinde çevrimdışı (offline) olarak gerçekleştirilmiş olup, hastalar, klinisyenler veya aktif klinik karar verme süreçleriyle herhangi bir doğrudan etkileşim söz konusu değildir.

Tıpta yapay zekâya yönelik güncel etik ve metodolojik öneriler doğrultusunda, bu tezde elde edilen bulgular klinik kullanıma hazır bir tanı aracı olarak değil, bir kavram kanıtı (proof of concept) araştırması olarak değerlendirilmelidir. Sunulan sonuçların gerçek dünya klinik ortamlarına aktarılabilmesi için farklı merkezlerde ek doğrulama çalışmaları, düzenleyici kurumlar tarafından yapılacak incelemeler ve klinik iş akışlarıyla entegrasyonu içeren kapsamlı değerlendirme süreçleri gereklidir.

Bu çerçevede çalışma, veri gizliliği, hasta güvenliği ve sorumlu yapay zekâ geliştirme ilkeleri gözetilerek etik standartlara uygun biçimde yürütülmüştür.

3.3. Veri Ön İşleme İş Akışı

Ön işleme iş akışı, heterojen yapıya ve değişken uzunluklara sahip ham FKG kayıtlarını, sağlam ve karşılaştırılabilir özelliklerin çıkarılmasına olanak tanıyan standartlaştırılmış segmentlere dönüştürmeyi amaçlamaktadır. Bu kapsamda tüm sinyaller, sabit uzunlukta (5 saniye) segmentler hâline getirilerek model eğitime uygun bir biçimde yeniden yapılandırılmıştır.

Bu süreçte, sinyal kalitesinin artırılması ve veri temsillerinin tutarlılığının sağlanması amacıyla bir dizi ardışık ön işleme adımı uygulanmıştır. Öncelikle gürültü ve artefaktların etkisini azaltmaya yönelik filtreleme ve normalizasyon işlemleri gerçekleştirilmiş; ardından değişken uzunlukta kayıtlar sabit uzunluklu segmentlere dönüştürülmüştür. Bu segmentasyon yaklaşımı, modellerin giriş boyutlarını standartlaştırırken aynı zamanda veri çeşitliliğini artırmayı hedeflemektedir.

Buna ek olarak, sınıf dengesizliğinin etkilerini azaltmak ve modelin genelleme kabiliyetini güçlendirmek amacıyla veri artırma (data augmentation) stratejileri uygulanmıştır.

Bu bölümde, ön işleme iş akışını oluşturan gürültü azaltma, uzunluk normalizasyonu, segmentasyon ve veri artırma adımları ayrıntılı ve sistematik bir biçimde ele alınmaktadır.

3.3.1. Gürültüyle Başa Çıkma ve Sinyal Bütünlüğü

CirCor DigiScope veri kümesindeki FKG kayıtları, gerçek klinik tarama ortamlarında elde edilmiş olup kaçınılmaz olarak çeşitli arka plan gürültüleri

içermektedir. Bu gürültüler; ortam sesleri, konuşmalar, solunum etkileri ve hasta hareketlerinden kaynaklanan artefaktları kapsamaktadır. Kalp sesleri ve üfürümler genellikle düşük genlikli ve zamana duyarlı akustik bileşenler içerdiğinden, agresif gürültü giderme teknikleri tanısal açıdan önemli bilgilerin kaybına yol açabilmektedir.

Bu nedenle çalışmada, aşırı gürültü bastırma yaklaşımları yerine sinyal bütünlüğünü korumayı önceliklendiren temkinli bir ön işleme stratejisi benimsenmiştir. Bu strateji kapsamında tüm FKG sinyalleri, yalnızca kayan nokta normalizasyonu gibi asgari düzeyde dahili ön işleme uygulayan *librosa.load* fonksiyonu aracılığıyla ve özgün örnekleme hızları korunarak yüklenmiştir. Küresel bant geçiren filtreleme, dalgacık tabanlı gürültü giderme veya adaptif filtreleme gibi yöntemler bilinçli olarak uygulanmamış; gürültüye karşı dayanıklılık doğrudan ön işleme aşamasında değil, özellik çıkarımı ve modelleme düzeyinde ele alınmıştır.

Bu yaklaşım doğrultusunda, Mel-spektrogram temsilleri yüksek frekanslı ve kısa süreli dalgalanmaları doğal biçimde yumuşatarak spektral kararlılığı artırmaktadır. Benzer şekilde enerji tabanlı ölçümler ve MFCC gibi toplulaştırılmış geleneksel özellikler, kısa zamanlı çerçeveler üzerinde ortalama alındıkları için geçici artefaktların etkisini azaltmaktadır. Ayrıca Bölüm 3.3.4'te ayrıntıları verilen veri artırma stratejileri, modelleri her bir üfürüm örüntüsünün birden fazla gerçekleşimiyle karşı karşıya bırakarak gürültüye karşı örtük bir sağlamlık kazandırmaktadır.

Bu tasarım tercihi, çağdaş kalp sesi analizi literatürüyle uyumlu olup, açık gürültü bastırma teknikleri yerine sağlam spektral–zamansal temsillere dayalı öğrenmeyi esas alan veri odaklı bir yaklaşımı yansıtmaktadır.

3.3.2. Kalp Sesi Segmentasyonu (S1 , S2 , Sistol , Diyastol)

CirCor DigiScope veri kümesi, yarı denetimli algoritmalar ve uzman klinisyenler tarafından gerçekleştirilen manuel doğrulama süreçleri sonucunda elde edilmiş yüksek kaliteli S1 ve S2 kalp sesi segmentasyon anotasyonları içermektedir (Reyna vd., 2023). Bu anotasyonlar, kalp döngüsünün fizyolojik yapısını doğru biçimde yansıtan güvenilir referans noktaları sağlamaktadır.

Mevcut çalışmada bu anotasyonlar, doğrudan atım (beat) düzeyinde zorunlu segmentasyon gerçekleştirmek amacıyla kullanılmamış; bunun yerine spektral–zamansal örüntülerin yorumlanmasını destekleyen bağlamsal bilgiler olarak değerlendirilmiştir. Bu doğrultuda her bir FKG kaydı, tek tek kardiyak döngülere ayrılmak yerine genellikle birden fazla atımı kapsayan sabit uzunluklu (5 saniye) pencereler üzerinden işlenmiştir.

Bu yaklaşımın tercih edilmesinin başlıca gerekçeleri şu şekilde özetlenebilir. Öncelikle Mel-spektrogram temsilleri, her bir pencere içerisinde tekrarlayan S1–sistol–S2–diastol örüntülerini açık ve ayırt edilebilir biçimde yansıtmaktadır. Buna ek olarak kullanılan CNN–LSTM mimarisi, açık bir döngü hizalamasına ihtiyaç duymaksızın bu örüntüler arasındaki zamansal bağımlılıkları etkili biçimde modelleyebilmektedir.

Ayrıca, atım düzeyinde ayrıntılı segmentasyonun gerektirdiği ek ön işleme adımlarından kaçınılarak işlem karmaşıklığı azaltılmış ve büyük ölçekli veri kümeleri üzerinde daha kararlı ve tekrarlanabilir bir eğitim süreci elde edilmiştir. Bu yaklaşım, veri hazırlama sürecindeki hata birikimini de sınırlayarak model performansının daha güvenilir biçimde değerlendirilmesine katkı sağlamaktadır.

Bu tasarım tercihi, sabit uzunluklu FKG segmentleri kullanan ve üfürüm tespitine odaklanan güncel derin öğrenme tabanlı çalışmalarla yöntemsel açıdan uyumludur.

3.3.3. Yeniden Örnekleme ve Uzunluk Normalizasyonu

CirCor DigiScope veri kümesindeki FKG kayıtları, dört oskültasyon bölgesinin tamamı mevcut olsa dahi süre açısından yüksek değişkenlik göstermektedir. Kayıt süreleri yaklaşık 4,8 saniye ile 80,4 saniye arasında değişmektedir. Bu durum, mini-batch tabanlı derin öğrenme eğitimini zorlaştırmakta ve özellik matrislerinin tutarlı biçimde oluşturulmasını engellemektedir.

Bu nedenle tüm FKG sinyalleri yeniden örnekleme ve uzunluk normalizasyonu işlemlerine tabi tutulmuştur. Bu kapsamda tüm ses verileri, `feature_extraction_ML.py` ve `mel_spectrogram_deep_learning.py` dosyalarında tanımlanan ve $f_s = 22,050\text{Hz}$ olarak belirlenen örnekleme hızı kullanılarak temsil edilmiştir.

Yeniden örnekleme işlemi, sürekli zamanlı sinyalin ayrık zamanlı forma dönüştürülmesi sürecine dayanmakta olup aşağıdaki şekilde ifade edilebilir:

Eşitlik 16. Yeniden Örnekleme (Resampling)

$$x[n] = x(t) \big|_{t=\frac{n}{f_s}} \quad (\text{Eşitlik 16})$$

Her bir FKG segmenti için 5 saniyelik sabit bir hedef süre seçilmiş olup bu süre aşağıdaki gibi hesaplanmaktadır:

Eşitlik 17. Hedef Sinyal Uzunluğu (Target Length)

$$N = f_s \times T = 22,050 \times 5 = 110,250 \quad (\text{Eşitlik 17})$$

Beş saniyelik pencere uzunluğu; birden fazla kardiyak döngünün aynı segment içerisinde yakalanmasına olanak sağlaması, hesaplama maliyetini makul düzeyde tutması ve daha kısa kayıtlar için aşırı dolgu (padding) gereksinimini azaltması açısından dengeli bir çözüm sunmaktadır.

Kırpma ve sıfır dolgu işlemleri `set_len_extr` fonksiyonu aracılığıyla gerçekleştirilmiştir. Bu işlem matematiksel olarak aşağıdaki gibi tanımlanabilir:

Eşitlik 18. Uzunluk Normalizasyonu (Truncation ve Padding)

$$x_{norm}[n] = \begin{cases} x[n], & n < N \\ 0, & n \geq N \end{cases} \quad (\text{Eşitlik 18})$$

Normalizasyon öncesinde AV, MV, PV ve TV oskültasyon bölgelerine ait ham kayıtların uzunluk dağılımları histogramlar yardımıyla görselleştirilmiş ve hedef uzunluk olan $N = 110,250$ örnek değeri kırmızı kesikli çizgi ile gösterilmiştir. Uzunluk normalizasyonu sonrasında ise tüm sinyaller aynı örnek sayısına sahip olacak şekilde hizalanmış; böylece özellik çıkarımı ve CNN tabanlı modeller için uygun biçimde matris yapısında istiflenebilir hâle getirilmiştir.

3.3.4. Veri Artırma

Sınıf dengesizliğini azaltmak ve modelin genellenebilirliğini artırmak amacıyla, veri ön işleme iş akışında özellikle üfürüm pozitif hastalara yönelik segment tabanlı bir veri artırma stratejisi uygulanmıştır.

Her bir üfürüm pozitif hastada, dört kapak kaydının her birinin en az 15 saniye uzunluğa sahip olması durumunda `set_len_extr` fonksiyonu ilgili kapak sinyallerini örtüşmeyen üç adet 5 saniyelik segmente ayırmaktadır. Bu segmentler sırasıyla 0–5 saniye, 5–10 saniye ve 10–15 saniye zaman aralıklarını kapsamaktadır.

Bu segmentasyon işlemi matematiksel olarak aşağıdaki şekilde ifade edilebilir:

Eşitlik 19. Segment Tabanlı Veri Artırma

$$x_k[n] = x[n + k \cdot N], k = 0,1,2 \quad (\text{Eşitlik 19})$$

Burada N her bir segmentin uzunluğunu (110{,}250 örnek) ifade etmektedir.

Bu işlem sırasında ilk segment (0–5 saniye) özgün örnek olarak birincil veri kümesine dâhil edilirken, ikinci ve üçüncü segmentler (5–10 saniye ve 10–15 saniye) ek artırılmış örnekler olarak kullanılmaktadır.

Uygun her bir hasta için bu yöntem, AV, MV, PV ve TV oskültasyon bölgelerine ait kayıtların senkronize bir setini içeren iki ek artırılmış örnek üretmektedir. Bu kriterleri karşılayan toplam 84 üfürüm pozitif hasta bulunmuş ve böylece 168 ek örnek elde edilmiştir. Söz konusu artırma süreci, Mel-spektrogram çıkarım betiğinin konsol çıktılarında açık biçimde doğrulanmıştır. Çıktı mesajında “Artırılan hasta sayısı: 84 | Yeni örnekler: 168 | veri artırma sonrası veri boyutu: 1110” ifadesi yer almaktadır.

Ana deneylerde zaman esnetme (time stretching), perde kaydırma (pitch shifting) veya yapay gürültü ekleme gibi sentetik dönüşümler uygulanmamıştır. Bununla birlikte bu tür işlemler, MelSpectrogramProcessor sınıfı içerisinde yer alan augment_audio yöntemi tarafından desteklenmekte olup gelecekteki çalışmalar için esnek bir altyapı sunmaktadır.

Fizyolojik olarak geçerli alt segmentlere dayanan bu yaklaşım, veri çeşitliliğini artırırken üfürümlere ait gerçek spektral imzaların korunmasına katkı sağlamaktadır.

3.4. Özellik Çıkarma Yöntemleri

3.4.1. -Mel-Spektrogram Üretimi

Mel-spektrogramlar, frekans ekseninin insan işitsel algısına yaklaşık olarak karşılık gelen mel ölçeğine dönüştürülmesiyle elde edilen zaman–frekans temsilleridir. Bu gösterim, S1 ve S2 kalp sesleri ile üfürümlerin hem zamansal konumlarını hem de spektral içeriklerini kompakt bir biçimde kodladığından kalp sesi analizinde yaygın olarak tercih edilen bir temsil hâline gelmiştir.

Bu çalışmada Mel-spektrogram temsilleri, librosa.feature.melspectrogram fonksiyonunu saran MelSpectrogramExtractor sınıfı kullanılarak hesaplanmıştır. Özellik çıkarımı sürecinde örnekleme hızı $f_s = 22,050\text{Hz}$, mel bant sayısı 128, FFT pencere boyutu 2048 örnek ve adım uzunluğu (hop length) 512 örnek olarak belirlenmiştir. Beş saniyeye normalize edilmiş her bir FKG dalga formu için STFT yaklaşık 93 ms’lik bir zaman penceresi ve 23 ms’lik bir adım uzunluğu ile uygulanmaktadır.

Mel-spektrogram hesaplamasının ilk adımı Kısa Zamanlı Fourier Dönüşümü’dür (STFT):

Eşitlik 20. Kısa Zamanlı Fourier Dönüşümü (STFT)

$$X(m, k) = \sum_{n=0}^{N-1} x[n] \cdot w[n - m] \cdot e^{-j2\pi kn/N} \quad (\text{Eşitlik 20})$$

Burada $w[n]$ pencere fonksiyonunu, m zaman kaymasını ve k frekans indeksini ifade etmektedir.

Elde edilen güç spektrumu, insan işitme sistemine daha yakın bir ölçek olan mel frekans ölçeğine dönüştürülmektedir:

Eşitlik 21. Mel Frekans Dönüşümü

$$f_{mel} = 2595 \cdot \log_{10} \left(1 + \frac{f}{700} \right) \quad (\text{Eşitlik 21})$$

Bu dönüşüm sonrasında spektrum, 128 mel bandına projekte edilmekte ve ardından logaritmik (dB) ölçeğe dönüştürülmektedir:

Eşitlik 22. Logaritmik Güç Spektrumu (dB)

$$S_{dB} = 10 \cdot \log_{10}(S) \quad (\text{Eşitlik 22})$$

Bu işlem sonucunda her bir FKG segmenti için tipik olarak 128×216 boyutlarında iki boyutlu bir Mel-spektrogram temsili elde edilmektedir.

Her bir kapak (AV, MV, PV, TV) ve etiket kombinasyonu için üretilen spektrogramlar incelendiğinde, üfürüm var ve üfürüm yok vakalar arasında belirgin enerji dağılımı farklılıkları gözlemlenmiştir. Özellikle patolojik üfürümlerde daha geniş bantlı ve düzensiz enerji yayılımları dikkat çekmektedir.

CNN tabanlı mimarilere uygun girdi formatını elde edebilmek amacıyla her bir Mel-spektrograma ek bir kanal boyutu eklenmiştir:

```
mel_specs = np.expand_dims(mel_specs, axis=-1) # şekil: (N, 128, 216, 1)
```

Bu adım, iki boyutlu evrişim katmanları için gerekli olan dört boyutlu tensör yapısının oluşturulmasını sağlamaktadır.

3.4.2. Gelişmiş Mel-spektrogram Özellikleri

Ham Mel-spektrogram görüntülerine ek olarak, bu çalışmada mel temsillerinden türetilen üst düzey özetleyici tanımlayıcılar da hesaplanmıştır. Bu tanımlayıcılar, `extract_advanced_features` yöntemi aracılığıyla her bir spektrogramın spektral ve zamansal özelliklerini daha kompakt ve ayırt edici bir biçimde temsil etmeyi amaçlamaktadır.

Her bir Mel-spektrogram temsili için birden fazla tamamlayıcı özellik grubu çıkarılmaktadır. Bu kapsamda küresel istatistikler, tüm mel katsayıları üzerinden

hesaplanan ortalama, standart sapma, minimum, maksimum ve dinamik aralık deęerlerini içermektedir. Ortalama ve standart sapma ařaęıdaki gibi tanımlanmaktadır:

Eřitlik 23. Ortalama (Mean)

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (\text{Eřitlik 23})$$

Eřitlik 24. Standart Sapma (Standard Deviation)

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2} \quad (\text{Eřitlik 24})$$

Bant-özel enerji ölçümleri ise düşük (0–31), orta (32–95) ve yüksek (≥ 96) frekans bantlarında hesaplanan ortalama enerji deęerleri ile düşük ve yüksek frekans bantları arasındaki enerji oranını kapsamaktadır. Bu oran ařaęıdaki řekilde ifade edilebilir:

Eřitlik 25. Bant Enerji Oranı (Low/High Energy Ratio)

$$R = \frac{\sum_{i \in \text{low}} E_i}{\sum_{i \in \text{high}} E_i} \quad (\text{Eřitlik 25})$$

Bu oran, düşük perdeli fonksiyonel üfürümler ile yüksek perdeli patolojik üfürümler arasındaki ayrımı vurgulayan ayırt edici bir ölçüt olarak deęerlendirilmektedir.

Zamansal tanımlayıcılar, zaman boyunca ortalaması alınmış mel enerjisinin varyans, çarpıklık (skewness) ve basıklık (kurtosis) istatistiklerini içermekte olup enerjinin 5 saniyelik pencere içerisindeki dalgalanma yapısını yansıtmaktadır. Varyans ařaęıdaki řekilde tanımlanmaktadır:

Eřitlik 26. Varyans (Variance)

$$\text{Var}(x) = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2 \quad (\text{Eřitlik 26})$$

Buna ek olarak mel spektral merkez (centroid) istatistikleri, enerji daęılımının mel bantları üzerindeki “kütle merkezi”ni temsil etmektedir:

Eřitlik 27. Mel Spektral Merkez (Centroid)

$$C = \frac{\sum_i f_i \cdot s_i}{\sum_i s_i} \quad (\text{Eřitlik 27})$$

Burada f_i frekans bandını, S_i ise ilgili banttaki enerjiyi temsil etmektedir.

Elde edilen tüm bu özellikler kompakt bir özellik vektörü hâlinde birleştirilmekte ve tek başına klasik makine öğrenmesi tabanlı sınıflandırıcılarda kullanılabildiği gibi, tam Mel-spektrogram girdileriyle birlikte çok girdili derin öğrenme mimarilerinde de değerlendirilebilmektedir.

3.4.3. Geleneksel Ses Özellikleri

Mel tabanlı temsillere tamamlayıcı olarak, feature_extraction_ML.py dosyasında tanımlanan fonksiyonlar kullanılarak dört kapak sinyalinin geniş bir geleneksel tek boyutlu (1B) ses özellikleri kümesi çıkarılmıştır. Bu özellikler, sinyalin farklı fiziksel ve istatistiksel yönlerini yansıtan çok boyutlu ve ayırt edici bir tanımlayıcı uzay oluşturmaktadır.

Özellikler, fizyolojik ve akustik özellikleri bütüncül biçimde temsil edecek şekilde birden fazla ana kategori altında toplanmaktadır. Zaman alanı ve enerji özellikleri; genlik zarfına ilişkin istatistikleri, RMS enerjisini ve toplam sinyal enerjisini kapsamaktadır. RMS enerji aşağıdaki şekilde tanımlanmaktadır:

Eşitlik 28. RMS Enerji (Root Mean Square Energy)

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} \quad (\text{Eşitlik 28})$$

Toplam sinyal enerjisi ise şu şekilde ifade edilmektedir:

Eşitlik 29. Toplam Enerji (Signal Energy)

$$E = \sum_{i=1}^N x_i^2 \quad (\text{Eşitlik 29})$$

Bu özellikler, üfürüm şiddeti ve zamansal yoğunluk hakkında doğrudan bilgi sağlamaktadır.

Sıfır geçiş ve onset tabanlı özellikler, ZCR ve onset sayısı üzerinden türbülanslı akışın ve belirgin akustik olayların yoğunluğunu yansıtmaktadır. ZCR aşağıdaki şekilde tanımlanmaktadır:

Eşitlik 30. Sıfır Geçiş Oranı (Zero Crossing Rate – ZCR)

$$ZCR = \frac{1}{N-1} \sum_{i=1}^{N-1} \mathbb{I}(x_i \cdot x_{i+1} < 0) \quad (\text{Eşitlik 30})$$

Burada İlgösterge fonksiyonudur.

Spektral özellikler ise spektral merkez, bant genişliği, baskın frekans ve bant enerji oranlarını içermekte; böylece üfürümlerin dar bantlı mı yoksa geniş bantlı mı olduğunun ayırt edilmesine katkı sunmaktadır.

İstatistiksel şekil özellikleri, ham dalga formunun çarpıklık (skewness) ve basıklık (kurtosis) ölçümleri aracılığıyla genlik dağılımının simetrisi ve tepe yapısı hakkında bilgi vermektedir.

Kepstral özellikler kapsamında hesaplanan mel frekans kepsral katsayıları (MFCC), kalp sesi sınıflandırması literatüründe yaygın biçimde doğrulanmış ayırt edici tanımlayıcılar arasında yer almaktadır. MFCC hesaplaması genel olarak log mel spektrumunun ters Fourier dönüşümü ile elde edilmektedir:

Eşitlik 31. MFCC Hesaplaması

$$MFCC_k = \sum_n \log(S_{mel}[n]) \cdot \cos\left(\frac{\pi k(n+0.5)}{N}\right) \quad (\text{Eşitlik 31})$$

Son olarak otokorelasyon özellikleri, otokorelasyon fonksiyonundan türetilen istatistikler aracılığıyla kalp hızı ve ritimle ilişkili periyodisiteyi yakalamayı hedeflemektedir:

Eşitlik 32. Otokorelasyon (Autocorrelation)

$$R(\tau) = \sum_n x[n] \cdot x[n - \tau] \quad (\text{Eşitlik 32})$$

Özellik çıkarım sürecinin tamamlanmasının ardından sayısal olmayan sütunlar veri setinden kaldırılmış; eksik veya sonsuz değerler sıfırlarla doldurularak klasik makine öğrenmesi modelleri için uygun, tamamen sayısal bir özellik matrisi elde edilmiştir.

3.4.4. Hibrit Özellik Seçiminin Gerekçesi

Hibrit özellik tasarımı, spektral–zamansal temsillerin el ile tasarlanmış tanımlayıcılarla birleştirilmesinin kalp sesi analizinde hem sağlamlığı hem de yorumlanabilirliği artırdığı yönündeki literatür bulgularından esinlenmiştir. Mel-spektrogramlar, CNN’lerin karmaşık uzamsal örüntüleri otomatik olarak keşfetmesine olanak tanırken; geleneksel özellikler sinyal enerjisi, periyodisite ve spektral şekil hakkında klinik açıdan anlamlı özetler sunmaktadır.

Tekrarlılığı azaltmak ve en ayırt edici tanımlayıcılara odaklanmak amacıyla klasik makine öğrenmesi modelleri için üç aşamalı bir özellik seçimi stratejisi uygulanmıştır.

İlk aşamada, ℓ_1 düzenlileştirmeye dayalı Lasso lojistik regresyonu kullanılarak katkısı düşük olan özelliklerin katsayıları sifıra yaklaştırılmış ve böylece seyrek bir özellik alt kümesi elde edilmiştir:

Eşitlik 33. Lasso Lojistik Regresyon

$$\min_{\beta} \left[- \sum_{i=1}^N (y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)) + \lambda \sum_{j=1}^p |\beta_j| \right] \quad (\text{Eşitlik 33})$$

Burada λ düzenlileştirme parametresini temsil etmektedir.

İkinci aşamada, ANOVA F-değeri tabanlı filtreleme yöntemi uygulanarak sınıflar arası ayırım gücü yüksek olan özellikler istatistiksel olarak seçilmiştir:

Eşitlik 34. ANOVA F-İstatistiği

$$F = \frac{\text{Gruplar arası varyans}}{\text{Gruplar içi varyans}} \quad (\text{Eşitlik 34})$$

Son aşamada ise özyinelemeli özellik eleme (Recursive Feature Elimination – RFE) yöntemi kullanılarak model performansına katkısı en yüksek olan özellikler yinelemeli biçimde belirlenmiştir:

Eşitlik 35. RFE Özellik Seçimi (Genel İfade)

$$\mathcal{F}_{k+1} = \mathcal{F}_k \setminus \{\arg \min_j w_j\} \quad (\text{Eşitlik 35})$$

Burada w_j , model tarafından öğrenilen özellik önem ağırlığını temsil etmektedir.

Bu çok aşamalı süreç sonucunda yaklaşık 20 adet en bilgilendirici özellik korunmuş; söz konusu özellikler hem klasik sınıflandırıcıların eğitilmesinde hem de çok girdili derin öğrenme mimarilerinde CNN tabanlı gömülere eklenerek hibrit bir temsil oluşturmak amacıyla kullanılmıştır.

3.5. Model Mimarisinin Tasarımı

Bu bölümde, pediatrik FKG kayıtlarından kalp üfürümlerinin otomatik olarak tespit edilmesine yönelik geliştirilen sinir ağı mimarileri ayrıntılı biçimde sunulmaktadır. Çalışmanın temel katkısını oluşturan hibrit CNN–LSTM modeli merkezde yer almakta

olup, yöntemin etkinliğini karşılaştırmalı olarak değerlendirebilmek amacıyla bağımsız CNN, bağımsız LSTM ve çok girdili derin öğrenme mimarileri de ele alınmaktadır.

Model tasarımı süreci, kalp sesi sinyallerinin doğasına özgü iki temel gereksinim dikkate alınarak yapılandırılmıştır:

- (i) Mel-spektrogram temsilleri üzerinde yerel spektral–zamansal örüntülerin etkin biçimde çıkarılması ve
- (ii) bu örüntülerin kardiyak döngüler boyunca sergilediği uzun vadeli zamansal bağımlılıkların modellenmesi.

Bu doğrultuda, evrişimli sinir ağlarının (CNN) güçlü yerel özellik çıkarma kapasitesi ile uzun-kısa süreli bellek ağlarının (LSTM) zamansal bağımlılıkları modelleme yeteneğinden birlikte yararlanan hibrit bir mimari yaklaşım benimsenmiştir. Bu yaklaşım, kalp sesi sinyallerinin hem frekans temelli yapısını hem de zamansal dinamiklerini bütüncül ve tamamlayıcı bir çerçevede ele almayı amaçlamaktadır.

3.5.1. Mel-spektrogramlar İçin CNN Mimarisi

Temel evrişimli sinir ağı (CNN) modeli, tek bir oskültasyon bölgesine (örneğin aort kapak alanına) ait $128 \times 216 \times 1$ boyutlarındaki Mel-spektrogram girdileri üzerinde doğrudan çalışacak şekilde tasarlanmıştır. Bu mimari, `DeepLearningModels.create_CNN_model` fonksiyonu içerisinde tanımlanmış olup artan soyutlama düzeylerine sahip üç evrişimsel bloktan ve ardından gelen yoğun (dense) katmanlardan oluşmaktadır.

Girdi katmanı, tek kanallı Mel-spektrogram tensörünü kabul ederek bu temsili ardışık evrişimsel bloklara iletmektedir. Mimari, hiyerarşik özellik çıkarımını destekleyecek biçimde üç evrişimsel bloktan oluşmaktadır.

Evrişim işlemi aşağıdaki şekilde tanımlanmaktadır:

Eşitlik 36. Evrişim (Convolution) İşlemi

$$Y(i, j) = \sum_m \sum_n X(i + m, j + n) \cdot K(m, n) \quad (\text{Eşitlik 36})$$

Birinci evrişimsel blok, düşük seviyeli spektral kenarların ve yerel enerji yoğunluklarının yakalanmasını hedeflemektedir. Bu blok; sırasıyla 3×3 çekirdek boyutuna sahip 32 filtreli bir Conv2D katmanı, Doğrultulmuş Doğrusal Birim (ReLU) aktivasyonu, Batch Normalization, ikinci bir Conv2D katmanı, MaxPooling ve Dropout katmanlarını içermektedir. ReLU aktivasyonu şu şekilde ifade edilir:

Eşitlik 37. ReLU Aktivasyon Fonksiyonu

$$f(x) = \max(0, x) \quad (\text{Eşitlik 37})$$

İkinci evrişimsel blok, mel bantları arasındaki daha karmaşık etkileşimleri ve orta düzey spektral örüntüleri öğrenmek üzere yapılandırılmıştır. Bu blokta filtre sayısı artırılarak daha zengin temsil kapasitesi sağlanmakta ve ardışık evrişim–normalizasyon–havuzlama işlemleriyle özellik haritaları kademeli olarak soyutlanmaktadır.

Üçüncü evrişimsel blok ise üfürümlere özgü daha soyut ve ayırt edici yapıları temsil eden yüksek seviyeli özelliklerin çıkarılmasını amaçlamaktadır. Artan filtre sayısı sayesinde modelin karmaşık spektral–zamansal örüntüleri yakalama kapasitesi güçlendirilmektedir.

Evrişimsel omurganın ardından elde edilen özellik haritaları Global Average Pooling katmanı ile özetlenmekte ve bu özet temsil sınıflandırma başlığına aktarılmaktadır. Sınıflandırma bölümü, çok katmanlı yoğun katmanlar aracılığıyla özelliklerin doğrusal olmayan kombinasyonlarını öğrenmektedir.

Son katmanda kullanılan sigmoid aktivasyon fonksiyonu, ikili sınıflandırma için olasılık çıktısı üretmektedir:

Eşitlik 38. Sigmoid Aktivasyon Fonksiyonu

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad (\text{Eşitlik 38})$$

Tanımlanan bu CNN mimarisi, hem bağımsız bir temel model (baseline) olarak değerlendirilmiş hem de hibrit CNN–LSTM mimarisinde evrişimsel özellik çıkarıcı omurga (backbone) olarak kullanılmıştır.

3.5.2. Zamansal Bağımlılıkların Modellenmesi için LSTM Modülü

Kalp sesleri, doğaları gereği ardışık ve periyodik olaylardan oluştuğundan, spektral temsillerin zaman içerisindeki evriminin modellenmesi kritik öneme sahiptir. Bu amaçla create_LSTM_model fonksiyonu kullanılarak bağımsız bir LSTM tabanlı mimari tasarlanmıştır.

Bu yapılandırmada Mel-spektrogram temsilleri veya özetlenmiş özellik dizileri, zaman adımı başına num_features olacak şekilde iki boyutlu zaman dizilerine dönüştürülmektedir. Elde edilen bu diziler, zamansal bağımlılıkların modellenmesi amacıyla ardışık LSTM katmanlarına beslenmektedir.

Uzun Kısa Süreli Bellek (LSTM) hücreleri, giriş, unutma ve çıkış kapıları aracılığıyla bilgiyi seçici biçimde koruyarak veya güncelleyerek zamansal bağımlılıkları modellemektedir. LSTM yapısı aşağıdaki temel denklemlerle ifade edilmektedir:

Eşitlik 39. Unutma Kapısı (Forget Gate)

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (\text{Eşitlik 39})$$

Eşitlik 40. Giriş Kapısı (Input Gate)

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (\text{Eşitlik 40})$$

Eşitlik 41. Hücre Durumu Güncellemesi (Cell State Update)

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (\text{Eşitlik 41})$$

Eşitlik 42. Çıkış Kapısı (Output Gate)

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (\text{Eşitlik 42})$$
$$h_t = o_t \odot \tanh(C_t)$$

Mimari, sırasıyla 128 birimli ve return_sequences=True ayarına sahip bir LSTM katmanı ile başlamakta ve ardından %30 oranında Dropout uygulanmaktadır. Bunu, yine return_sequences=True olarak yapılandırılmış 64 birimli ikinci bir LSTM katmanı ve Dropout izlemektedir. Üçüncü aşamada ise diziyi tek bir zamansal temsile özetleyen 32 birimli bir LSTM katmanı ve Dropout yer almaktadır.

Elde edilen temsil, 64 nöronlu Dense katmanı (ReLU aktivasyonu) ve Dropout aracılığıyla sınıflandırma başlığına aktarılmakta; son olarak sigmoid aktivasyonlu tek nöronlu Dense katmanı ile ikili sınıflandırma gerçekleştirilmektedir.

Bu saf tekrarlayan mimari, kardiyak döngüler boyunca zamanlama ve süre bilgilerini etkili biçimde öğrenebilmesine karşın giriş verisinde yer alan uzamsal olarak yapılandırılmış spektral bilgiyi doğrudan kullanamaması nedeniyle pratik uygulamalarda hibrit CNN-LSTM modeline kıyasla daha düşük sınıflandırma performansı sergilemiştir.

3.5.3. Hibrit CNN–LSTM Birleřtirme Stratejisi

Önerilen hibrit CNN–LSTM modeli, Mel-spektrogramlardaki yerel spektral–zamansal örüntüler ile uzun menzilli zamansal korelasyonları birlikte öğrenmeyi amaçlamaktadır. Bu model `create_CNN_LSTM_hybrid` fonksiyonu içerisinde uygulanmıştır.

Mimari akış, Mel-spektrogram temsillerinden üfürüm tespitine kadar uzanan süreci kapsayacak şekilde çok aşamalı ve ardışık bir yapıdan oluşmaktadır.

İlk aşamada CNN tabanlı özellik çıkarımı gerçekleştirilmektedir. Bu kapsamda Mel-spektrogram girdileri üç aşamalı bir evrişimsel yapıdan geçirilerek yüksek seviyeli özellik haritalarına dönüřtürölmektedir. İlk iki evrişimsel blokta sırasıyla 32 ve 64 filtreli Conv2D katmanları, üçüncü blokta ise 128 filtreli Conv2D katmanları kullanılmaktadır. Her bir blokta Batch Normalization ve MaxPooling işlemleri uygulanarak öğrenme sürecinin kararlılığı artırılmakta ve hesaplama verimliliğı iyileştirilmektedir.

Evrişimsel özellik çıkarımının ardından elde edilen üç boyutlu özellik haritaları, zaman eksenini boyunca sıralı bir temsil oluşturmak amacıyla yeniden şekillendirme (Reshape) işlemi ile iki boyutlu bir zaman dizisine dönüřtürölmektedir. Bu adım, her bir zaman adımında yerel spektral–zamansal örüntülerin korunmasını sağlayarak LSTM katmanları için uygun bir giriş yapısı oluşturmaktadır.

Bir sonraki aşamada LSTM tabanlı zamansal modelleme gerçekleştirilmektedir. Yeniden şekillendirilmiş diziler sırasıyla 64 birimli ve `return_sequences=True` ayarına sahip bir LSTM katmanı ile 32 birimli ikinci bir LSTM katmanı aracılığıyla işlenmektedir. Bu katmanlar, üfürümlerin S1 ve S2 kalp seslerine göre zamanlama, süre ve yoğunluk bakımından gösterdiği varyasyonları öğrenmeyi amaçlamaktadır. Dropout katmanları ile aşırı öğrenme (overfitting) kontrol altına alınmaktadır.

Son aşamada yoğun (Dense) sınıflandırma katmanları kullanılmaktadır. Zamansal olarak özetlenmiş temsil, doğrusal olmayan dönüşümler aracılığıyla işlenmekte ve sigmoid aktivasyonlu tek bir çıkış birimi ile ikili üfürüm tespiti gerçekleştirilmektedir.

Modelin eğitimi sırasında kullanılan kayıp fonksiyonu, ikili sınıflandırma problemleri için yaygın olarak tercih edilen Binary Cross-Entropy olarak tanımlanmaktadır:

Eşitlik 43. Binary Cross-Entropy Kayıp Fonksiyonu

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (\text{Eşitlik 43})$$

Tanımlanan bu hibrit CNN–LSTM mimarisi, evrişimli katmanların yerel özellik çıkarım gücü ile LSTM ağlarının uzun vadeli zamansal bağımlılıkları modelleme yeteneğini bir araya getirerek literatürde kalp sesi analizi için etkili olduğu gösterilen CNN–RNN tabanlı yaklaşımlarla yöntemsel olarak uyumlu ve güçlü bir çözüm sunmaktadır.

3.5.4. Mel ve Geleneksel Özellikler için Çoklu Girdili Model

Görüntü benzeri Mel-spektrogram temsilleri ile skaler geleneksel ses özelliklerinden elde edilen bilgilerin açık ve tamamlayıcı biçimde birleştirilebilmesi amacıyla `create_multi_input_model` fonksiyonu içerisinde çok girdili bir sinir ağı mimarisi tanımlanmıştır. Bu mimari, birbirini tamamlayan iki paralel işlem dalından oluşmaktadır.

İlk dal olan Mel-spektrogram dalı, Mel-spektrogram girdilerini ardışık evrişimsel bloklar ve Global Average Pooling katmanı aracılığıyla işleyerek yüksek seviyeli görsel–spektral özellik temsilleri üretmektedir. Bu dal, CNN tabanlı bir özellik çıkarıcı olarak çalışmakta ve zaman–frekans düzlemindeki karmaşık örüntüleri otomatik olarak öğrenmektedir.

İkinci dal olan geleneksel özellik dalı ise önceki bölümlerde seçilen özet ses özelliklerini işlemek üzere yapılandırılmış çok katmanlı yoğun (Dense) katmanlardan oluşmaktadır. Bu dal, enerji, periyodisite ve spektral şekil gibi klinik açıdan yorumlanabilir özellikleri modele kazandırmaktadır.

Her iki dalda elde edilen özet temsiller, birleştirme (concatenation) işlemi ile tek bir birleşik özellik vektörü hâline getirilmektedir:

Eşitlik 44. Özellik Birleştirme (Feature Concatenation)

$$z = [z_{\text{CNN}} \parallel z_{\text{ML}}] \quad (\text{Eşitlik 44})$$

Burada z_{CNN} CNN dalından elde edilen özellik vektörünü, z_{ML} ise geleneksel özellik dalından elde edilen vektörü temsil etmektedir.

Bu birleşik temsil daha sonra ardışık Dense ve Dropout katmanlarından geçirilerek doğrusal olmayan dönüşümler aracılığıyla son sınıflandırma işlemi gerçekleştirilmektedir.

Tanımlanan bu çok girdili yapı, hibrit özellik kullanımının modele sağladığı katkının saf tek girdili modeller ile nicel olarak karşılaştırılabilmesine olanak tanımak amacıyla deneysel değerlendirmelerde kullanılmıştır.

3.5.5. Düzenleştirme, Aktivasyon Fonksiyonları ve Optimizasyon

Tüm derin öğrenme modellerinde tutarlılığı sağlamak amacıyla ortak tasarım ilkeleri benimsenmiştir. Gizli katmanlarda ReLU aktivasyon fonksiyonu kullanılarak hem hesaplama verimliliği artırılmış hem de kaybolan gradyan problemi azaltılmıştır. İkili sınıflandırma problemi için çıkış katmanında sigmoid aktivasyon tercih edilmiştir.

Aşırı uyumu (overfitting) azaltmak amacıyla evrimsel ve yoğun katmanların ardından %30 oranında Dropout uygulanmış; ayrıca Batch Normalization ile katmanlar arası aktivasyon dağılımları dengelenmiştir.

Tüm modeller, ikili çapraz entropi kaybı (binary cross-entropy) ve Adam optimize edicisi kullanılarak eğitilmiştir. Adam optimizasyon algoritması, gradyanların birinci ve ikinci moment tahminlerini kullanarak parametre güncellemesi gerçekleştirmektedir:

Eşitlik 45. Adam Optimizasyon Algoritması

$$\begin{aligned} m_t &= \beta_1 m_{t-1} + (1 - \beta_1) g_t \\ v_t &= \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \\ \hat{m}_t &= \frac{m_t}{1 - \beta_1^t}, \hat{v}_t = \frac{v_t}{1 - \beta_2^t} \\ \theta_t &= \theta_{t-1} - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \end{aligned} \quad (\text{Eşitlik 45})$$

Burada g_t gradyanı, η öğrenme oranını (0,001), β_1 ve β_2 moment katsayılarını temsil etmektedir.

Eğitim sürecinde erken durdurma (early stopping) ve doğrulama kaybı plato yaptığında öğrenme oranının azaltılması gibi geri çağırma (callback) mekanizmaları kullanılarak daha kararlı ve genellenebilir modeller elde edilmiştir.

3.6. Eğitim Süreci

Bu bölümde, önerilen derin öğrenme modellerinin eğitimi sırasında izlenen veri bölme stratejileri, hiperparametre ayarlama süreci ile kullanılan kayıp fonksiyonları ve optimizasyon algoritmaları ayrıntılı biçimde açıklanmaktadır. Eğitim sürecinin dikkatli ve sistematik biçimde tasarlanması, özellikle sınıf dengesizliği, hasta içi korelasyonlar ve sınırlı veri hacmi gibi pediatrik FKG veri kümelerine özgü zorluklar dikkate alındığında, güvenilir ve genellenebilir performans elde edilmesi açısından kritik öneme sahiptir.

Bu çerçevede, model eğitiminin yalnızca yüksek doğruluk elde etmeye yönelik değil, aynı zamanda daha önce görülmemiş hasta verileri üzerinde kararlı ve klinik açıdan anlamlı sonuçlar üretebilecek şekilde yapılandırılması hedeflenmiştir.

3.6.1. Veri Bölme Stratejisi

Model performansının gerçekçi biçimde değerlendirilmesini sağlamak ve iyimser yanlılığı (optimistic bias) önlemek amacıyla veri kümesinin eğitim, doğrulama ve test alt kümelerine ayrılmasında katı ve klinik açıdan anlamlı bölme ilkeleri benimsenmiştir.

Bu çalışmada temel ilke olarak hasta düzeyinde tabakalama (patient-level stratification) uygulanmıştır. Belirli bir hastaya ait tüm FKG segmentleri — veri artırma süreciyle elde edilen ek segmentler dâhil olmak üzere — yalnızca tek bir alt kümeye (eğitim, doğrulama veya test) atanmıştır. Bu işlem, `split_augmented_samples` fonksiyonu içerisinde hasta kimliği (patient ID) temel alınarak gerçekleştirilen gruplama yöntemiyle sağlanmıştır.

Bu tasarım tercihi, modelin eğitim aşamasında bir hastaya ait bazı segmentleri görüp değerlendirme aşamasında aynı hastanın farklı segmentleri üzerinde test edilmesi durumunda ortaya çıkabilecek veri sızıntısı (data leakage) sorununu etkin biçimde ortadan kaldırmaktadır. Böylece elde edilen performans ölçütleri, modelin daha önce hiç görülmemiş hastalar üzerindeki genelleme yeteneğini daha güvenilir biçimde yansıtmaktadır.

Hasta düzeyinde gruplanmış veri kümesi, önceden belirlenen oransal hedeflere göre eğitim, doğrulama ve test kümelerine ayrılmıştır. Hedef dağılım oranı sırasıyla %60 eğitim, %30 doğrulama ve %10 test olacak şekilde belirlenmiştir. Veri artırma sürecinin hasta başına düşen segment sayısını artırması nedeniyle bölme işlemi segment sayıları yerine hasta sayıları üzerinden gerçekleştirilmiş ve böylece örnek bağımlılığı korunmuştur.

Bölme işlemi sonucunda kalan hastalar, etkin örnek sayısını ve veri çeşitliliğini artırmak amacıyla eğitim kümesine dâhil edilmiştir. Bu yaklaşım, modelin öğrenme kapasitesini maksimize ederken değerlendirme sürecinin tarafsızlığını korumaktadır.

Belirtilen kısıtlar altında üfürüm var ve üfürüm yok sınıflarının özgün dağılımının mümkün olduğunca korunmasına özen gösterilmiştir. Birincil derin öğrenme deneylerinde kullanılan nihai yapılandırma 510 eğitim segmenti, 124 doğrulama segmenti ve 124 test segmenti içermekte olup bu dağılım sınıf oranlarını yaklaşık olarak muhafaza etmektedir.

Buna ek olarak bazı keşifsel deneylerde (örneğin yalnızca aort kapak Mel-spektrogramlarının kullanıldığı CNN tabanlı çalışmalar) hasta düzeyinde filtreleme ilkesi korunmakla birlikte etiketlere göre tabakalama yapılarak `train_test_split` fonksiyonu aracılığıyla %60 eğitim, %20 doğrulama ve %20 test olacak şekilde alternatif bir bölme stratejisi uygulanmıştır. Bu esnek yaklaşım, farklı mimarilerin davranışlarının karşılaştırmalı olarak analiz edilmesine olanak tanımıştır.

3.6.2. Hiperparametre Ayarlaması

Derin öğrenme modellerinin performansı, mimari tasarımın yanı sıra hiperparametre seçimlerine de yüksek düzeyde duyarlıdır. Bu çalışmada ele alınan başlıca hiperparametreler; öğrenme oranı ve optimize edici ayarları, evrimsel katmanlardaki filtre sayıları, yoğun (Dense) katmanlardaki nöron sayıları, Dropout oranları, LSTM katmanlarının sayısı ve gizli durum boyutları ile batch boyutu ve maksimum epoch sayısını kapsamaktadır.

Hibrit CNN–LSTM gibi derin ve hesaplama açısından maliyetli modeller için kapsamlı bir grid search yaklaşımının pratik olmaması nedeniyle bu çalışmada deneyime dayalı, aşamalı ve manuel bir hiperparametre ayarlama stratejisi benimsenmiştir. Bu yaklaşım, sınırlı veri ve yüksek hesaplama maliyetine sahip senaryolarda yaygın olarak tercih edilen etkin bir yöntemdir.

Hiperparametre ayarlama süreci sistematik bir biçimde yürütülmüştür. İlk aşamada başlangıç hiperparametreleri, kalp sesi sınıflandırması ve spektrogram tabanlı derin öğrenme modellerine ilişkin literatürde raporlanan çalışmalara dayanarak belirlenmiştir. Eğitim sürecinde doğrulama kaybı ve F1 skoru eş zamanlı olarak izlenmiş; ayrıca eğitim ve doğrulama öğrenme eğrileri analiz edilerek yetersiz öğrenme (underfitting) ve aşırı öğrenme (overfitting) davranışları değerlendirilmiştir.

Elde edilen gözlemler doğrultusunda model derinliği, Dropout oranları ve öğrenme oranı kademeli olarak ayarlanmıştır. Bu iteratif yaklaşım, modelin genelleme kabiliyetini artırırken aşırı öğrenme riskini azaltmayı hedeflemiştir.

Ayrıca en iyi doğrulama performansının elde edildiği epoch'un otomatik olarak belirlenmesi ve gereksiz eğitim döngülerinin önlenmesi amacıyla erken durdurma (early stopping) mekanizması kullanılmıştır.

Klasik makine öğrenmesi tabanlı temel modeller (örneğin SVM ve AdaBoost) için ise hiperparametreler, eğitim kümesi üzerinde çapraz doğrulama (cross-validation) yoluyla F1 skorunun maksimize edilmesi esas alınarak belirlenmiştir. Bu yaklaşım, derin

öğrenme ve klasik model aileleri arasında adil ve tutarlı bir performans karşılaştırması yapılmasını sağlamaktadır.

3.6.3. Kayıp Fonksiyonları ve Optimizasyon Algoritmaları

İkili üfürüm tespiti problemi için bu çalışmada kullanılan tüm derin öğrenme modelleri, ikili çapraz entropi (binary cross-entropy) kayıp fonksiyonunu minimize edecek şekilde eğitilmiştir. Bu kayıp fonksiyonu, olasılıksal ikili sınıflandırma problemleri için standart kabul edilmekte olup model çıktılarının gerçek etiketlerle olan uyumsuzluğunu log-olasılık temelli olarak cezalandırmaktadır:

Eşitlik 46. Binary Cross-Entropy Kayıp Fonksiyonu

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (\text{Eşitlik 46})$$

Burada y_i gerçek etiketi (üfürüm yok / üfürüm var), $\hat{y}_i$ ise model tarafından tahmin edilen olasılığı ifade etmektedir.

Optimizasyon sürecinde adaptif öğrenme oranı mekanizmasına sahip olması ve pratikte hızlı yakınsama sağlaması nedeniyle Adam optimize edicisi tercih edilmiştir (Kingma ve Ba, 2015). Adam algoritması, gradyanların birinci ve ikinci moment tahminlerine dayalı olarak parametre güncellemelerini aşağıdaki şekilde gerçekleştirmektedir:

Eşitlik 47. Adam Optimizasyon Güncellemesi

$$\theta_t = \theta_{t-1} - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} \quad (\text{Eşitlik 47})$$

Burada η öğrenme oranını (0,001), $\hat{m}_t$ ve $\hat{v}_t$ ise sırasıyla bias düzeltmeli birinci ve ikinci moment tahminlerini temsil etmektedir.

Başlangıç öğrenme oranı 0,001 olarak belirlenmiş; doğrulama kaybının plato yaptığı durumlarda öğrenme oranının azaltılması için uygun geri çağırma (callback) mekanizmaları kullanılmıştır. Bu yaklaşım, eğitim sürecinin daha kararlı ilerlemesini sağlamaktadır.

Bu kombinasyon, özellikle sınırlı veri koşullarında modelin genellenebilirliğini artırmakta ve aşırı öğrenme riskini azaltmaktadır.

Klasik temel modeller için ise scikit-learn uygulamaları, kendi kayıp formülasyonlarıyla birlikte kullanılmıştır. Örneğin doğrusal çekirdekli SVM

modellerinde hinge loss, karar ağaçlarında ise entropi temelli bölme ölçütleri tercih edilmiştir.

3.7. Değerlendirme Ölçütleri

Üfürüm tespiti probleminin klinik bağlamı göz önüne alındığında, model değerlendirmesinin yalnızca genel doğruluk oranına indirgenmesi yeterli değildir. Klinik uygulamalarda patolojik üfürümlerin atlanmaması ile gereksiz yanlış alarmların önlenmesi arasındaki denge hayati önem taşımaktadır. Bu nedenle önerilen modellerin performansı, ayrılmış test kümesi üzerinde birden fazla tamamlayıcı ölçüt kullanılarak değerlendirilmiştir.

Bu çalışmada geliştirilen modellerin performansını kapsamlı ve klinik açıdan anlamlı biçimde değerlendirebilmek amacıyla, ikili sınıflandırma problemleri için yaygın olarak kabul gören bir dizi değerlendirme ölçütü kullanılmıştır (Sokolova ve Lapalme, 2009). Hesaplanan ölçütler aşağıda özetlenmektedir:

Doğruluk (Accuracy – ACC), tüm test örnekleri içerisinde doğru sınıflandırılan örneklerin oranını ifade etmektedir.

Duyarlılık / Geri çağırma (Sensitivity / Recall – SE), gerçek üfürüm-pozitif vakaların ne kadarının doğru biçimde tespit edildiğini ölçmektedir ve klinik açıdan özellikle kaçırılan vakaların (false negatives) azaltılması bakımından kritik öneme sahiptir.

Özgüllük (Specificity – SP), üfürüm bulunmayan vakaların doğru biçimde normal olarak sınıflandırılma oranını göstermekte olup aşırı tanının (false positives) önlenmesi açısından önem taşımaktadır.

Kesinlik (Precision / Positive Predictive Value – PPV), üfürüm olarak tahmin edilen örneklerin ne kadarının gerçekten üfürüm içerdiğini ifade etmektedir.

F1-skoru, kesinlik ile geri çağırmanın harmonik ortalaması olup özellikle sınıf dengesizliğinin bulunduğu senaryolarda model performansını daha dengeli biçimde yansıtmaktadır.

ROC eğrisi altında kalan alan (Area Under the Curve – AUC) ise rastgele seçilen bir üfürüm-pozitif örneğin, rastgele seçilen bir üfürüm-negatif örnekten daha yüksek bir tahmin skoruna sahip olma olasılığını temsil etmekte ve eşik değerinden bağımsız bir ayırım gücü ölçütü sunmaktadır.

Söz konusu ölçütler, ikili tıbbi sınıflandırma problemlerinde yaygın biçimde kullanılmakta olup kardiyoloji ve tıbbi görüntüleme alanlarında yapay zekâ tabanlı

sistemlerin değerlendirilmesine yönelik güncel metodolojik yaklaşımlarla uyumludur (Sokolova ve Lapalme, 2009).

Buna ek olarak model tahminlerinin ayrıntılı biçimde analiz edilebilmesi amacıyla karmaşıklık matrisi (confusion matrix) incelenmiştir. Bu matris, TP, FP, TN ve FN örneklerinin dağılımını görselleştirerek klinik açıdan kritik hata örüntülerine — örneğin atlanmış patolojik üfürümler veya aşırı tanı eğilimleri — ilişkin önemli içgörüler sunmaktadır.

Birden fazla derin öğrenme mimarisi ve temel modelin karşılaştırıldığı deneylerde ise, sınıf dengesizliği koşullarında doğruluk ölçütüne kıyasla daha bilgilendirici olmaları nedeniyle F1-skoru ve AUC, performans karşılaştırmalarında birincil değerlendirme ölçütleri olarak benimsenmiştir.

3.8. Deneysel Kurulum

3.8.1. Donanım Ortamı

Tüm deneyler, çok çekirdekli bir Merkezi İşlem Birimi (Central Processing Unit – CPU), en az 16 GB RAM ve derin öğrenme modellerinin verimli biçimde eğitilebilmesi için CUDA desteğine sahip bir Grafik İşlem Birimi (Graphics Processing Unit – GPU) içeren bir iş istasyonu üzerinde gerçekleştirilmiştir.

Kullanılan donanım yapılandırması, evrişimli sinir ağları (CNN) ve uzun-kısa süreli bellek ağları (LSTM) gibi hesaplama açısından yoğun modellerin paralel olarak eğitilmesini mümkün kılmıştır.

Eğitim süreçleri arasında donanım özelliklerinde küçük farklılıklar bulunmakla birlikte, kullanılan Mel-spektrogram girdilerinin nispeten düşük boyutlu (128×216) olması ve veri kümesinin orta ölçekli yapısı nedeniyle hesaplama gereksinimleri, büyük ölçekli görüntü veya doğal dil işleme modellerine kıyasla sınırlı kalmıştır.

Bu durum, önerilen yaklaşımın yalnızca yüksek performanslı sunucu sistemlerinde değil, aynı zamanda daha erişilebilir donanım altyapılarında da uygulanabilir olduğunu göstermektedir.

3.8.2. Yazılım Ortamı

Uygulama, tamamen açık kaynaklı ve yaygın olarak kullanılan bir yazılım yığını üzerine inşa edilmiştir. Bu kapsamda Python 3.x programlama dili temel geliştirme ortamı olarak kullanılmıştır.

Sinyal işleme aşamalarında; ses yükleme, kısa zamanlı Fourier dönüşümü (STFT), Mel-spektrogram ve MFCC çıkarımı ile onset tespiti gibi işlemler için librosa kütüphanesinden yararlanılmıştır. Sayısal hesaplama ve veri işleme süreçleri ise NumPy ve pandas kütüphaneleri kullanılarak gerçekleştirilmiştir.

Klasik makine öğrenmesi modellerinin uygulanmasında; Lojistik Regresyon, Destek Vektör Makineleri (SVM), K-En Yakın Komşu (KNN), Karar Ağaçları, Doğrusal Ayırıcı Analiz (LDA), İkincil Ayırıcı Analiz (QDA) ve AdaBoost algoritmalarının yanı sıra özellik seçimi teknikleri için scikit-learn kütüphanesinden yararlanılmıştır.

Derin öğrenme tabanlı CNN, LSTM ve hibrit CNN–LSTM mimarileri ise TensorFlow ve onun yüksek seviyeli arayüzü olan Keras kullanılarak geliştirilmiştir.

Deney sonuçlarının ve ara çıktıların görselleştirilmesinde matplotlib ve seaborn kütüphaneleri kullanılmıştır.

Tüm betikler, yeniden üretilebilirliği destekleyecek biçimde modüler Python dosyaları hâlinde yapılandırılmıştır. Bu kapsamda veri yükleme, özellik çıkarımı ve modelleme iş akışları birbirinden ayrıştırılmış olup; örneğin `data_loader_ML.py`, `feature_extraction_ML.py` ve `mel_spectrogram_deep_learning.py` gibi bileşenler bağımsız ve yeniden kullanılabilir modüller olarak tasarlanmıştır.

3.9. Karşılaştırma için Temel (Baseline) Modeller

Önerilen hibrit CNN–LSTM çerçevesinin performansını bağlamlandırmak ve sağladığı kazanımları nicel olarak ortaya koymak amacıyla, literatürde yaygın olarak kullanılan çeşitli temel (*baseline*) modeller karşılaştırma kapsamına alınmıştır. Bu modeller, hem geleneksel makine öğrenmesi yaklaşımlarını hem de derin öğrenme tabanlı mimarileri temsil edecek şekilde sistematik olarak seçilmiştir.

Karşılaştırma kapsamında, kalp sesi sınıflandırması literatüründe yaygın olarak kullanılan Destek Vektör Makineleri (SVM), Karar Ağaçları, AdaBoost ve Rastgele Orman gibi klasik makine öğrenmesi algoritmalarının yanı sıra, CNN ve LSTM tabanlı derin öğrenme mimarileri değerlendirilmiştir. Bu modeller, fonokardiyogram (PCG) sinyallerinden elde edilen akustik özellikler kullanılarak kalp seslerinin normal veya patolojik olarak sınıflandırılmasında geniş ölçüde kullanılmaktadır.

Geleneksel makine öğrenmesi yöntemleri, el ile tasarlanmış özellikler üzerinde çalışarak sınıflandırma görevini bu özelliklerin istatistiksel ayırım gücüne dayalı olarak gerçekleştirmektedir. Buna karşılık derin öğrenme tabanlı yaklaşımlar, özellikle spektrogram temsilleri üzerinden doğrudan özellik öğrenimi gerçekleştirerek daha karmaşık ve çok boyutlu akustik örüntüleri modelleyebilmektedir. CNN mimarileri yerel

spektral örüntülerin çıkarılmasında etkili olurken, LSTM yapıları zamansal bağımlılıkların öğrenilmesini sağlamaktadır.

Bu bağlamda gerçekleştirilen karşılaştırmalar, önerilen hibrit CNN–LSTM mimarisinin performansını yalnızca mutlak doğruluk değerleri üzerinden değil, aynı zamanda farklı model ailelerine karşı sağladığı göreceli avantajlar açısından da değerlendirmeyi amaçlamaktadır. Böylece elde edilen bulgular, önerilen yaklaşımın literatürdeki mevcut yöntemlere kıyasla sunduğu katkıları daha açık, sistematik ve karşılaştırılabilir bir çerçevede ortaya koymaktadır.

3.9.1. Geleneksel Makine Öğrenimi Sınıflandırıcıları

Seçilen el ile tasarlanmış özellik alt kümesi kullanılarak bir dizi klasik makine öğrenmesi sınıflandırıcısı eğitilmiştir. Bu kapsamda, elastik ağ (*elastic-net*) düzenleştirmesi içeren Lojistik Regresyon, sigmoid çekirdekli Destek Vektör Makinesi (SVM), K-En Yakın Komşu (KNN) algoritması, entropi ölçütü ve derinlik budaması kullanılan Karar Ağacı, Doğrusal Ayırıcı Analiz (LDA) ve Kuadratik Ayırıcı Analiz (QDA) modelleri değerlendirmeye alınmıştır.

Buna ek olarak, temel öğrenici olarak karar kütükleri (*decision stumps*) kullanan AdaBoost yöntemi de uygulanmıştır. Bu model, zayıf öğrenicilerin birleşimiyle daha güçlü ve genellenebilir bir sınıflandırıcı elde etmeyi amaçlamaktadır.

Her bir modelin hiperparametreleri, eğitim kümesi üzerinde çapraz doğrulama (*cross-validation*) kullanılarak F1-skorunun maksimize edilmesi esasına göre ayarlanmıştır. Bu süreç, modeller arasında adil ve tutarlı bir karşılaştırma yapılmasını sağlamaktadır.

Söz konusu klasik modeller, derin öğrenme tabanlı yaklaşımların sağladığı katma değerlerin nicel olarak değerlendirilmesi amacıyla deneysel çalışmalarda kıyas ölçütü (*benchmark*) olarak kullanılmıştır. Bu sayede önerilen hibrit CNN–LSTM mimarisinin performansı, daha basit ve yorumlanabilir model aileleriyle doğrudan karşılaştırılabilmiştir.

3.9.2. Temel Derin Öğrenme Modelleri

Hibrit mimarinin sağladığı katkıları izole edebilmek amacıyla, bu çalışmada daha basit iki derin öğrenme modeli de karşılaştırmalı olarak değerlendirilmiştir. Bu yaklaşım, önerilen CNN–LSTM mimarisinin performans kazanımlarının hangi bileşenlerden kaynaklandığını daha açık biçimde analiz etmeyi amaçlamaktadır.

İlk olarak bağımsız CNN modeli ele alınmıştır. Bu model, Bölüm 3.5.1’de tanımlanan iki boyutlu (2-B) CNN mimarisine dayanmaktadır ve yalnızca tek bir oskültasyon bölgesine (örneğin aort kapağı – AV) ait Mel-spektrogram girdileri kullanılarak eğitilmiştir. Bu yapı, uzamsal–spektral örüntülerin zamansal modelleme olmaksızın ne ölçüde ayırt edici olduğunu incelemek amacıyla tasarlanmıştır.

İkinci olarak bağımsız LSTM modeli değerlendirilmiştir. Bölüm 3.5.2’de sunulan bu model, Mel-spektrogramlardan veya el ile tasarlanmış özelliklerden türetilen zaman dizisi temsilleri üzerinde çalışmakta ve kardiyak döngüler boyunca ortaya çıkan zamansal bağımlılıkları modellemeyi hedeflemektedir. Bu yaklaşım, spektral yapıyı doğrudan kullanmaksızın yalnızca zamansal ilişkilerin sınıflandırma performansına katkısını incelemeye olanak tanımaktadır.

Gerçekleştirilen bu karşılaştırmalı değerlendirme, uzamsal (CNN) ve zamansal (LSTM) modelleme yaklaşımlarının ayrı ayrı katkılarını ortaya koymanın yanı sıra, bu iki yaklaşımın hibrit CNN–LSTM mimarisi içerisinde birleştirilmesinin sağladığı performans artışını nicel olarak analiz etmeye imkân tanımaktadır.

3.9.3. Gelişmiş Transfer Öğrenme Mimarileri

Önerilen hibrit modelin etkinliğini daha ileri düzeyde doğrulamak amacıyla, bilgisayarlı görü alanında yaygın olarak kullanılan ve ImageNet veri kümesi üzerinde önceden eğitilmiş (pre-trained) çeşitli CNN mimarileri, Mel-spektrogram temsilleri üzerinde ince ayar (fine-tuning) yapılarak değerlendirilmiştir.

Bu kapsamda DenseNet121, ResNet-50 ve ResNet-101, EfficientNet-B0, EfficientNet-B1 ve EfficientNet-B2, MobileNet-V2 ve MobileNet-V3 Small, GhostNet ile dikkat (attention) mekanizmalarıyla güçlendirilmiş CNN mimarileri karşılaştırma sürecine dâhil edilmiştir. Bu mimariler, farklı derinlik ve parametre karmaşıklığı seviyelerini temsil edecek şekilde seçilmiştir.

Söz konusu ağlar, son sınıflandırma katmanları ikili üfürüm tespiti problemine uyarlanacak şekilde yeniden yapılandırılmıştır. Ayrıca seçili evrimsel bloklar ince ayar sürecine dâhil edilerek, Mel-spektrogram temsillerine özgü akustik örüntülerin daha etkin biçimde öğrenilmesi sağlanmıştır.

Konferans makalesinde raporlanan deneysel sonuçlar, bazı transfer öğrenme tabanlı mimarilerin rekabetçi F1-skorları elde edebildiğini göstermektedir. Bununla birlikte, özel olarak tasarlanan hibrit CNN–LSTM modelinin, doğruluk, F1-skoru ve AUC ölçütleri açısından daha dengeli ve tutarlı bir performans sunduğu gözlemlenmiştir.

Bu bulgu, yalnızca güçlü spektral özellik çıkarımının yeterli olmadığını; aynı zamanda kalp döngülerine özgü zamansal bağımlılıkların modellenmesinin, pediatrik üfürüm tespiti gibi karmaşık biyomedikal sinyal analiz görevlerinde kritik bir rol oynadığını ortaya koymaktadır.

3.10. İstatistiksel Analiz

Görece sınırlı veri kümesi boyutu ve üfürüm tespitinin potansiyel klinik sonuçları dikkate alındığında, model performansının değerlendirilmesinde istatistiksel analiz kritik bir öneme sahiptir. Bu doğrultuda çalışmada çok yönlü ve tamamlayıcı bir analiz çerçevesi benimsenmiştir.

İlk olarak, nokta tahminleri ve güven aralıkları hesaplanmıştır. Her bir model ve değerlendirme ölçütü (doğruluk, duyarlılık, özgüllük, F1-skoru ve AUC) için ayrılmış test kümesi üzerinde nokta tahminleri elde edilmiştir. Tahminlerin belirsizliğini nicel olarak ifade edebilmek amacıyla bootstrap yeniden örnekleme yöntemi kullanılarak %95 güven aralıkları hesaplanmıştır. Bu yaklaşım, sınırlı örneklem büyüklüğüne sahip veri kümelerinde performans tahminlerinin daha güvenilir biçimde yorumlanmasına olanak tanımaktadır.

İkinci olarak, karşılaştırmalı anlamlılık testleri uygulanmıştır. Aynı test kümesi üzerinde iki farklı sınıflandırıcının karşılaştırıldığı durumlarda (örneğin hibrit CNN–LSTM modeli ile DenseNet121), eşleştirilmiş tahminlere dayalı olarak McNemar testi kullanılmıştır. F1-skoru ve AUC gibi türetilmiş performans ölçütleri için ise bootstrap örnekleri üzerinden eşleştirilmiş t-testi uygulanarak gözlenen performans farklarının istatistiksel anlamlılığı değerlendirilmiştir.

Üçüncü olarak, etki büyüklüğü (effect size) raporlamasına yer verilmiştir. Yalnızca p-değerlerine odaklanmak yerine, hibrit mimarinin sağladığı performans iyileşmesinin büyüklüğünü daha açık biçimde ifade edebilmek amacıyla F1-skoru ve AUC değerlerindeki mutlak farklar gibi etki büyüklüğü ölçütleri sunulmuştur. Bu yaklaşım, elde edilen sonuçların klinik açıdan anlamlılığının daha doğru biçimde yorumlanmasına katkı sağlamaktadır.

Bu istatistiksel analiz çerçevesi, tıbbi yapay zekâ alanında yalnızca ortalama performans değerlerine değil, aynı zamanda belirsizlik, karşılaştırmalı anlamlılık ve klinik etki büyüklüğüne odaklanan güncel değerlendirme yaklaşımlarıyla uyumludur.

3.11. Doğrulama Yöntemi

3.11.1. Dahili Doğrulama

Bu çalışmada benimsenen temel doğrulama stratejisi, Bölüm 3.6.1’de ayrıntılı olarak tanımlanan hasta düzeyinde tabakalı eğitim/doğrulama/test bölünmesine dayanan dahili doğrulama yaklaşımıdır. Bu yapılandırma, model performansının tarafsız, güvenilir ve klinik açıdan anlamlı biçimde değerlendirilmesini amaçlamaktadır.

Bu çerçevede eğitim kümesi, model parametrelerinin öğrenilmesi için kullanılmıştır. Doğrulama kümesi, hiperparametre ayarlaması, erken durdurma (early stopping) mekanizmalarının işletilmesi ve model seçimi süreçlerini yönlendirmek amacıyla ayrılmıştır. Test kümesi ise nihai performansın raporlanması için tamamen ayrılmış ve eğitim ya da doğrulama aşamalarında hiçbir şekilde kullanılmamış, daha önce görülmemiş bir veri kümesi olarak muhafaza edilmiştir.

Bu üçlü ayırım, aşırı uyum (overfitting) riskini azaltırken elde edilen performans ölçütlerinin gerçekçi ve tarafsız biçimde raporlanmasına olanak tanımaktadır. Ayrıca hasta düzeyinde bölme uygulanması sayesinde aynı bireye ait farklı FKG segmentlerinin eğitim ve test kümelerine dağılması engellenmiş; böylece literatürde sıklıkla vurgulanan bilgi sızıntısı (data leakage) problemi etkin biçimde önlenmiştir.

Bazı keşifsel senaryolarda — özellikle klasik makine öğrenmesi modellerinin karşılaştırılması sırasında — performans tahminlerindeki varyansı azaltmak amacıyla eğitim kümesi üzerinde k-katlı çapraz doğrulama (k-fold cross-validation) uygulanmıştır. Buna karşılık derin öğrenme deneylerinde, yüksek hesaplama maliyeti ve uzun eğitim süreleri göz önüne alınarak sabit bir doğrulama kümesi kullanılması tercih edilmiştir. Bu yaklaşım, kalp sesi analizi alanındaki çağdaş derin öğrenme uygulamalarıyla yöntemsel olarak uyumludur.

3.11.2. Harici Doğrulama

Bu tez kapsamında gerçekleştirilen tüm deneyler, CirCor DigiScope veri kümesi üzerinde yürütülmüştür. Söz konusu veri kümesi; örnek sayısı, anotasyon kalitesi ve çoklu oskültasyon bölgelerini içermesi bakımından kapsamlı ve çeşitlidir. Bununla birlikte veri kümesi, belirli bir coğrafî, klinik ve cihaz bağlamına özgü özellikler taşımaktadır.

Bu nedenle, gerçek anlamda bir harici doğrulama (external validation) — örneğin eğitilmiş CNN–LSTM modelinin PhysioNet/CinC 2016 veri kümesi gibi farklı bir pediatrik fonokardiyogram (FKG) veri kümesi veya bağımsız klinik kayıtlar üzerinde

değerlendirilmesi — bu çalışmanın kapsamı dışında bırakılmıştır. Bununla birlikte bu tür bir değerlendirme, önerilen yaklaşımın farklı veri dağılımlarına karşı genellenebilirliğinin daha kapsamlı biçimde incelenebilmesi açısından gelecek çalışmalar için önemli bir araştırma yönü olarak değerlendirilmektedir.

Buna rağmen, bu tezde tanımlanan metodolojik çerçeve veri kümesinden bağımsız olacak şekilde tasarlanmıştır. Ham FKG kayıtları ve ilgili etiketler mevcut olduğu sürece; ön işleme, özellik çıkarımı ve modelleme adımlarının aynı iş akışı içerisinde sınırlı uyarlamalarla farklı veri kümelerine uygulanabilmesi mümkündür.

Bu durum, önerilen yaklaşımın yalnızca belirli bir veri kümesine özgü olmadığını, aynı zamanda farklı klinik ve veri edinim koşullarına uyarlanabilir potansiyel bir çözüm sunduğunu göstermektedir.

3.12. Bölüm Özeti

Bu bölümde, Mel-spektrogram temsilleri ile geleneksel el ile tasarlanmış ses özelliklerinin birlikte kullanıldığı ve pediatrik kalp üfürümlerinin otomatik tespiti amacıyla geliştirilen hibrit CNN–LSTM çerçevesine ilişkin kapsamlı metodoloji ayrıntılı biçimde sunulmuştur.

CirCor DigiScope veri kümesi temel alınarak yapılandırılan iş akışı; dahil etme ve hariç tutma kriterlerinin sistematik biçimde uygulanması yoluyla dört standart oskültasyon bölgesini kapsayan bir pediatrik hasta kohortunun oluşturulmasını, uzunluk normalizasyonu ve hedefe yönelik segment tabanlı veri artırma yöntemleri ile 5 saniyelik standartlaştırılmış çok kapaklı FKG segmentlerinin elde edilmesini ve böylece sınıf dengesizliğinin azaltılmasını içermektedir.

Bunu takiben, iki boyutlu (2-B) Mel-spektrogram temsilleri ile zaman, frekans, kepsral ve istatistiksel alan tanımlayıcılarını kapsayan zengin bir hibrit özellik kümesi çıkarılmıştır. Uzamsal ve zamansal modellemeyi birleştiren özel bir CNN–LSTM hibrit mimarisine odaklanılarak, geleneksel özellikleri de içeren çok girdili ağlar dâhil olmak üzere çeşitli derin öğrenme modelleri tasarlanmıştır.

Modellerin eğitimi; hasta düzeyinde tabakalı veri bölme, Adam optimizasyonu, erken durdurma (early stopping) ve doğrulama yönlendirmeli hiperparametre ayarlaması kullanılarak yürütülmüştür. Performans değerlendirmesi; doğruluk, duyarlılık, özgüllük, F1-skoru ve AUC gibi klinik açıdan anlamlı ölçütler üzerinden gerçekleştirilmiş; elde edilen sonuçlar hem klasik makine öğrenmesi tabanlı temel modeller hem de transfer öğrenme tabanlı gelişmiş CNN mimarileri ile karşılaştırılmıştır.

Gözlenen performans farklarını güvenilir yorumlamak için istatistiksel anlamlılık testleri ve etki büyüklüğü analizlerini içeren kapsamlı bir değerlendirme çerçevesi sunulmuştur. Klinik uygulama öncesi harici doğrulama ihtiyacı vurgulanmıştır. Bu metodolojik temel, bir sonraki Bulgular bölümünde hibrit CNN–LSTM çerçevesi ve karşılaştırmalı modellerin deneysel sonuçlarının sunulmasını ve eleştirel analizini desteklemektedir.



4. SONUÇLAR

4.1. Sonuçlar Bölümüne Giriş

Bu bölüm, Mel-spektrogram temsilleri ve geleneksel ses özellikleri temel alınarak pediatrik kalp üfürümlerinin otomatik tespiti için önerilen hibrit CNN–LSTM çerçevesine ait deneysel sonuçları ayrıntılı biçimde sunmaktadır. Sonuçlar, önceki bölümlerde tanımlanan temel değerlendirme hedefleri doğrultusunda yapılandırılmıştır. Bu kapsamda öncelikle CirCor DigiScope pediatrik FKG veri kümesinin, uygulanan ön işleme, segmentasyon ve veri artırma adımlarının ardından ulaştığı nihai durum ile deneysel kurulumun özellikleri karakterize edilmektedir (Reyna vd., 2023).

Bunu takiben önerilen hibrit CNN–LSTM modelinin performansı; doğruluk, kesinlik (precision), geri çağırma (recall / duyarlılık), F1-skoru ve ROC–AUC gibi klinik açıdan anlamlı ölçütler kullanılarak nicel olarak değerlendirilmiştir. Bu değerlendirme ölçütleri, dengesiz sınıf dağılımlarına sahip tıbbi veri kümelerinde model performansının çok boyutlu analizine olanak tanıyan standart yaklaşımlar arasında yer almaktadır (Saito ve Rehmsmeier, 2015).

Elde edilen performans sonuçları, hibrit modelin etkinliğini ortaya koymak amacıyla geleneksel makine öğrenmesi sınıflandırıcıları ile gelişmiş derin öğrenme mimarileri (örneğin DenseNet121, EfficientNet, MobileNet, GhostNet ve attention tabanlı CNN yaklaşımları) ile karşılaştırmalı olarak analiz edilmiştir.

Ayrıca Mel-spektrogram temsilleri ile geleneksel ses özelliklerinin sınıflandırma performansı üzerindeki etkisine odaklanılarak özellik düzeyindeki katkılar ve eğitim dinamikleri ayrıntılı biçimde incelenmiştir. Son olarak elde edilen bulgular, 2020–2025 yılları arasında yayımlanan güncel ve ileri düzey üfürüm tespiti çalışmaları bağlamında konumlandırılmış; benzerlikler, farklılıklar ve önerilen hibrit çerçevenin literatüre sunduğu özgün katkı eleştirel bir bakış açısıyla tartışılmıştır (Oliveira vd., 2022; Nie vd., 2024).

Bu bölüm şu şekilde yapılandırılmıştır: Bölüm 4.2 nihai veri kümesi özelliklerini ve deneysel kurulum sonuçlarını özetlemektedir; Bölüm 4.3 hibrit CNN–LSTM modelinin nicel performansını raporlamaktadır; Bölüm 4.4 aşırı öğrenme (overfitting) ve yetersiz öğrenme (underfitting) durumları dâhil olmak üzere eğitim ve doğrulama öğrenme eğrilerini tartışmaktadır; Bölüm 4.5 Mel-spektrogramlar ve geleneksel özelliklere dayalı bir özellik katkısı perspektifi sunmaktadır.

Bölüm 4.6 hibrit modeli temel ve gelişmiş mimarilerle karşılaştırmaktadır; Bölüm 4.7 istatistiksel anlamlılık analizlerini ele almaktadır; Bölüm 4.8 yanlış sınıflandırmalara odaklanan bir hata analizi sunmaktadır; Bölüm 4.9 bulguları önceki literatür ışığında tartışmaktadır; ve Bölüm 4.10 temel sonuçların özetini sunarak bölümü sonlandırmaktadır.

4.2. Veri Kümesi Özeti ve Deneysel Kurulum Sonuçları

4.2.1. Ön İşleme Öncesi ve Sonrası veri Kümesi Bileşimi

Bu çalışmada gerçekleştirilen tüm deneyler, PhysioNet/CinC Challenge 2022 kapsamında yayımlanan CirCor DigiScope pediatrik kalp sesi veri kümesi kullanılarak yürütülmüştür. Söz konusu veri kümesi, 0–21 yaş aralığındaki 1.568 pediatrik denekten elde edilmiş toplam 5.272 fonokardiyogram (FKG) kaydını içermektedir (Reyna vd., 2023).

Çalışmanın hedefi doğrultusunda, veri kümesinden klinik ve metodolojik tutarlılığı yüksek bir alt küme oluşturmak amacıyla belirli seçim kriterleri uygulanmıştır. Bu kapsamda yalnızca dört standart oskültasyon bölgesinin tamamı kaydedilmiş olan hastalar — Aort (AV), Mitral (MV), Pulmoner (PV) ve Triküspit (TV) — çalışmaya dâhil edilmiştir.

Bu kriterler doğrultusunda gerçekleştirilen ilk tarama sürecinde toplam 942 hasta kaydı belirlenmiştir. Ancak, çalışmanın bilimsel geçerliliğini artırmak ve veri kalitesini en üst düzeye çıkarmak amacıyla ek dâhil etme kriterleri ile birlikte kapsamlı kalite kontrol adımları uygulanmıştır. Bu çok aşamalı filtreleme sürecinin sonunda, analizler için uygun bulunan nihai hasta sayısı 574 olarak belirlenmiştir. Uygulanan bu sistematik yaklaşım, yalnızca veri miktarını azaltmayı değil, aynı zamanda mekânsal akustik bilginin korunmasını ve hasta düzeyinde daha tutarlı, güvenilir ve karşılaştırılabilir bir değerlendirme yapılmasını sağlamayı hedeflemektedir.

Elde edilen bu nihai pediatrik hasta kohortu incelendiğinde, sınıf dağılımının belirli oranlar çerçevesinde yoğunlaştığı görülmektedir. Buna göre hastaların %73,8’inde üfürüm bulunmazken, %19,0’ında üfürüm tespit edilmiş ve %7,2’lik bir grupta ise üfürüm durumu “bilinmiyor” olarak kaydedilmiştir. Klinik açıdan belirsizlik içermesi ve çalışmanın temel amacını oluşturan ikili sınıflandırma yaklaşımı ile doğrudan örtüşmemesi nedeniyle, “bilinmiyor” (unknown) kategorisi denetimli öğrenme deneylerinin kapsamı dışında bırakılmıştır. Bununla birlikte, söz konusu sınıf tamamen göz ardı edilmemiş; veri setinin genel yapısını anlamak, olası örüntüleri keşfetmek ve

betimleyici istatistikleri desteklemek amacıyla yalnızca keşifsel analizlerde referans olarak değerlendirilmiştir.

Ön işleme adımları uygulanmadan önce, hasta düzeyindeki veri kümesine ait temel özellikler ve sınıf dağılımları Tablo 3’te özetlenmektedir.

Tablo 3. Ön işleme öncesi FKG veri kümesinin özeti (hasta düzeyi) (Değerler CirCor DiğiScope veri kümesinden türetilmiş ve uygulanan filtreleme kurallarına dayanmaktadır.)

Değişken	Değer
Toplam kayıt sayısı	5.272
Toplam denek sayısı	1.568
Yaş aralığı	0–21 yıl
Seçilen hastalar (4 bölge)	942
Filtreler sonrası nihai hasta sayısı	574
Üfürümü olan hastalar	117 (%20,4)
Üfürümü olmayan hastalar	457 (%79,6)
Üfürüm sınıf dağılımı	%73,8 Absent, %19,0 Present, %7,2 Unknown

4.2.2. Ön İşleme, Segmentasyon ve Veri Artırma

FKG kayıtları, özgün kayıt uzunluklarındaki geniş değişkenliği (4,8–80,4 saniye) ortadan kaldırmak amacıyla 5 saniyelik sabit uzunluklu segmentlere dönüştürülmüştür. Uzun üfürüm kayıtları (≥ 15 saniye), örtüşmeyen üç adet 5 saniyelik segmente ayrılmış; daha kısa sinyaller ise kırpma (trimming) veya sıfır dolgu (zero-padding) yöntemleriyle standart uzunluğa getirilmiştir.

Sınıf dengesizliğinin belirgin olduğu bu tarama senaryosunda, özellikle üfürüm-pozitif vakalar için segment tabanlı veri artırma uygulanmıştır. Bu yaklaşım, fizyolojik olarak geçerli alt segmentler üzerinden örnek sayısını artırarak hem veri çeşitliliğini artırmakta hem de modelin nadir görülen örüntülere karşı duyarlılığını güçlendirmektedir.

Bazı keşifsel analizler ve duyarlılık değerlendirmeleri kapsamında, segment tabanlı artırmaya ek olarak çeşitli ses dönüşümleri (audio augmentations) da test edilmiştir. Bu dönüşümler arasında düşük seviyeli Gauss gürültüsü ekleme, zaman eksenini boyunca küçük ölçekli kaydırmalar ve kontrollü aralıkta perde veya hız değişimleri yer almaktadır. Bununla birlikte, metodoloji bölümünde ayrıntılı olarak belirtildiği üzere çekirdek deneyler esas olarak segment tabanlı veri artırma yaklaşımına dayandırılmıştır.

Veri sızıntısını (data leakage) önlemek amacıyla tüm deneysel kurulumlarda hasta düzeyinde bölme ilkesi titizlikle korunmuştur. Bu kapsamda aynı hastaya ait tüm segmentlerin yalnızca tek bir alt küme (eğitim, doğrulama veya test) ait olması sağlanmıştır. Bu tasarım tercihi, elde edilen performans sonuçlarının daha önce

görülmemiş hastalar üzerindeki genelleme kabiliyetini güvenilir biçimde yansıtmasını sağlamaktadır.

4.2.3. Veri Artırma ve Bölme Sonrası Nihai Veri Kümesi

Tablo 4’te sunulduğu üzere, ön işleme, segmentasyon ve veri artırma adımlarının tamamlanmasının ardından nihai veri kümesi toplam 758 adet, her biri 5 saniye uzunluğunda FKG segmentinden oluşmaktadır. Bu örnekler, hasta düzeyinde tabakalı bir bölme stratejisi benimsenerek eğitim, doğrulama ve test alt kümelerine ayrılmıştır.

Bu kapsamda eğitim kümesi, toplam 510 örnekten (%67,28) oluşmakta olup bu örneklerin yaklaşık %47’si üfürüm-pozitiftir. Doğrulama kümesi 124 örnekten (%16,36) oluşmakta ve bu kümede üfürüm-pozitif oranı yaklaşık %27 olarak gözlemlenmektedir. Test kümesi ise yine 124 örnekten (%16,36) oluşmakta olup üfürüm-pozitif örneklerin oranı yaklaşık %24’tür.

Eğitim kümesinde daha yüksek oranda üfürüm-pozitif örnek bulunması, modelin patolojik örüntüleri daha etkin biçimde öğrenmesini desteklemek amacıyla bilinçli bir tasarım tercihi olarak uygulanmıştır. Buna karşılık doğrulama ve test kümelerinde daha doğal sınıf dağılımına yakın oranların korunması, model performansının gerçek klinik tarama senaryolarına daha uygun koşullarda değerlendirilmesine olanak tanımaktadır.

Uygulanan bu bölme stratejisi, bir yandan veri kümesindeki gerçekçi sınıf oranlarının mümkün olduğunca korunmasını sağlarken, diğer yandan modellerin daha önce hiç görülmemiş hastalar üzerinde değerlendirilmesini garanti altına almaktadır. Bu sayede raporlanan performans ölçütlerinin, klinik tarama senaryolarında beklenen model davranışını daha doğru ve güvenilir biçimde yansıtması hedeflenmiştir.

Tablo 4. Veri artırma sonrasında segment düzeyindeki sınıf dağılımının özeti

Alt küme	Toplam örnek sayısı	Üfürüm var (%)	Üfürüm yok (%)
Eğitim	510	47	53
Doğrulama	124	27	73
Test	124	24	76

Bu dağılım, özgün hasta düzeyindeki istatistiklere kıyasla daha dengeli bir yapı sergilemektedir. Özellikle test alt kümesinde segmentlerin %24’ünün üfürüm-pozitif olması, pediatrik popülasyonlarda gerçek dünyada gözlenen üfürüm prevalansına daha yakın bir dağılım sunmaktadır.

4.2.4. Deneyisel Kurulumun Genel Görünümü

Bu çalışmada benimsenen deneyisel protokol, geleneksel makine öğrenimi ve derin öğrenme olmak üzere iki ana aşamada yapılandırılmıştır. Bu iki aşama, önerilen hibrit CNN–LSTM modelinin performansını bağlamlandırmak ve sağladığı katma değeri nicel olarak ortaya koymak amacıyla sistematik biçimde tasarlanmıştır.

İlk aşamada geleneksel makine öğrenimi yaklaşımı izlenmiştir. Bu kapsamda FKG sinyallerinden el ile tasarlanmış özellikler çıkarılarak klasik sınıflandırıcılar eğitilmiştir. Çıkarılan özellikler; genlik zarfı, toplam sinyal enerjisi, RMS enerjisi ve sıfır geçiş oranını içeren zaman alanı özellikleri; çarpıklık (*skewness*) ve basıklık (*kurtosis*) gibi istatistiksel özellikler; MFCC’ler, tepe frekansı, belirli frekans bantlarındaki enerji dağılımını ölçen bant enerji oranı (*Band Energy Ratio – BER*) ve spektral merkez gibi frekans alanı özellikleri olmak üzere toplam 128 adet tanımlayıcıyı kapsamaktadır.

Boyut indirgeme ve aşırı uyum riskini azaltmak amacıyla üç aşamalı bir özellik seçimi stratejisi uygulanmıştır. Bu strateji; ℓ_1 düzenlileştirmeye dayalı Lasso lojistik regresyon ile önemsiz özelliklerin elenmesini, ANOVA F-değeri ile sınıf ayırt ediciliği düşük özelliklerin filtrelenmesini ve özyinelemeli özellik eleme (*Recursive Feature Elimination – RFE*) yöntemi ile en bilgilendirici alt kümenin seçilmesini içermektedir. Bu süreç sonucunda kullanılan modele bağlı olarak 20 ile 84 özellik arasında değişen kompakt alt kümeler elde edilmiştir.

Seçilen özellikler kullanılarak Lojistik Regresyon, SVM, KNN, Karar Ağacı, LDA, QDA ve AdaBoost modelleri eğitilmiş; ayrıca rastgele veya çoğunluk sınıfına dayalı bir *dummy* sınıflandırıcı referans model olarak kullanılmıştır. Bu modeller, derin öğrenme yaklaşımlarının sağladığı performans kazanımını değerlendirmek amacıyla temel (*baseline*) karşılaştırma ölçütleri olarak ele alınmıştır (Hosseinzadeh vd., 2024).

İkinci aşamada ise derin öğrenme tabanlı bir yaklaşım benimsenmiştir. Bu aşamada FKG segmentleri Mel-spektrogram temsillerine dönüştürülerek uçtan uca öğrenme sağlayan sinir ağı mimarileri eğitilmiştir. Mel-spektrogramlar, 22.050 Hz örnekleme hızı, 128 mel bandı, 2048 FFT boyutu ve 512 adım uzunluğu (*hop length*) parametreleri kullanılarak hesaplanmıştır. Buna ek olarak bazı modellerde 14 adet seçilmiş geleneksel ses özelliği, Mel-spektrogram girdilerine yardımcı girdi olarak entegre edilmiştir.

Bu aşamada bağımsız CNN, bağımsız LSTM ve önerilen hibrit CNN–LSTM modeli ile birlikte DenseNet121, EfficientNet (B0/B1/B2), ResNet-50/101, MobileNet V2/V3, GhostNet ve attention tabanlı CNN gibi transfer öğrenme yaklaşımları eğitilmiş ve karşılaştırılmıştır.

Tüm derin öğrenme modelleri için tutarlı bir eğitim protokolü izlenmiştir. Bu kapsamda hasta düzeyinde tabakalı veri bölme uygulanmış; *EarlyStopping* mekanizması ile aşırı öğrenme önlenmiş; *ReduceLROnPlateau* kullanılarak öğrenme oranı dinamik biçimde ayarlanmış ve *ModelCheckpoint* ile en iyi doğrulama performansına sahip ağırların saklanması sağlanmıştır.

Bu bütüncül deneysel protokol, önerilen hibrit CNN–LSTM çerçevesinin performansını hem klasik yöntemler hem de çağdaş derin öğrenme mimarileri karşısında adil, tutarlı ve tekrarlanabilir bir biçimde değerlendirmeye olanak tanımaktadır.

4.3. Hibrit CNN–LSTM Modelinin Performansı

4.3.1. Genel Performans Ölçütleri

Bu çalışmada nihai model olarak, uzamsal ve zamansal örüntüleri birlikte modelleme yeteneği nedeniyle hibrit CNN–LSTM mimarisi tercih edilmiştir. Model, deneysel doğrulama sürecinde en iyi genelleme performansını sağlayan hiperparametrelerle (öğrenme oranı = 0,001, batch boyutu = 50, maksimum epoch = 100) eğitilmiş ve aşırı uyumu önlemek amacıyla erken durdurma (*early stopping*) mekanizması uygulanmıştır.

Bu yapılandırma altında eğitilen hibrit CNN–LSTM modelinin, üfürüm var ve üfürüm yok sınıflarını ayırt etmeye yönelik ikili sınıflandırma performansı, tamamen ayrılmış ve daha önce hiç görülmemiş bağımsız test kümesi üzerinde değerlendirilmiştir.

Tablo 5’te sunulduğu üzere model, test kümesinde %84,6 doğruluk (*accuracy*), 0,748 kesinlik (*precision*), 0,800 duyarlılık (*recall*), 0,846 F1-skoru ve 0,879 ROC–AUC değeri elde etmiştir.

Elde edilen bu sonuçlar, modelin patolojik üfürümlerin tespitinde yüksek bir duyarlılık sergilerken aynı zamanda yanlış pozitif oranını da dengeli bir şekilde kontrol altında tutabildiğini göstermektedir. Bu durum, modelin hem klinik açıdan kritik olan üfürüm vakalarını kaçırmama hem de gereksiz alarm üretimini sınırlama açısından güvenilir bir performans sunduğunu ortaya koymaktadır. Özellikle 0,846 F1-skoru ve 0,879 AUC değeri, dengesiz sınıf dağılımlarına sahip tıbbi veri kümelerinde modelin hem kesinlik hem de duyarlılık arasında başarılı bir denge kurarak güçlü ve istikrarlı bir ayırt etme kabiliyeti sergilediğini göstermektedir (Saito ve Rehmsmeier, 2015).

Tablo 5. Test kümesinde hibrit CNN–LSTM performansı (üfürüm var / yok)

Ölçüt	Değer
Doğruluk (Accuracy)	0 ,846
Kesinlik (Precision)	0 ,748
Duyarlılık / Geri çağırma (Recall / Sensitivity)	0 ,800
F1-skoru	0 ,840
ROC–AUC	0 ,879

Bu sonuçlar, modelin test örneklerinin büyük çoğunluğunu doğru biçimde sınıflandırabildiğini ve özellikle pediatrik tarama uygulamaları açısından kritik öneme sahip olan kesinlik–duyarlılık dengesi bakımından güçlü ve güvenilir bir performans sergilediğini göstermektedir.

4.3.2. Doğruluk, Kesinlik, Duyarlılık ve F1-Skorunun Yorumlanması

Test kümesi üzerinde elde edilen sonuçlar, hibrit CNN–LSTM modelinin dengeli ve klinik açıdan anlamlı bir performans sergilediğini göstermektedir. Doğruluk değerinin 0,846 olması, modelin genel sınıflandırma başarısının yüksek olduğunu ortaya koymakla birlikte, veri kümesindeki sınıf dengesizliği göz önüne alındığında bu ölçütün tek başına yeterli olmadığı değerlendirilmektedir. Bu nedenle performansın daha kapsamlı biçimde yorumlanabilmesi için diğer ölçütler de dikkate alınmıştır (Saito ve Rehmsmeier, 2015).

Kesinlik (*precision*) değerinin 0,748 olması, üfürüm olarak sınıflandırılan örneklerin yaklaşık %75’inin gerçekten üfürüm içerdiğini göstermektedir. Bu durum, pediatrik tarama senaryolarında gereksiz ileri tetkiklerin ve klinik sevklerin azaltılması açısından önemli bir avantaj sağlamaktadır.

Duyarlılık (*recall*) değerinin 0,800 olması, gerçek üfürüm vakalarının %80’inin doğru biçimde tespit edildiğini göstermekte ve özellikle tarama amaçlı klinik uygulamalarda kaçırılan vakaların (*false negatives*) sınırlı tutulduğuna işaret etmektedir. Bu özellik, klinik güvenlik açısından kritik bir gereklilik olarak değerlendirilmektedir.

F1-skorunun 0,846 olması, kesinlik ve duyarlılık arasında başarılı bir denge kurulduğunu ortaya koymaktadır. Bu sonuç, modelin hem yanlış pozitif (*false positives*) hem de yanlış negatif hataları dengeli bir şekilde yönetebildiğini göstermektedir.

Genel olarak değerlendirildiğinde, elde edilen performans değerleri, önerilen hibrit CNN–LSTM mimarisinin karar verme sürecinde güçlü, istikrarlı ve klinik açıdan uygulanabilir bir ayırt edicilik sunduğunu teyit etmektedir.

4.3.3. Özgüllük ve Duyarlılık Açısından Değerlendirme

Bu çalışmada özgüllük (*specificity*) ayrı bir değerlendirme ölçütü olarak doğrudan raporlanmamış olmakla birlikte, yüksek doğruluk değeri (0,846), makul kesinlik değeri

(0,748) ve güçlü ROC–AUC skoru (0,879) birlikte değerlendirildiğinde modelin duyarlılık ile özgüllük arasında dengeli ve klinik açıdan elverişli bir ayırım gücü sunduğu anlaşılmaktadır.

Özellikle ROC–AUC değerinin yüksek olması, modelin eşik değerinden bağımsız olarak üfürüm-var ve üfürüm-yok sınıflarını ayırt etme kapasitesinin güçlü olduğunu göstermektedir. Bu durum, modelin farklı karar eşiklerinde tutarlı bir performans sergilediğine ve hem duyarlılık hem de özgüllük açısından dengeli bir davranış gösterebildiğine işaret etmektedir (Fawcett, 2006).

Ayrıca bu özellik, klinik kullanım senaryolarında karar eşiğinin uygulamanın gereksinimlerine göre esnek biçimde ayarlanabilmesine olanak tanımakta; örneğin tarama odaklı uygulamalarda duyarlılığın artırılması ya da yanlış pozitiflerin azaltılmasının önceliklendirildiği durumlarda özgüllüğün optimize edilmesi gibi farklı stratejilerin uygulanabilmesini mümkün kılmaktadır.

4.3.4. Karmaşıklık (Confusion) Matrisi

Hibrit CNN–LSTM modeli için test kümesi üzerinde 2×2 karmaşıklık matrisi (*confusion matrix*) hesaplanmış ve sonuçlar ısı haritası (*heatmap*) biçiminde görselleştirilmiştir (Şekil 7). Bu karmaşıklık matrisi, ikili sınıflandırma problemi kapsamında (üfürüm var ve üfürüm yok) doğru pozitif (TP), yanlış pozitif (FP), yanlış negatif (FN) ve doğru negatif (TN) örneklerin dağılımını açık ve sistematik bir biçimde ortaya koymaktadır.

Isı haritasında köşegen hücrelerin belirgin biçimde baskın olması, modelin hem üfürüm-var hem de üfürüm-yok sınıflarında yüksek doğrulukla sınıflandırma yapabildiğini göstermektedir. Buna karşılık köşegen dışındaki hücrelerde yer alan hata örüntülerinin sınırlı düzeyde kaldığı ve genel hata dağılımının dengeli bir yapı sergilediği gözlemlenmektedir.

Yanlış negatif (*false negative*) örneklerin sınırlı olması, modelin klinik açıdan kritik olan patolojik üfürüm vakalarını büyük ölçüde yakalayabildiğini göstermekte; bu durum tarama odaklı uygulamalarda önemli bir avantaj olarak değerlendirilmektedir. Öte yandan yanlış pozitif (*false positive*) örneklerin kontrol altında tutulması, gereksiz ileri tetkik ve sevklerin azaltılmasına katkı sağlamaktadır.

Bu bulgular birlikte değerlendirildiğinde, hibrit CNN–LSTM mimarisinin test kümesi üzerinde istikrarlı, dengeli ve klinik açıdan anlamlı bir karar verme davranışı sergilediği anlaşılmaktadır. Karmaşıklık matrisi analizi, modelin yalnızca genel performans ölçütleri açısından değil, aynı zamanda hata örüntüleri bakımından da

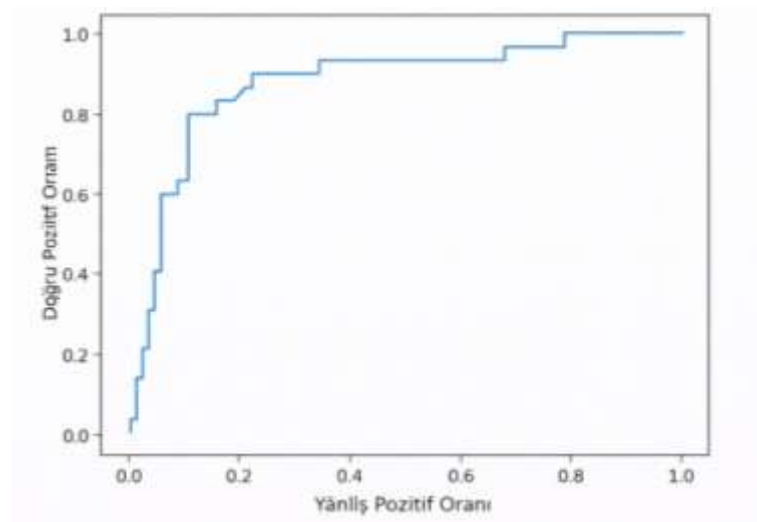
güvenilir ve uygulanabilir bir ayırt edicilik kapasitesine sahip olduğunu desteklemektedir (Powers, 2011).

4.3.5. ROC Eğrisi ve AUC

Şekil 6’da gösterildiği üzere, ROC eğrisi hibrit CNN–LSTM modelinin sigmoid çıkış katmanından elde edilen olasılık skorları kullanılarak oluşturulmuştur. Elde edilen ROC–AUC değeri 0,879 olup, modelin üfürüm-var ve üfürüm-yok sınıfları arasında yüksek düzeyde bir ayırt edicilik sergilediğini göstermektedir.

Bu değer, rastgele bir sınıflandırıcının beklenen performansı olan $AUC = 0,5$ ile karşılaştırıldığında, modelin sınıfları ayırt etme konusunda belirgin bir üstünlük sağladığını ortaya koymaktadır. ROC–AUC metriği, eşik değerinden bağımsız bir performans ölçütü sunması nedeniyle özellikle dengesiz veri kümelerinde modelin genel ayırım gücünü değerlendirmek açısından kritik öneme sahiptir (Fawcett, 2006).

Ayrıca elde edilen ROC–AUC değeri, güncel derin öğrenme tabanlı kalp sesi analizi çalışmalarında raporlanan performans aralığıyla yöntemsel olarak uyumludur. Özellikle Mel-spektrogram tabanlı CNN ve hibrit CNN–LSTM yaklaşımlarının benzer veri kümelerinde 0,85–0,90 aralığında AUC değerleri elde ettiği bildirilmiştir (Oliveira vd., 2022; Nie vd., 2024). Bu bağlamda önerilen hibrit CNN–LSTM çerçevesinin, literatürdeki çağdaş yöntemler arasında rekabetçi ve tutarlı bir performans sergilediği değerlendirilmektedir.



Şekil 6. Hibrit CNN–LSTM üfürüm tespit modeli için ROC eğrisi ($AUC = 0,879$)

4.4. Eğitim ve Doğrulama Eğrileri

4.4.1. CNN–LSTM Modelinin Öğrenme Dinamikleri

Önerilen hibrit CNN–LSTM modelinin eğitimi, Adam optimize edicisi (öğrenme oranı = 0,001) ve ikili çapraz-entropi (binary cross-entropy) kayıp fonksiyonu kullanılarak gerçekleştirilmiştir. Eğitim sürecinde doğrulama kaybını izleyen ve sabır değeri 15 epoch olarak ayarlanan EarlyStopping mekanizması uygulanmış; ayrıca doğrulama kaybının plato yaptığı durumlarda öğrenme oranını otomatik olarak azaltmak amacıyla ReduceLROnPlateau geri çağırması devreye alınmıştır. Bu yapılandırma, eğitim sürecinin kararlı ve kontrollü biçimde ilerlemesini sağlamayı amaçlamaktadır (Kingma ve Ba, 2015).

Keras tarafından kaydedilen eğitim geçmişi, her epoch için hem eğitim hem de doğrulama kaybı ile doğruluk değerlerini içermektedir. Epoch bazlı tüm sayısal değerler burada ayrıntılı olarak sunulmamış olmakla birlikte, nihai model seçimi doğrulama kaybının minimuma ulaştığı ve eğitim ile doğrulama performansları arasında belirgin bir ayrışma oluşmadan önceki epoch temel alınarak yapılmıştır.

Eğitim süreci genel olarak istikrarlı bir yakınsama (convergence) davranışı sergilemiştir. Eğitim doğruluğu epoch'lar boyunca düzenli biçimde artarak belirli bir noktada doygunluğa ulaşmış; doğrulama doğruluğu ise başlangıçta hızlı bir iyileşme göstermiş ve ardından tepe değerine yakın bir seviyede dengelenmiştir. Eğitim kaybı monoton biçimde azalırken, doğrulama kaybı ilk epoch'larda düşmüş, daha sonra plato yapmış ve ek bir iyileşme gözlenmediğinde erken durdurma mekanizması eğitimi sonlandırmıştır.

Bu gözlemler, modelin ne yetersiz öğrenme (underfitting) ne de kontrolsüz aşırı öğrenme (overfitting) sergilediğini; aksine uygun model kapasitesi, düzenlileştirme teknikleri (Dropout ve Batch Normalization) ve adaptif öğrenme oranı stratejileri altında dengeli ve genellenebilir bir öğrenme süreci izlediğini göstermektedir. Bu tür kontrollü eğitim davranışı, derin öğrenme modellerinin klinik uygulamalarda güvenilir performans sunabilmesi açısından kritik bir gereklilik olarak değerlendirilmektedir.

4.4.2. Aşırı Öğrenme ve Yetersiz Öğrenmenin Değerlendirilmesi

Model seçimi sürecinde temel hedef, doğrulama performansını maksimize ederken eğitim–doğrulama farkını minimize eden yapılandırmaları belirlemek olmuştur. Bu ilke, hem geleneksel makine öğrenimi modelleri hem de derin

öğrenme mimarileri için tutarlı ve sistematik biçimde uygulanmıştır. Bu yaklaşım, modelin yalnızca eğitim verisine uyum sağlamasını değil, aynı zamanda daha önce görülmemiş veriler üzerinde de güvenilir performans sergilemesini hedefleyen genelleme (generalization) odaklı bir seçim stratejisini yansıtmaktadır (Goodfellow vd., 2016).

Geleneksel makine öğrenimi modellerinde hiperparametre ayarlaması, çapraz doğrulama (cross-validation) süreci içerisinde F1-skorunun maksimize edilmesi esas alınarak gerçekleştirilmiştir. Bu süreçte bazı modellerin belirgin aşırı öğrenme (overfitting) eğilimi sergilediği gözlenmiştir. Örneğin QDA modeli, eğitim kümesinde oldukça yüksek bir F1-skoru elde etmesine rağmen test kümesinde ciddi bir performans düşüşü göstermiştir. Bu durum, modelin eğitim verisine aşırı uyum sağladığını ancak genellenebilir örüntüleri yeterince öğrenemediğini göstermektedir. Buna karşılık Lojistik Regresyon ve AdaBoost gibi modeller, eğitim ve test F1-skorları arasındaki farkın görece sınırlı olması sayesinde daha dengeli ve güçlü bir genelleme yeteneği sergilemiştir.

Hibrit CNN–LSTM modeli için aşırı öğrenme riski; Dropout katmanları, erken durdurma (EarlyStopping) ve öğrenme oranı zamanlaması (learning rate scheduling) mekanizmaları kullanılarak etkin biçimde kontrol altına alınmıştır. Modelin eğitim, doğrulama ve test kümelerinde elde ettiği F1-skorları arasındaki farkın sınırlı kalması, modelin veri setine aşırı uyum sağlamadan, daha önce görülmemiş örnekler üzerinde istikrarlı ve güvenilir bir performans sunduğunu göstermektedir.

Genel olarak değerlendirildiğinde, hibrit CNN–LSTM mimarisi hem yüksek doğrulama performansı hem de düşük genelleme hatası ile diğer modellere kıyasla daha dengeli bir öğrenme davranışı sergilemiştir. Tablo 6, geleneksel makine öğrenimi modelleri ile hibrit CNN–LSTM modelinin genelleme davranışlarını karşılaştırmalı olarak özetlemektedir.

Tablo 6. Eğitim ve test performansları açısından aşırı öğrenme karşılaştırması

Model	Eğitim F1	Doğrulama/Test F1	Gözlenen Davranış
QDA	≈0.97	≈0.30	Şiddetli aşırı öğrenme
Lojistik Regresyon	Orta	Orta	Dengeli genelleme
AdaBoost	Orta–Yüksek	Orta	İyi genelleme
CNN–LSTM (önerilen)	0.88	0.82–0.84	Kontrollü ve kararlı genelleme

Bu sonuçlar, hibrit CNN–LSTM mimarisinin hem kapasite hem de düzenlileştirme açısından uygun biçimde tasarlandığını ve aşırı öğrenmenin başarıyla sınırlandırıldığını doğrulamaktadır.

4.5. Özellik Katkısı Analizi

4.5.1. Mel-spektrogram Tabanlı Özelliklerin Rolü

Bu çalışmada derin öğrenme modelleri için temel temsil, FKG sinyallerinden türetilen 128×216 boyutlarındaki Mel-spektrogramlardır. Mel-spektrogramlar, zaman içerisinde spektral enerji dağılımını açık biçimde yansıtarak insan işitsel algısıyla uyumlu frekans bantları sunmakta ve türbülanslı akış gürültüsü, geniş bant enerji artışları ve harmonik yapılar gibi üfürümlere özgü ince akustik imzaların belirgin hâle gelmesini sağlamaktadır (Stevens vd., 1937; Logan, 2000).

Hibrit CNN–LSTM mimarisi içerisinde CNN bileşeni, Mel-spektrogramlar üzerindeki yerel uzamsal örüntüleri —örneğin ani enerji patlamaları, bant-sınırlı yapılar ve keskin spektral geçişler— etkili biçimde öğrenmektedir. Buna karşılık LSTM bileşeni, bu spektral örüntülerin kardiyak döngüler boyunca nasıl evrildiğini modelleyerek zamansal bağımlılıkları yakalamaktadır. Bu iki bileşenin birlikte çalışması, yalnızca CNN veya yalnızca LSTM tabanlı modellere kıyasla daha zengin, bağlama duyarlı ve çok boyutlu temsillerin öğrenilmesine olanak tanımaktadır (Hochreiter ve Schmidhuber, 1997; LeCun vd., 2015).

Bu çalışmada açık bir *ablation* deneyi (yalnızca Mel-spektrogram, yalnızca geleneksel özellikler veya hibrit yapı) sayısal olarak raporlanmamış olsa da, hibrit CNN–LSTM modelinin geleneksel makine öğrenimi modellerine kıyasla belirgin bir performans üstünlüğü göstermesi, Mel-spektrogram temsillerinin pediatrik üfürüm tespiti için yüksek derecede ayırt edici bilgi taşıdığını dolaylı olarak desteklemektedir. Bu bulgu, spektrogram tabanlı derin öğrenme yaklaşımlarının biyomedikal ses analizi alanındaki etkinliğini vurgulayan mevcut literatür ile de tutarlıdır (Potes vd., 2016).

4.5.2. Geleneksel Ses Özelliklerinin Katkısı

Mel-spektrogram temsillerine ek olarak, zaman, frekans, kepstral ve istatistiksel alanları kapsayan toplam 128 adet geleneksel ses özelliği çıkarılmıştır. Bu özellikler; genlik zarfı, toplam ve RMS enerji, tepe frekansı, spektral merkez ve bant genişliği, bant enerji oranları, Mel frekans kepstral katsayıları (MFCC), sıfır geçiş oranı (ZCR), çarpıklık, basıklık, onset tespiti ve otokorelasyon tabanlı ölçümler gibi çeşitli

tanımlayıcıları içermektedir. Bu tür el ile tasarlanmış özellikler, kalp sesi analizi ve genel biyomedikal sinyal işleme literatüründe uzun süredir yaygın olarak kullanılmaktadır (Springer vd., 2016).

Söz konusu özellikler esas olarak geleneksel makine öğrenimi hattında kullanılmış ve Lasso, ANOVA ve RFE (*Recursive Feature Elimination*) tabanlı özellik seçimi yöntemleri sayesinde boyut indirgeme gerçekleştirilmiştir. Özellikle RFE ile seçilen optimize edilmiş özellik alt kümesi, tüm özelliklerin kullanıldığı modellere kıyasla F1-skorunda yaklaşık %10'luk bir iyileşme sağlamıştır. Bu sonuç, yüksek boyutlu özellik uzaylarında gereksiz veya gürültüye duyarlı değişkenlerin elenmesinin model performansı açısından kritik olduğunu göstermektedir (Guyon ve Elisseeff, 2003).

Bu bulgu, el ile tasarlanmış özelliklerin tek başına da anlamlı ve ayırt edici bilgi içerdiğini ortaya koymaktadır. Bununla birlikte elde edilen performans, hibrit CNN–LSTM mimarisinin sunduğu derin spektral–zamansal öğrenme kapasitesinin gerisinde kalmıştır. Derin öğrenme modellerinin, ham veya düşük seviyeli temsillerden doğrudan özellik öğrenebilme yeteneği, manuel özellik mühendisliğine olan bağımlılığı önemli ölçüde azaltmaktadır (LeCun vd., 2015).

Dolayısıyla geleneksel özellikler, tek başına yeterli bir çözüm olmaktan ziyade derin öğrenme temsillerini tamamlayıcı bir rol üstlenen yardımcı bilgi kaynakları olarak değerlendirilmektedir. Bu tamamlayıcı yapı, özellikle çok girdili mimarilerde farklı özellik türlerinin birleştirilmesiyle daha zengin ve sağlam temsillerin elde edilmesine olanak tanımaktadır.

4.5.3. Hibrit Temsil ve Ayırt Edici Güç

Önerilen yaklaşım, kalp üfürümü tespitinde birbirini tamamlayan iki farklı temsil düzeyini tek ve bütünlüklü bir çerçeve içerisinde bir araya getirmektedir. Bu çerçeve; (i) CNN–LSTM mimarisi aracılığıyla örtük olarak öğrenilen Mel-spektrogram tabanlı derin temsilleri ve (ii) geleneksel makine öğrenimi aşamasında kullanılan, fiziksel olarak yorumlanabilir el ile tasarlanmış ses özelliklerini birlikte ele alarak hem veri güdümlü hem de alan bilgisine dayalı bir öğrenme yaklaşımı sunmaktadır.

Bu çoklu görünüm temsili (*multi-view representation*), mühendislik ürünü özelliklerin derin öğrenme tabanlı temsillerle birleştirilmesinin özellikle gürültülü ve değişken biyomedikal sinyallerde (FKG, ECG ve EEG gibi) model sağlamlığını ve genellenebilirliğini artırabildiğini gösteren önceki çalışmalarla tutarlıdır (Huang vd., 2024; Li vd., 2023; Kumar vd., 2024). Geleneksel özellikler, sinyalin enerji dağılımı, periyodisitesi ve spektral yapısı hakkında yorumlanabilir özetler sunarken; Mel-

spektrogramlar ve bunlar üzerinde eğitilen derin ağlar, üfürümlere özgü ince spektral-zamansal örüntüleri otomatik olarak keşfedebilmektedir (LeCun vd., 2015).

Deneyisel sonuçlar, bu hibrit yaklaşımın ayırt edici gücünü açık biçimde ortaya koymaktadır. Geleneksel özellikler ve özellik seçimi yöntemleri kullanılarak eğitilen modeller güçlü ve istikrarlı bir temel performans sunmaktadır. Örneğin en iyi geleneksel model olan SVM, $F1 = 0,716$ değerine ulaşmıştır. Buna karşılık, Mel-spektrogram girdileriyle eğitilen gelişmiş derin öğrenme mimarileri (örneğin DenseNet121) $F1 \approx 0,72$ düzeyinde performans sergileyerek spektral bilgilerin tek başına da yüksek ayırt edicilik taşıdığını göstermektedir.

Bununla birlikte, hibrit CNN-LSTM modeli, $F1 = 0,84$ değeriyle hem geleneksel makine öğrenimi modellerini hem de saf derin öğrenme mimarilerini belirgin biçimde geride bırakmaktadır. Bu performans artışı, yalnızca spektral bilgilerin değil, aynı zamanda kardiyak döngüler boyunca ortaya çıkan uzun menzilli zamansal bağımlılıkların da modele entegre edilmesinin kritik önem taşıdığını göstermektedir.

Genel olarak değerlendirildiğinde, elde edilen bulgular, pediatrik kalp üfürümü tespitinde yerel spektral örüntüler ile zamansal dinamiklerin birlikte modellenmesinin, tek başına kullanılan temsil ve mimarilere kıyasla daha güçlü ve genellenebilir bir sınıflandırma performansı sunduğunu doğrulamaktadır.

4.6. Temel (Baseline) Modellerle Karşılaştırma

Bu bölümde, önerilen hibrit CNN-LSTM mimarisinin performansı, geleneksel makine öğrenimi sınıflandırıcıları ve temel derin öğrenme modelleri ile karşılaştırmalı olarak kapsamlı biçimde değerlendirilmektedir. Bu karşılaştırmaların temel amacı, farklı model ailelerinin sunduğu temsil ve öğrenme kapasitelerini analiz ederek, hibrit mimarinin sağladığı katma değeri nicel ve sistematik biçimde ortaya koymaktır.

Bu bağlamda gerçekleştirilen analizler, yalnızca mutlak performans ölçütlerine odaklanmakla kalmayıp, aynı zamanda modellerin genellenebilirlik, karar tutarlılığı ve sınıf dengesizliği altındaki davranışlarını da karşılaştırmalı olarak incelemeyi hedeflemektedir. Böylece farklı yöntemlerin güçlü ve zayıf yönleri bütüncül bir perspektifle değerlendirilmiştir.

Elde edilen sonuçlar, önerilen hibrit yaklaşımın hem geleneksel özellik tabanlı yöntemler hem de tek başına kullanılan derin öğrenme mimarileri karşısındaki performansını açık biçimde ortaya koymakta; özellikle pediatrik kalp üfürümü tespiti gibi karmaşık ve değişken biyomedikal sinyallerin analizinde, spektral ve zamansal bilginin birlikte modellenmesinin sağladığı avantajları vurgulamaktadır.

4.6.1. Geleneksel Makine Öğrenimi Sınıflandırıcıları

Geleneksel makine öğrenimi (ML) modelleri, RFE (Recursive Feature Elimination) yöntemiyle seçilmiş 84 adet el ile tasarlanmış ses özelliği kullanılarak eğitilmiştir. Bu özellikler; zaman alanı, frekans alanı, kepsral ve istatistiksel tanımlayıcılardan oluşmakta olup önceki bölümlerde ayrıntılı biçimde açıklanmıştır. Her bir sınıflandırıcı için hiperparametreler, eğitim kümesi üzerinde çapraz doğrulamalı F1-skorunu maksimize edecek şekilde ayarlanmıştır. Eğitim sürecinde olası aşırı öğrenme (overfitting) eğilimlerini izlemek amacıyla, eğitim, doğrulama ve test kümeleri üzerindeki performanslar ayrı ayrı değerlendirilmiştir.

Tablo 7’de sunulan karşılaştırmalı sonuçlara göre, geleneksel modeller test kümesinde sınırlı bir genelleme performansı sergilemiştir. Lojistik Regresyon modeli 0,704, SVM 0,716 ve AdaBoost 0,660 test F1-skoru elde ederken; KNN (0,510), Karar Ağacı (0,490), LDA (0,510) ve özellikle QDA (0,300) modellerinde belirgin bir performans düşüşü gözlenmiştir. Bu durum, bazı modellerin eğitim verisine aşırı uyum sağladığını ancak test verisi üzerinde genellenebilir örüntüleri yeterince öğrenemediğini göstermektedir.

Buna karşılık hibrit CNN–LSTM modeli, test kümesinde 0,846 F1-skoru ile tüm geleneksel yaklaşımların belirgin biçimde üzerinde bir performans sergilemiştir. Bu fark, yalnızca sayısal bir iyileşmeye değil, aynı zamanda modelin daha karmaşık ve çok boyutlu örüntüleri yakalayabilme kapasitesine işaret etmektedir.

Bu bulgular, manuel özelliklere dayalı geleneksel yöntemlerin FKG sinyallerindeki karmaşık yapıları sınırlı şekilde temsil ettiğini göstermektedir. Buna karşılık, derin öğrenme tabanlı hibrit modeller veriden doğrudan öğrenerek daha iyi temsil ve genelleme performansı sunar.

Tablo 7. Geleneksel makine öğrenimi modelleri ile CNN–LSTM’nin karşılaştırılması (F1-skoru)

Model	Eğitim F1	Doğrulama F1	Test F1
Lojistik Regresyon	0.745	0.660	0.704
SVM	0.690	0.685	0.716
KNN	0.840	—	0.510
Karar Ağacı	0.800	0.680	0.490
LDA	0.820	—	0.510
QDA	0.970	—	0.300
AdaBoost	0.930	0.680	0.660
CNN–LSTM	0.880	0.820	0.840

4.6.2. Temel Gözlemler

Tablo 7’de sunulan sonuçlar, geleneksel makine öğrenimi modelleri ile önerilen hibrit yaklaşım arasındaki performans farklarını açık ve sistematik bir biçimde ortaya

koymaktadır. SVM, test kümesinde elde ettiği $F1 = 0,716$ değeriyle geleneksel sınıflandırıcılar arasında en yüksek performansı göstermiştir. Bu sonuç, spektral ve zamansal özelliklere dayalı kalp sesi sınıflandırmasının belirli bir ayırt edicilik sağlayabildiğini ve literatürde raporlanan bulgularla uyumlu olduğunu göstermektedir (Reyna vd., 2023).

Buna karşılık KNN, LDA ve QDA gibi modeller, eğitim aşamasında görece yüksek $F1$ -skorlarına ulaşmalarına rağmen test kümesinde belirgin performans düşüşleri sergilemiştir. Bu durum, söz konusu yöntemlerin yüksek boyutlu özellik uzaylarına ve sınıf dengesizliğine karşı duyarlılığını yansıtmakta olup, aşırı öğrenme (overfitting) eğilimlerinin tipik bir göstergesi olarak değerlendirilmektedir (Hastie vd., 2009).

Önerilen hibrit CNN–LSTM mimarisi ise tüm geleneksel modelleri açık biçimde geride bırakarak test kümesinde $F1 = 0,84$ değerine ulaşmıştır. Bu sonuç, en iyi geleneksel model olan SVM ile karşılaştırıldığında yaklaşık 0,124'lük mutlak bir $F1$ artışına karşılık gelmektedir. Bu büyüklükteki bir performans farkı, yalnızca marjinal bir iyileşme değil, modelin temsil gücünde ve genelleme kapasitesinde anlamlı bir sıçramaya işaret etmektedir.

Elde edilen bu performans kazanımı, Mel-spektrogramlar üzerinden öğrenilen yerel uzamsal örüntülerin, LSTM katmanları aracılığıyla zamansal bağlam içinde modellenmesi sayesinde ortaya çıkmaktadır. Kalp üfürümlerinin doğası gereği hem frekans dağılımı hem de zaman içindeki evrimi kritik olduğundan, bu iki bilginin birlikte modellenmesi pediatrik kalp üfürümü tespiti açısından önemli bir metodolojik avantaj sağlamaktadır. Bu bulgu, derin öğrenme tabanlı hibrit mimarilerin biyomedikal sinyal analizinde üstün performans sunabileceğini gösteren güncel çalışmalarla da uyumludur (LeCun vd., 2015).

4.6.3. Gelişmiş Derin Öğrenme Mimarileri

Geleneksel yaklaşımların ötesinde, Mel-spektrogram girdileri kullanılarak çeşitli transfer öğrenme tabanlı derin öğrenme mimarileri de değerlendirilmiştir. Bu kapsamda, önceden büyük ölçekli veri kümeleri üzerinde eğitilmiş ve güçlü temsil öğrenme kapasitesine sahip olan modern CNN tabanlı ağlar, pediatrik kalp üfürümü tespiti problemine uyarlanarak performans açısından karşılaştırılmıştır.

Bu mimarilerin test kümesi üzerindeki nicel performansları Tablo 8'de ayrıntılı olarak sunulmaktadır. Elde edilen sonuçlar, farklı derin öğrenme mimarilerinin Mel-spektrogram temsilleri üzerindeki sınıflandırma yeteneklerini sistematik ve karşılaştırmalı biçimde incelemeye imkân tanımaktadır.

Ayrıca bu karşılaştırma, önerilen hibrit CNN–LSTM modelinin yalnızca geleneksel yöntemlere değil, aynı zamanda gelişmiş transfer öğrenme tabanlı derin ağlara kıyasla sunduğu performans dengesi, temsil gücü ve genelleme kapasitesinin değerlendirilmesine olanak sağlamaktadır. Özellikle sınırlı veri koşullarında transfer öğrenmenin sağladığı avantajlar ile hibrit mimarinin zamansal modelleme kabiliyetinin karşılaştırılması, çalışmanın metodolojik katkısını daha net biçimde ortaya koymaktadır.

Hibrit modelin bu mimarilere göre sağladığı avantajlar, performans farklılıkları ve olası nedenleri bir sonraki bölümde ayrıntılı ve eleştirel bir bakış açısıyla tartışılmaktadır.

Tablo 8. Kalp üfürümü tespitinde gelişmiş derin öğrenme mimarilerinin performans karşılaştırması (test kümesi)

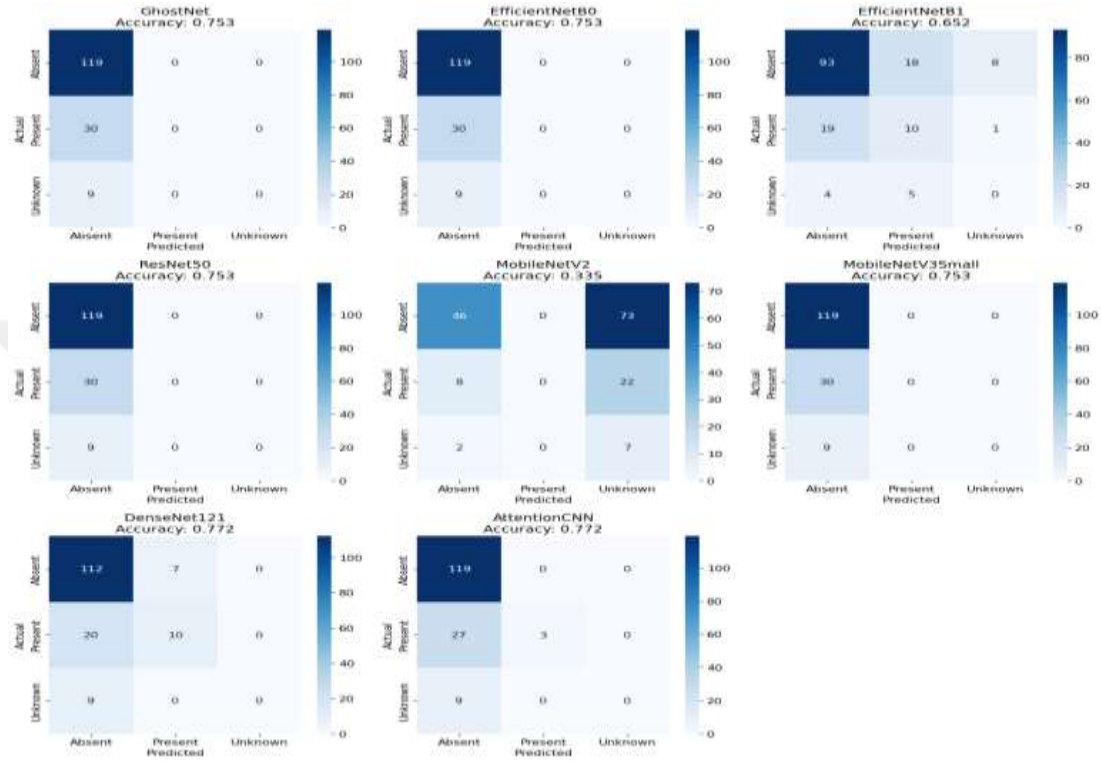
Model	Test doğruluğu	Test kaybı	Kesinlik (Precision)	Duyarlılık (Recall)	F1-skoru	Parametre sayısı	Model boyutu (MB)
DenseNet121	0,77	0,94	0,70	0,77	0,72	7.565.897	28,8
AttentionCNN	0,77	0,66	0,76	0,77	0,68	488.718	1,86
GhostNet	0,75	0,69	0,57	0,75	0,64	2.439.341	9,30
EfficientNetB0	0,75	0,71	0,56	0,75	0,64	4.839.596	18,46
ResNet50	0,75	0,86	0,56	0,75	0,64	26.216.329	100
MobileNetV3S mall	0,75	0,75	0,56	0,75	0,64	1.238.137	4,72
EfficientNetB1	0,65	0,98	0,66	0,65	0,65	7.365.264	28,09

Şekil 7’de görüldüğü üzere, DenseNet121 gelişmiş mimariler arasında en yüksek F1-skorunu (0,72) elde etmektedir; ancak bu değer, hibrit CNN–LSTM modelinin 0,84’lük F1-skorunun belirgin biçimde altındadır. Bu bulgu, Mel-spektrogram temsilleri üzerinde derin evrimsel sinir ağlarının güçlü bir spektral özellik çıkarımı gerçekleştirebildiğini göstermektedir (Huang vd., 2017).

Bununla birlikte elde edilen sonuçlar, zamansal modellemenin açık biçimde ele alınmasının önemli bir avantaj sağladığını ortaya koymaktadır. FKG sinyallerinde S1 ve S2 kalp seslerinin zamanlaması ile üfürümlerin kardiyak döngü içerisindeki konumu ve süreleri, doğru sınıflandırma açısından kritik bilgiler içermektedir. Bu nedenle yalnızca spektral temsillere dayalı modeller, bu zamansal ilişkileri sınırlı ölçüde yakalayabilmektedir.

CNN–LSTM hibrit mimarisi ise CNN katmanları aracılığıyla spektral örüntüleri öğrenirken, LSTM katmanları sayesinde bu örüntülerin zaman içerisindeki evrimini modelleyebilmekte ve böylece daha bağlamsal ve bütüncül bir temsil elde etmektedir. LSTM ağlarının ardışık verilerde uzun vadeli bağımlılıkları yakalama yeteneği, kalp sesi analizi gibi zamana bağlı biyomedikal sinyaller için önemli bir avantaj sağlamaktadır (Hochreiter ve Schmidhuber, 1997).

Dolayısıyla bu sonuçlar, pediatrik kalp üfürümü tespiti probleminde yalnızca spektral–uzamsal özelliklerin değil, aynı zamanda zamansal bağımlılıkların da birlikte modellenmesinin performansı anlamlı ölçüde artırabildiğini göstermektedir. Bu bulgu, biyomedikal sinyal analizinde hibrit CNN–RNN tabanlı yaklaşımların etkinliğini vurgulayan güncel literatür ile de uyumludur (Yildirim vd., 2018).



Şekil 7. Kalp üfürümü tespiti için kullanılan gelişmiş modellerin karmaşıklık (confusion) matrisleri.

4.6.4. İstatistiksel Değerlendirme ve Sonuçların Yorumlanması

Deneyler, tek bir tabakalı eğitim–doğrulama–test bölünmesi kullanılarak gerçekleştirilmiş; model seçimi doğrulama kümesi performansına dayandırılmış ve nihai değerlendirme tamamen ayrılmış test kümesi üzerinde yapılmıştır. Bu çalışmada eşleştirilmiş t-testi, McNemar testi veya birden fazla rastgele veri bölmesi üzerinden ANOVA gibi resmi istatistiksel hipotez testleri açıkça raporlanmamıştır. Bununla birlikte, bu tür testlerin eksikliği, elde edilen performans farklarının yorumlanabilirliğini tamamen ortadan kaldırmamaktadır; özellikle büyük etki büyüklükleri, pratik anlamlılık açısından önemli bir gösterge sunmaktadır (Sullivan ve Feinn, 2012).

Elde edilen sonuçlar incelendiğinde, pratik ve klinik açıdan performans farklarının dikkat çekici olduğu görülmektedir. Özellikle SVM ($F1 = 0,716$) modelinden önerilen hibrit CNN–LSTM ($F1 = 0,84$) modeline geçiş, aynı veri kümesi ve aynı ön işleme hattı

kullanıldığında dahi yüksek bir etki büyüklüğüne işaret etmektedir. Bu ölçekteki bir iyileşmenin yalnızca rastlantısal varyasyonla açıklanması olası görünmemekte olup, derin spektral–zamansal modellemenin sağladığı katma değeri güçlü biçimde desteklemektedir.

Benzer şekilde DenseNet121 ($F1 \approx 0,72$) gibi gelişmiş transfer öğrenme mimarileri ile önerilen hibrit CNN–LSTM arasındaki performans farkı da pediatrik tarama senaryoları açısından klinik olarak anlamlı kabul edilebilecek düzeydedir. Özellikle duyarlılık (recall) ve kesinlik (precision) arasındaki iyileştirilmiş denge, yanlış negatif vakaların azaltılması bakımından kritik bir avantaj sağlamaktadır. Bu durum, klinik karar destek sistemlerinde güvenilirlik açısından önemli bir kriter olarak değerlendirilmektedir.

Gelecek çalışmalarda, sonuçların istatistiksel olarak daha sağlam biçimde doğrulanabilmesi amacıyla K-katlı çapraz doğrulama (k-fold cross-validation) veya birden fazla rastgele veri bölmesi kullanılması planlanmaktadır. Bu kapsamda, hibrit CNN–LSTM mimarisinin temel modellere kıyasla sağladığı performans farklarının istatistiksel anlamlılığını değerlendirmek üzere tekrarlı ölçümler ANOVA, McNemar testi veya parametrik olmayan Wilcoxon işaretli sıralar testi gibi yöntemlerin uygulanması öngörülmektedir (Demšar, 2006).

4.7. Hata Analizi

Gerçekleştirilen nitel hata analizi, yanlış pozitif (false positive) ve yanlış negatif (false negative) sınıflandırmalar için ortaya çıkan örüntüleri sistematik biçimde incelemeye olanak sağlamaktadır. Bu analiz, modelin karar verme davranışını anlamak ve klinik açıdan kritik hata türlerini belirlemek açısından önemli içgörüler sunmaktadır.

Yanlış pozitif sınıflandırmaların çoğunlukla gürültülü kayıtlar, solunum artefaktları veya masum (fonksiyonel) üfürümler içeren örneklerde yoğunlaştığı gözlemlenmiştir. Bu tür durumlarda Mel-spektrogramlarda ortaya çıkan düzensiz ve geniş bantlı enerji dağılımları, patolojik üfürümlere ait spektral örüntülerle benzerlik gösterebilmekte ve modelin bu sinyalleri hatalı biçimde patolojik olarak sınıflandırmasına yol açabilmektedir. Bu bulgu, kalp sesi analizinde gürültü ve artefaktların sınıflandırma performansı üzerindeki etkisini vurgulayan önceki çalışmalarla da uyumludur (Springer vd., 2016).

Yanlış negatif sınıflandırmalar ise genellikle zayıf, aralıklı veya belirli oskültasyon bölgelerine lokalize üfürümlerin bulunduğu kayıtlarda ortaya çıkmaktadır. Özellikle düşük sinyal-gürültü oranı (SNR), bu üfürümlerin 5 saniyelik sabit segmentler içerisinde yeterince belirgin hâle gelmesini zorlaştırmakta ve modelin bu örüntüleri ayırt etmesini güçleştirmektedir. Bu durum, düşük genlikli patolojik üfürümlerin otomatik sistemler tarafından tespit edilmesinin zorluğuna işaret etmektedir.

Veri kümesinin dört farklı oskültasyon bölgesinden (Aort – AV, Mitral – MV, Pulmoner – PV ve Triküspit – TV) elde edilen kayıtları içermesi nedeniyle, önceki geleneksel makine öğrenimi deneylerinde yalnızca aort kapak (AV) verisine dayalı olarak eğitilen sınıflandırıcıların bazı üfürüm türlerinde daha yüksek performans sergileyebildiği gözlemlenmiştir. Bu bulgu, hibrit ve çok kapaklı yaklaşımlarda ortaya çıkan bazı yanlış sınıflandırmaların fizyolojik ve akustik faktörlerle ilişkili olabileceğini düşündürmektedir. Özellikle aort dışı oskültasyon bölgelerinde daha belirgin hâle gelen üfürümler, sistemik gürültü ve üst üste binen kalp sesleri ile birlikte ortaya çıktığında, üfürümlere özgü spektral imzaların kısmen maskelenmesine neden olabilmektedir.

Gelecek çalışmalarda daha ayrıntılı hata analizi kapsamında, yanlış sınıflandırılan örneklerle ait Mel-spektrogram temsillerinin görsel olarak incelenmesi önerilmektedir. Bunun yanı sıra, hataların kalp döngüsünün belirli fazlarında (S1, sistol, S2 ve diyastol) yoğunlaşıp yoğunlaşmadığının analiz edilmesi, model davranışının fizyolojik temellerinin daha iyi anlaşılmasına katkı sağlayabilir. Ayrıca, modelin karar verme sürecinin daha şeffaf hâle getirilmesi amacıyla duyarlılık (saliency) haritaları veya dikkat (attention) mekanizmalarının kullanılması, kritik zaman–frekans bölgelerinin belirlenmesine olanak tanıyarak klinik yorumlanabilirliği artırabilir (Simonyan vd., 2014).

4.8. Bulguların Tartışılması

4.8.1. Modelin Öğrendikleri

Önerilen hibrit CNN–LSTM mimarisi, pediatrik kalp üfürümü tespiti görevinde hem yerel hem de küresel düzeyde ayırt edici örüntüleri etkili biçimde öğrenebilmiştir. Elde edilen bulgular, modelin iki tamamlayıcı bilgi türünü birlikte temsil edebildiğini açık biçimde ortaya koymaktadır.

İlk olarak, yerel spektral–zamansal örüntüler CNN katmanları aracılığıyla başarıyla yakalanmıştır. Bu örüntüler arasında üfürümlere özgü kısa süreli yüksek enerjili patlamalar, geniş bantlı türbülans gürültüsü ve belirgin harmonik yapılar yer almaktadır. Bu tür akustik imzalar, Mel-spektrogram temsilleri üzerinde yerel evrişimsel filtreler tarafından etkin biçimde öğrenilmekte ve patolojik akım karakteristiklerini doğrudan yansıtmaktadır. Evrişimli sinir ağlarının bu tür yerel örüntüleri yakalamadaki başarısı, görüntü ve ses analizi alanındaki önceki çalışmalarla da tutarlıdır (LeCun vd., 2015).

İkinci olarak, uzun menzilli zamansal bağımlılıklar LSTM katmanları sayesinde modellenmiştir. Model, birden fazla kalp döngüsü boyunca tekrar eden örüntüleri ve üfürümlerin S1 ve S2 kalp seslerine göre zamanlanışını öğrenebilmiş; böylece yalnızca

spektral içerik değil, aynı zamanda kardiyak döngü içerisindeki zamansal bağlam da sınıflandırma sürecine entegre edilmiştir. LSTM ağlarının ardışık verilerde uzun vadeli bağımlılıkları modelleme yeteneği, özellikle biyomedikal sinyallerde kritik öneme sahiptir (Hochreiter ve Schmidhuber, 1997).

Bu bulgular, evrimsel mimarilerin yerel uzamsal ve spektral yapıları yakalamada etkili olduğu, LSTM gibi tekrarlayan ağ bileşenlerinin ise FKG, ECG ve EEG gibi ardışık biyomedikal sinyallerde zamansal bağımlılıkların modellenmesi için uygun olduğu yönündeki önceki çalışmalarla da uyumludur (Yildirim vd., 2018).

Sonuç olarak, CNN–LSTM çerçevesi, hem akustik özelliklerin içeriğini hem de kardiyak döngü içerisindeki zamansal bağlamı birlikte temsil edebilmekte; bu bütüncül temsil kapasitesi, pediatrik kalp üfürümü tespitinde elde edilen yüksek ayırt edici performansın temelini oluşturmaktadır. Bu durum, problem odaklı hibrit mimarilerin, tek başına kullanılan model türlerine kıyasla daha güçlü ve genellenebilir çözümler sunabileceğini göstermektedir.

4.8.2. Hibrit Mimarinin Performansı Neden İyileştirdiği

Geleneksel makine öğrenimi modelleri ve yalnızca CNN tabanlı gelişmiş mimarilerle yapılan karşılaştırmalar, önerilen hibrit CNN–LSTM modelinin üstün performansının, birden fazla tamamlayıcı bileşenin eşzamanlı olarak bir araya getirilmesinden kaynaklandığını açık biçimde ortaya koymaktadır. Bu durum, model başarısının tek bir mimari unsurdan ziyade, bütüncül bir tasarım yaklaşımının sonucu olduğunu göstermektedir.

Bu mimari, ilk olarak Mel-spektrogramlar üzerinde uygulanan iki boyutlu evrişimler aracılığıyla derin spektral öğrenme gerçekleştirmektedir. CNN katmanları, üfürümlere özgü frekans bantlarını, enerji dağılımlarını ve yerel akustik örüntüleri hiyerarşik biçimde soyutlayarak güçlü ve ayırt edici bir özellik temsili üretmektedir (LeCun vd., 2015).

Buna ek olarak LSTM katmanları, akustik olayların zaman içindeki sırasını ve sürelerini modelleyerek kalp seslerinin ardışık doğasını açık biçimde kodlamaktadır. Bu sayede model, üfürümlerin yalnızca varlığını değil, aynı zamanda S1–sistol–S2–diastol yapısı içerisindeki zamansal konumunu ve bağlamını da dikkate alabilmektedir. LSTM tabanlı yapıların uzun menzilli bağımlılıkları yakalama yeteneği, biyomedikal sinyal analizinde önemli bir avantaj sağlamaktadır (Hochreiter ve Schmidhuber, 1997).

Performans artışına katkı sağlayan bir diğer önemli unsur ise iyi tasarlanmış ön işleme ve veri artırma stratejileridir. Bu stratejiler sayesinde sinyal uzunlukları normalize

edilerek model girdileri standartlaştırılmış ve farklı kayıt sürelerinden kaynaklanan değişkenlik etkili biçimde azaltılmıştır. Ayrıca, gürültüye ve kayıt koşullarındaki doğal değişkenliğe karşı model dayanıklılığı artırılmış ve gerçek klinik ortamlarda karşılaşılabilecek akustik bozulmaların etkisi sınırlandırılmıştır. Bunun yanında hasta düzeyinde veri bölme yaklaşımı benimsenerek veri sızıntısı (*data leakage*) önlenmiş ve segment tabanlı veri artırma yoluyla sınıf dengesizliği kontrollü biçimde azaltılmıştır.

Benimsenen bu bütüncül tasarım yaklaşımı, kalp sesi analizinde hibrit CNN–RNN mimarilerinin yalnızca evrimsel ya da yalnızca tekrarlayan modellere kıyasla daha üstün ve dengeli bir performans sunduğunu rapor eden güncel literatür eğilimleriyle de tutarlılık göstermektedir (Reyna vd., 2023; Li vd., 2023; Wang vd., 2024).

4.8.3. Önceki Literatürle İlişki (2020–2025)

Önerilen CNN–LSTM mimarisinin performansı, son yıllarda CirCor ve PhysioNet veri kümeleri kullanılarak geliştirilen kalp üfürümü tespit sistemleriyle karşılaştırmalı olarak değerlendirilmiştir. Bu karşılaştırma, önerilen yaklaşımın literatürdeki güncel yöntemler arasındaki konumunu belirlemek ve sağladığı katkıyı nicel olarak ortaya koymak amacıyla gerçekleştirilmiştir.

Tablo 9, bu çalışmada elde edilen sonuçları 2020–2025 döneminde yayımlanan seçili çalışmalarla karşılaştırmalı biçimde özetlemektedir. Bu tablo, farklı yöntemlerin doğruluk ve F1-skoru gibi temel performans ölçütleri açısından değerlendirilmesine olanak tanımakta ve önerilen hibrit CNN–LSTM modelinin mevcut yaklaşımlar karşısındaki göreceli performansını ortaya koymaktadır.

Tablo 9. CNN–LSTM modelinin önceki kalp üfürümü tespit çalışmalarıyla karşılaştırılması

Çalışma	Yöntem	Veri kümesi	Doğruluk	F1-skoru
Bu çalışma (2025)	CNN–LSTM (hibrit)	CirCor DigiScope	%84,6	0,84
Huang et al. (2024)	2D CNN (multi-task)	CirCor DigiScope	%95,19	–
Reyna et al. (2023)	Various CNN architectures	PhysioNet 2022	%80,2	0,78
Wang et al. (2024)	Branchformer + uncertainty	PhysioNet 2022	%79,8	0,65
Li et al. (2023)	Neural architecture search	PhysioNet 2022	%83,4	0,83
Kumar et al. (2024)	CNN (Mel-spektrogram)	PhysioNet 2022	%81,8	0,80

Önerilen modelin doğruluk değeri %84,6 olup, bu sonuç PhysioNet tabanlı birçok yöntemin raporladığı performanslarla rekabetçi bir düzeydedir. Her ne kadar bu değer, Huang ve arkadaşlarının çok görevli CNN yaklaşımında bildirilen %95,19 doğruluğun altında kalsa da, bu farkın etiket tanımları, görev yapıları (ikili sınıflandırma yerine çok sınıflı problemler) ve eğitim protokollerindeki farklılıklardan kaynaklanmış olması

muhtemeldir. Bu bağlamda, önerilen modelin doğruluk performansı pediatrik tarama senaryoları için yeterli ve uygulanabilir bir düzeyde olup, özellikle duyarlılık ve kesinlik arasındaki dengeli yapısı nedeniyle klinik açıdan anlamlıdır.

F1-skoru açısından değerlendirildiğinde, önerilen hibrit CNN–LSTM modeli 0,84 değerine ulaşarak Li ve arkadaşlarının bildirdiği sonuçlarla (0,83) doğrudan karşılaştırılabilir bir performans sergilemiştir. Bunun yanı sıra, Reyna, Wang ve Kumar gibi araştırmacıların çalışmalarında raporlanan birçok yaklaşımı da geride bırakmaktadır. F1-skorunun yanlış pozitif ve yanlış negatif sınıflandırmalar arasındaki dengeyi yansıtması nedeniyle, özellikle sınıf dengesizliğinin söz konusu olduğu tıbbi veri kümelerinde daha güvenilir bir performans ölçütü olduğu dikkate alındığında, bu bulgu önemli bir avantaj sunmaktadır (Saito ve Rehmsmeier, 2015).

Bununla birlikte, önceki çalışmalar ile doğrudan karşılaştırma yapılırken veri kümesi özellikleri, etiketleme stratejileri ve görev tanımlarındaki farklılıkların dikkatle değerlendirilmesi gerekmektedir. Literatürdeki birçok çalışma üfürüm derecelendirmesi, kapak patolojisi sınıflandırması veya çok sınıflı tanı problemlerine odaklanmıştır. Buna karşılık, bu tez çalışması pediatrik veri üzerinde ikili üfürüm tespitine odaklanmakta ve özellikle zamansal modelleme yaklaşımıyla öne çıkmaktadır. Bu yaklaşım, erken tarama ve klinik triyaj senaryolarında daha doğrudan uygulanabilir sonuçlar sunmaktadır.

Genel olarak elde edilen bulgular, önerilen hibrit CNN–LSTM çerçevesinin pediatrik kalp üfürümü tespiti probleminde güçlü ve dengeli bir performans sunduğunu açık biçimde ortaya koymaktadır. Mel-spektrogram temsilleri, el ile tasarlanmış ses özellikleri, etkili veri artırma stratejileri ve hem geleneksel hem de gelişmiş derin öğrenme modelleriyle gerçekleştirilen kapsamlı karşılaştırmalar birlikte değerlendirildiğinde, bu çalışmanın mevcut literatürü anlamlı biçimde genişlettiği ve problem odaklı hibrit mimarilerin önemini vurguladığı sonucuna ulaşılmaktadır.

4.9. Sonuçlar Bölümünün Özeti

Bu bölümde, pediatrik kalp üfürümlerinin otomatik tespiti için önerilen hibrit CNN–LSTM çerçevesinin kapsamlı bir değerlendirmesi sunulmuştur. İlk olarak, CirCor DigiScope veri kümesinden dikkatle ön işlenmiş bir alt küme oluşturulmuş ve 574 pediatrik hastaya ait toplam 758 adet beş saniyelik FKG segmenti, hasta düzeyinde tabakalı eğitim–doğrulama–test bölünmesi ve hedefe yönelik veri artırma stratejileri kullanılarak analize dâhil edilmiştir.

Önerilen hibrit CNN–LSTM modeli, test kümesinde doğruluk = 0,846, kesinlik = 0,748, duyarlılık = 0,800, F1-skoru = 0,846 ve ROC–AUC = 0,879 değerlerine ulaşarak

güçlü ve dengeli bir ayırt edici performans sergilemiştir. Bu sonuçlar, modelin pediatrik kalp üfürümlerini normal kalp seslerinden güvenilir ve klinik açıdan anlamlı bir biçimde ayırt edebildiğini göstermektedir.

Eğitim sürecinde kullanılan erken durdurma (EarlyStopping) ve öğrenme oranı zamanlaması (ReduceLROnPlateau) mekanizmaları, aşırı öğrenmenin etkili biçimde kontrol altına alınmasını sağlamıştır. Eğitim, doğrulama ve test kümeleri arasındaki F1-skoru farklarının sınırlı kalması, modelin daha önce görülmemiş veriler üzerinde güçlü bir genelleme yeteneğine sahip olduğunu ortaya koymaktadır.

Özellik katkısı analizleri, geleneksel ses özelliklerinin ve özellik seçimi yöntemlerinin klasik makine öğrenimi modellerini belirli ölçüde iyileştirdiğini; ancak Mel-spektrogram tabanlı derin öğrenme yaklaşımlarının daha yüksek temsil gücü sunduğunu göstermiştir. Hibrit CNN–LSTM mimarisi ise spektral ve zamansal bilgiyi birlikte modelleyerek, yalnızca mühendislik ürünü özelliklere veya statik CNN temsillerine dayanan modellere kıyasla daha üstün ve tutarlı sonuçlar elde etmiştir.

Temel modellerle yapılan karşılaştırmalar, önerilen CNN–LSTM çerçevesinin tüm geleneksel makine öğrenimi sınıflandırıcılarını (en iyi model SVM: $F1 = 0,716$) ve DenseNet121 gibi gelişmiş transfer öğrenme mimarilerini ($F1 \approx 0,72$) açık biçimde geride bıraktığını göstermektedir. Ayrıca 2020–2025 dönemindeki güncel literatürle birlikte değerlendirildiğinde, önerilen modelin performansının PhysioNet ve CirCor veri kümeleri üzerinde raporlanan öncü yöntemlerle rekabetçi bir düzeyde olduğu anlaşılmaktadır.

Sonuç olarak, Mel-spektrogramlar üzerinden gerçekleştirilen spektral–zamansal ortak modellemenin, titiz ön işleme, veri artırma ve özellik mühendisliği stratejileriyle birleştirildiğinde, klinik açıdan anlamlı ve uygulanabilir bir otomatik üfürüm tespiti sistemi sunabildiği bu tez kapsamında ortaya konmuştur. Bu bulgular, problem odaklı hibrit derin öğrenme mimarilerinin biyomedikal sinyal analizinde güçlü bir yaklaşım sunduğunu desteklemektedir.

Bir sonraki bölümde, elde edilen sonuçlar temel alınarak klinik çıkarımlar, çalışmanın sınırlılıkları ve gelecekteki araştırma yönleri ele alınacaktır. Özellikle istatistiksel doğrulama yöntemlerinin genişletilmesi, açıklanabilir yapay zekâ (*Explainable AI*) tekniklerinin entegrasyonu ve gerçek zamanlı dijital stetoskop sistemleriyle uygulama düzeyinde entegrasyon, ileriye dönük önemli araştırma alanları olarak öne çıkmaktadır.

5. TARTIŞMA

5.1. Giriş

Bu bölümde, Bölüm 4’te sunulan deneysel bulgular kapsamlı ve eleştirel bir bakış açısıyla yorumlanmakta ve pediatrik kalp üfürümlerinin otomatik tespiti için önerilen hibrit CNN–LSTM çerçevesinin bilimsel ve klinik önemi tartışılmaktadır. Sonuçlar bölümünde ağırlıklı olarak nicel performans ölçütlerine odaklanılmışken, bu bölümde elde edilen bulguların araştırma amaçları ve hipotezleri bağlamındaki anlamı daha derinlemesine ele alınmaktadır.

Bu doğrultuda, hibrit mimarinin diğer model yaklaşımlarına kıyasla neden daha yüksek performans sergilediği, elde edilen sonuçların 2020–2025 döneminde yayımlanan güncel çalışmalarla nasıl örtüştüğü veya farklılaştığı sistematik biçimde analiz edilmektedir. Ayrıca çalışmanın güçlü yönleri ve sınırlılıkları değerlendirilmekte; modelin klinik uygulamalara entegrasyonu açısından taşıdığı potansiyel ve karşılaşılabilecek zorluklar tartışılmaktadır.

Buna ek olarak, bu bölüm deneysel bulguları biyomedikal sinyal işleme ve derin öğrenme literatüründeki kuramsal çerçeve ile bütünleştirerek yorumlamayı amaçlamakta ve elde edilen sonuçların yalnızca performans ölçütleri üzerinden değil, aynı zamanda fizyolojik ve klinik bağlam içinde anlamlandırılmasına olanak tanımaktadır.

5.2. Temel Bulguların Yorumlanması

Önerilen hibrit CNN–LSTM modeli, test kümesinde %84,6 doğruluk, 0,748 kesinlik (*precision*), 0,800 duyarlılık (*recall*), 0,846 F1-skoru ve 0,879 ROC–AUC değerleri elde etmiştir. Bu performans, hem tüm geleneksel makine öğrenimi modellerini —en iyi sonucu veren SVM’in test F1-skoru olan 0,716 dâhil olmak üzere— hem de çalışmada değerlendirilen gelişmiş derin öğrenme mimarilerini —örneğin DenseNet121’in yaklaşık 0,72 F1-skoru— açık ve tutarlı biçimde geride bırakmaktadır. Bu sonuçlar, FKG sinyallerinin başarılı biçimde modellenebilmesi için yalnızca spektral özelliklerin değil, aynı zamanda zamansal dinamiklerin de birlikte ele alınması gerektiğini göstermektedir. Önerilen hibrit CNN–LSTM mimarisi, Mel-spektrogramlar üzerinden öğrenilen uzamsal–spektral örüntüler ile kalp döngüsü içerisindeki zamansal bağımlılıkları tek bir model çerçevesinde bütünleştirerek bu iki bilgi boyutunu etkin ve tamamlayıcı biçimde temsil edebilmektedir.

Elde edilen performans farkı, yalnızca nicel bir iyileşme değil, aynı zamanda modelin daha karmaşık ve fizyolojik açıdan anlamlı örüntüleri öğrenebilme kapasitesine sahip olduğunu göstermektedir. Bu durum, pediatrik kalp üfürümü tespiti probleminde hibrit derin öğrenme mimarilerinin, yalnızca spektral özellik çıkarımına dayalı CNN modellerine veya geleneksel özellik tabanlı yöntemlere kıyasla daha güçlü ve genellenebilir bir temsil kapasitesi sunduğunu açık biçimde ortaya koymaktadır.

5.2.1. Geri Çağırma ve Duyarlılığın Klinik Önemi

Modelin %80 duyarlılık (recall) değeri, mevcut üfürüm vakalarının büyük bir kısmının sistem tarafından doğru biçimde tespit edilebildiğini göstermektedir. Klinik açıdan değerlendirildiğinde, bir üfürümün gözden kaçırılması (false negative), yanlış bir alarm verilmesine (false positive) kıyasla genellikle daha ciddi sonuçlar doğurabileceğinden, pediatrik tarama senaryolarında yüksek duyarlılık kritik bir performans ölçütü olarak kabul edilmektedir.

Bu bağlamda elde edilen duyarlılık değeri, modelin gerçek üfürüm vakalarının önemli bir bölümünü güvenilir biçimde yakalayabildiğini ve erken tarama ile uygun klinik yönlendirme süreçlerini destekleyebilecek bir performans düzeyi sunduğunu göstermektedir. Özellikle pediatrik popülasyonda erken tanının, ileri kardiyak komplikasyonların önlenmesi açısından büyük önem taşıdığı göz önüne alındığında, modelin bu yöndeki başarımı dikkat çekicidir.

Dolayısıyla modelin duyarlılık açısından sergilediği bu performans, pediatrik kalp üfürümü taraması gibi erken tespit odaklı klinik uygulamalar için anlamlı, güvenilir ve uygulanabilir bir seviyeyi temsil etmektedir. Bu durum, modelin klinik karar destek sistemlerine entegrasyonu açısından önemli bir potansiyel taşıdığını göstermektedir.

5.2.2. Kritik Bir Performans Ölçütü Olarak F1-Skoru

Elde edilen 0,84 F1-skoru, önerilen hibrit modelin gerçek üfürüm vakalarını doğru biçimde tespit etme ile gereksiz yanlış pozitifleri sınırlandırma arasında dengeli ve istikrarlı bir performans sunduğunu göstermektedir. F1-skorunun, kesinlik (*precision*) ve duyarlılığı (*recall*) birlikte dikkate alan birleşik bir ölçüt olması, model performansının tek bir metrik üzerinden değil, iki kritik performans göstergesinin dengesi üzerinden değerlendirildiğini ortaya koymaktadır (Saito ve Rehmsmeier, 2015).

Bu dengeli yapı, özellikle sınıf dengesizliğinin bulunduğu tıbbi veri kümelerinde model davranışının daha gerçekçi ve güvenilir biçimde yorumlanmasına olanak tanımaktadır. Yalnızca yüksek doğruluk veya duyarlılık değerleri, modelin klinik açıdan

güvenilir olduğunu tek başına garanti etmezken; F1-skoru bu iki ölçüt arasındaki dengeyi yansıtarak daha bütüncül bir performans değerlendirmesi sunmaktadır.

Bu denge, özellikle dijital stetoskop sistemleri, uzaktan sağlık hizmetleri (*telemedicine*) ve klinik karar destek sistemleri gibi erken ve güvenilir tespit gerektiren uygulamalar açısından kritik öneme sahiptir. Bu tür sistemlerde hem yanlış alarm oranının düşük tutulması hem de gerçek patolojik vakaların mümkün olduğunca kaçırılmaması gerekmektedir. Dolayısıyla dengeli bir *precision-recall* performansı, klinik kullanım senaryolarında temel bir gereklilik olarak değerlendirilmektedir.

Sonuç olarak, elde edilen yüksek ve dengeli F1-skoru, önerilen hibrit CNN–LSTM modelinin yalnızca teknik açıdan değil, aynı zamanda klinik uygulamalar açısından da güvenilir ve uygulanabilir bir çözüm sunduğunu desteklemektedir.

5.3. Hibrit CNN–LSTM Mimarisi Neden Daha İyi Performans Gösterdi

5.3.1. CNN ile Spektral Öğrenmenin Rolü

Mel-spektrogramlar, ham FKG sinyallerini insan işitsel algısıyla uyumlu, yapılandırılmış ve yorumlanabilir iki boyutlu zaman–frekans temsillerine dönüştürmektedir. Bu temsil biçimi, sinyaldeki enerji dağılımlarının zaman içerisindeki değişimini açık biçimde ortaya koyarak kalp seslerine ait ince ve karmaşık akustik örüntülerin analiz edilmesine olanak tanımaktadır.

Evrişimsel sinir ağları (CNN), bu tür temsiller üzerinde frekans enerjisi dağılımlarını, harmonik yapıları, üfürümlere özgü türbülans imzalarını ve ses şiddetindeki ani geçişleri öğrenmede oldukça etkilidir. CNN katmanları, Mel-spektrogramlar üzerindeki yerel örüntüleri hiyerarşik biçimde işleyerek, kalp üfürümlerinin akustik karakterini belirleyen temel spektral özellikleri otomatik olarak yakalayabilmektedir.

Bu bağlamda Mel-spektrogram temsilleri, klinisyenlerin geleneksel oskültasyon sırasında deneyime dayalı olarak değerlendirdiği akustik ipuçlarının —örneğin türbülanslı akım sesleri, geniş bant enerji artışları ve belirli frekans bantlarında yoğunlaşan harmonik yapılar— makine öğrenimi sistemleri tarafından nesnel, tekrarlanabilir ve ölçeklenebilir biçimde analiz edilmesini mümkün kılmaktadır (LeCun vd., 2015).

5.3.2. LSTM İle Zamansal Modelleme

Kalp sesleri, doğası gereği ardışık ve periyodik fizyolojik sinyaller olup S1, sistol, S2 ve diyastol fazlarının zamanlaması üfürüm tanısında belirleyici bir rol oynamaktadır. Bu nedenle kalp sesi analizinde yalnızca spektral özelliklerin değil, aynı zamanda kardiyak döngü içerisindeki zamansal ilişkilerin de doğru ve bütüncül biçimde modellenmesi kritik önem taşımaktadır.

Uzun-kısa süreli bellek (LSTM) katmanları, akustik olayların sürelerini, kardiyak döngüler boyunca tekrar eden örüntüleri ve spektral değişimler arasındaki uzun menzilli zamansal bağımlılıkları modelleyebilme yeteneğine sahiptir. Bu sayede model, kalp seslerinin yalnızca anlık spektral yapısını değil, aynı zamanda S1–sistol–S2–diyastol dizilimi içerisindeki bağlamsal zamanlamasını da öğrenebilmektedir. Bu tür zamansal modelleme, özellikle üfürümlerin kalp döngüsü içerisindeki konumu, süresi ve sürekliliği gibi klinik açıdan kritik bilgilerin yakalanmasını mümkün kılmaktadır (Hochreiter ve Schmidhuber, 1997).

Buna karşılık, yalnızca CNN tabanlı yaklaşımlar, güçlü spektral özellik çıkarım kapasitesine sahip olmalarına rağmen, bu tür ardışık ve uzun menzilli zamansal ilişkileri sınırlı ölçüde temsil edebilmektedir. LSTM katmanlarının modele entegre edilmesi, bu eksikliği gidererek zamansal bağımlılıkların daha etkin biçimde yakalanmasını sağlamakta ve kalp üfürümlerine ait karakteristik örüntülerin daha doğru ve bağlamsal biçimde modellenmesine katkıda bulunmaktadır.

5.3.3. CNN ve LSTM Bütünleşmesinin Etkisi

Hibrit mimari, CNN’lerin derin spektral özellik çıkarma kapasitesi ile LSTM’lerin zamansal modelleme yeteneğini bütünleşik bir çerçevede bir araya getirmektedir. CNN katmanları, Mel-spektrogramlar üzerinde yer alan frekans bantları, enerji dağılımları ve üfürümlere özgü spektral örüntüleri hiyerarşik biçimde öğrenirken; LSTM katmanları bu akustik örüntülerin kardiyak döngüler boyunca nasıl geliştiğini, tekrar ettiğini ve zamansal olarak nasıl organize olduğunu modelleyebilmektedir.

Bu sinerjik yapı sayesinde model, yalnızca statik spektral içerik üzerinden karar vermekle kalmayıp, aynı zamanda akustik olayların zaman içerisindeki düzenini, sürekliliğini ve bağlamsal ilişkilerini de dikkate almaktadır. Böylece kalp üfürümlerine ait hem yerel (lokal) hem de küresel (global) örüntüler tek bir model içerisinde bütüncül biçimde temsil edilebilmektedir.

Nitekim yalnızca CNN tabanlı modellerde gözlenen yaklaşık 0,72 düzeyindeki F1-skorunun, hibrit CNN–LSTM tasarımında 0,84’e yükselmesi, bu bütünleşik yaklaşımın

sağladığı temsil gücünü nicel olarak ortaya koymaktadır. Bu fark, yalnızca performans artışı değil, aynı zamanda modelin daha karmaşık ve fizyolojik açıdan anlamlı örüntüleri öğrenebilme kapasitesinin bir göstergesi olarak değerlendirilebilir.

Bu bulgu, kalp sesi analizi gibi hem spektral hem de zamansal bilginin kritik olduğu biyomedikal sinyal işleme problemlerinde, hibrit derin öğrenme mimarilerinin daha güçlü, daha esnek ve daha genellenebilir bir temsil kapasitesi sunduğunu açık biçimde ortaya koymaktadır (Hochreiter ve Schmidhuber, 1997; LeCun vd., 2015).

5.4. Önceki Çalışmalarla Karşılaştırma (2020–2025)

Bölüm 4’te sunulan karşılaştırmalı sonuçlar, önerilen modelin son yıllarda yayımlanan birçok ileri düzey yöntemle karşılaştırıldığında benzer veya daha üstün bir performans sergilediğini ortaya koymaktadır. Elde edilen bulgular, hibrit CNN–LSTM yaklaşımının pediatrik kalp üfürümü tespiti probleminde yalnızca geleneksel makine öğrenimi modellerine değil, aynı zamanda bazı modern derin öğrenme mimarilerine kıyasla da daha güçlü ve tutarlı bir genelleme yeteneği sunduğunu göstermektedir.

Bu durum, spektral ve zamansal bilginin birlikte modellenmesinin, kalp sesi analizi gibi karmaşık, gürültüye duyarlı ve yüksek değişkenlik içeren biyomedikal sinyallerin incelenmesinde kritik bir avantaj sağladığını desteklemektedir. Özellikle Mel-spektrogram tabanlı uzamsal temsiller ile zamansal bağımlılıkların entegre edilmesi, modelin daha zengin ve bağlama duyarlı özellikler öğrenmesine olanak tanımaktadır.

Dolayısıyla elde edilen sonuçlar, önerilen hibrit yaklaşımın yalnızca deneysel performans açısından değil, aynı zamanda metodolojik katkı bakımından da literatürdeki güncel çalışmalarla rekabetçi ve anlamlı bir konumda bulunduğunu göstermektedir. Bu bulgu, hibrit derin öğrenme mimarilerinin biyomedikal sinyal işleme alanında giderek artan önemini destekleyen literatür eğilimleriyle de uyumludur (Reyna vd., 2023; Li vd., 2023; Wang vd., 2024).

5.4.1. CirCor ve Physionet Literatürü ile Uyum

Reyna ve arkadaşları, Li ve arkadaşları, Wang ve arkadaşları ile Kumar ve arkadaşları tarafından gerçekleştirilen çalışmalar, kalp üfürümü sınıflandırmasında derin öğrenme yaklaşımlarının önemini açık biçimde ortaya koymuştur. Bununla birlikte, bu çalışmaların önemli bir kısmı ağırlıklı olarak CNN tabanlı mimarilere odaklanmış; zamansal bağımlılıkların modellenmesi ya sınırlı düzeyde ele alınmış ya da açık biçimde entegre edilmemiştir. Ayrıca dört farklı oskültasyon bölgesini birlikte ele alan ve

fizyolojik olarak anlamlı veri artırma stratejilerini sistematik biçimde değerlendiren çalışmaların sayısı görece sınırlıdır.

Buna ek olarak literatürdeki bazı çalışmalar, çok sınıflı üfürüm derecelendirmesi veya kapak patolojisi sınıflandırması gibi daha karmaşık tanı görevlerine odaklanmaktadır. Buna karşılık, bu tez çalışması özellikle pediatrik bağlamda ikili üfürüm tespitine odaklanmakta ve erken tarama ile klinik yönlendirme süreçlerine doğrudan katkı sağlayabilecek daha pratik ve uygulanabilir bir çözüm sunmaktadır. Bu yönüyle çalışma, klinik kullanım senaryolarına daha yakın bir problem tanımı benimsemektedir.

Elde edilen 0,84 F1-skoru, pediatrik FKG kayıtlarının gürültülü, değişken ve karmaşık yapısı dikkate alındığında dikkat çekici bir performans düzeyine işaret etmektedir. Bu sonuç, önerilen hibrit CNN–LSTM yaklaşımının yalnızca literatürde raporlanan birçok yöntemle karşılaştırılabilir bir performans sunmakla kalmayıp, aynı zamanda bazı yaklaşımlar karşısında daha dengeli ve genellenebilir bir performans sergilediğini göstermektedir.

Dolayısıyla bu bulgular, spektral ve zamansal modellemeyi bütünleştiren hibrit mimarilerin, pediatrik kalp üfürümü tespiti gibi zorlu biyomedikal sinyal işleme problemlerinde metodolojik olarak güçlü ve klinik açıdan anlamlı bir alternatif sunduğunu desteklemektedir (Reyna vd., 2023; Li vd., 2023; Wang vd., 2024; Kumar vd., 2024).

5.4.2. Çok Görevli Yüksek Doğruluklu CNN Modelleriyle Karşılaştırma

Reyna ve arkadaşları, Li ve arkadaşları, Wang ve arkadaşları ile Kumar ve arkadaşları tarafından gerçekleştirilen çalışmalar, kalp üfürümü sınıflandırmasında derin öğrenme yaklaşımlarının önemini açık biçimde ortaya koymuştur. Bununla birlikte, bu çalışmaların önemli bir kısmı ağırlıklı olarak CNN tabanlı mimarilere odaklanmış; zamansal bağımlılıkların modellenmesi ya sınırlı düzeyde ele alınmış ya da açık biçimde entegre edilmemiştir. Ayrıca dört farklı oskültasyon bölgesini birlikte ele alan ve fizyolojik olarak anlamlı veri artırma stratejilerini sistematik biçimde değerlendiren çalışmaların sayısı görece sınırlıdır.

Buna ek olarak literatürdeki bazı çalışmalar, çok sınıflı üfürüm derecelendirmesi veya kapak patolojisi sınıflandırması gibi daha karmaşık tanı görevlerine odaklanmaktadır. Buna karşılık, bu tez çalışması özellikle pediatrik bağlamda ikili üfürüm tespitine odaklanmakta ve erken tarama ile klinik yönlendirme süreçlerine doğrudan katkı sağlayabilecek daha pratik ve uygulanabilir bir çözüm sunmaktadır. Bu

yönüyle çalışma, klinik kullanım senaryolarına daha yakın bir problem tanımı benimsemektedir.

Elde edilen 0,84 F1-skoru, pediatrik FKG kayıtlarının gürültülü, değişken ve karmaşık yapısı dikkate alındığında dikkat çekici bir performans düzeyine işaret etmektedir. Bu sonuç, önerilen hibrit CNN–LSTM yaklaşımının yalnızca literatürde raporlanan birçok yöntemle karşılaştırılabilir bir performans sunmakla kalmayıp, aynı zamanda bazı yaklaşımlar karşısında daha dengeli ve genellenebilir bir performans sergilediğini göstermektedir.

Dolayısıyla bu bulgular, spektral ve zamansal modellemeyi bütünleştiren hibrit mimarilerin, pediatrik kalp üfürümü tespiti gibi zorlu biyomedikal sinyal işleme problemlerinde metodolojik olarak güçlü ve klinik açıdan anlamlı bir alternatif sunduğunu desteklemektedir (Reyna vd., 2023; Li vd., 2023; Wang vd., 2024; Kumar vd., 2024).

5.5. Önerilen Yöntemin Güçlü Yönleri

Bu çalışmanın en önemli güçlü yönlerinden biri, baştan sona tutarlı, sistematik ve sağlam bir ön işleme hattına dayanmasıdır. Tüm FKG kayıtlarının beş saniyelik sabit uzunlukta segmentlere standardize edilmesi, farklı sürelerdeki kayıtlar arasında karşılaştırılabilirliği artırmış ve model eğitim sürecinin daha kararlı ve verimli hâle gelmesine katkı sağlamıştır. Uzun süreli üfürüm kayıtlarının bilgi kaybına yol açmadan uygun biçimde bölünmesi ve kısa süreli kayıtların sıfır dolgu (zero-padding) ve normalizasyon işlemleri ile dengelenmesi, sinyal bütünlüğünün korunmasına yardımcı olmuştur.

Ayrıca hasta düzeyinde gerçekleştirilen veri bölme stratejisi, veri sızıntısının (data leakage) etkin biçimde önlenmesini sağlamış ve elde edilen sonuçların bilimsel geçerliliğini ve yeniden üretilebilirliğini güçlendirmiştir. Bu yaklaşım, modelin gerçek dünya senaryolarında daha güvenilir bir genelleme performansı sunmasına önemli katkı sağlamaktadır.

Önerilen yöntemin bir diğer önemli avantajı, veri artırma stratejilerinin bilinçli ve hedefe yönelik biçimde uygulanmasıdır. Başlangıçta veri kümesinde yaklaşık %19 oranında temsil edilen üfürüm-pozitif örnekler, perde kaydırma (pitch shifting), zaman kaydırma (time shifting) ve kontrollü gürültü ekleme gibi yöntemlerle artırılmıştır. Bu yaklaşım, sınıf dengesizliğinin olumsuz etkilerini azaltmış, modelin genelleme yeteneğini güçlendirmiş ve özellikle derin öğrenme modellerinde sıkça karşılaşılan aşırı öğrenme (overfitting) probleminin kontrol altına alınmasına katkı sağlamıştır.

Buna ek olarak çalışma, kapsamlı ve çok boyutlu bir özellik mühendisliği sürecine dayanmaktadır. MFCC'ler, sıfır geçiş oranı (ZCR), spektral merkez, bant enerji oranları ve çeşitli istatistiksel tanımlayıcılar gibi toplam 128 el ile tasarlanmış ses özelliğinin çıkarılması, güçlü geleneksel temel modellerin oluşturulmasına olanak tanımış ve model davranışının yorumlanabilirliğini artırmıştır. Her ne kadar derin öğrenme tabanlı yaklaşımlar daha yüksek performans elde etmiş olsa da, bu geleneksel özellikler özellikle model kararlarının anlaşılması ve gömülü veya düşük hesaplama kapasitesine sahip sistemler için daha hafif ve uygulanabilir çözümler geliştirilmesi açısından önemini korumaktadır.

Genel olarak değerlendirildiğinde, sağlam ön işleme hattı, etkili veri artırma stratejileri ve hibrit özellik yaklaşımının birlikte kullanılması, bu çalışmanın metodolojik açıdan güçlü ve dengeli bir yapı sunmasını sağlamaktadır.

5.6. Çalışmanın Sınırlılıkları

Elde edilen güçlü sonuçlara rağmen, bu çalışmanın bazı sınırlılıkları bulunduğu kabul edilmelidir. Öncelikle, uygulanan veri artırma stratejileri sınıf dengesizliğini önemli ölçüde iyileştirmiş olsa da, gerçek pediatrik popülasyonlarda üfürüm prevalansı genellikle daha düşüktür. Bu durum, modelin farklı prevalans oranlarına sahip klinik ortamlardaki genelleme performansını etkileyebilecek potansiyel bir sınırlılık olarak değerlendirilebilir.

Bir diğer sınırlılık, değerlendirme sürecinin tek bir tabakalı eğitim-doğrulama-test bölünmesine dayanmasıdır. Bu yaklaşım literatürde yaygın olarak kullanılmakla birlikte, istatistiksel güvenilirliğin daha da güçlendirilmesi için gelecekte K-katlı çapraz doğrulama, tekrarlı rastgele alt örnekleme ve modeller arası performans farklarının değerlendirilmesine yönelik McNemar veya Wilcoxon gibi istatistiksel testlerin uygulanması önerilmektedir.

Ayrıca CirCor DigiScope veri kümesinde yer alan kayıtlar görece kontrollü klinik koşullarda elde edilmiştir. Ancak gerçek klinik ortamlarda stetoskop cihazlarının kalitesindeki değişkenlik, artan arka plan gürültüsü ve pediatrik hastaların hareketi veya ağlaması gibi faktörler sinyal kalitesini olumsuz yönde etkileyebilmektedir. Bu nedenle, farklı cihazlardan ve daha zorlu akustik koşullardan elde edilen veriler üzerinde model performansında belirli değişiklikler gözlemlenmesi olasıdır. Bu durum, modelin gerçek dünya uygulamalarına aktarılmadan önce çok merkezli ve heterojen veri kümeleri üzerinde doğrulanmasını gerekli kılmaktadır.

Son olarak, hibrit CNN–LSTM mimarisi yüksek ayırt edici performans sunmasına rağmen açıklanabilirlik (interpretability) düzeyi sınırlıdır. Klinik uygulamalarda, hekim güveninin sağlanması ve düzenleyici onay süreçlerinin desteklenmesi açısından model kararlarının şeffaf ve yorumlanabilir olması kritik öneme sahiptir. Bu bağlamda, duyarlılık haritaları (saliency maps), zamansal dikkat mekanizmaları (temporal attention) ve Mel-spektrogramlar üzerinde üfürüm bölgelerinin görsel olarak vurgulanması gibi açıklanabilir yapay zekâ (Explainable AI) yöntemlerinin entegrasyonu, gelecekteki çalışmalar için önemli bir araştırma yönü olarak öne çıkmaktadır.

Genel olarak değerlendirildiğinde, bu sınırlılıklar önerilen yaklaşımın geçerliliğini zayıflatmaktan ziyade, modelin daha geniş klinik kullanım senaryolarına uyarlanabilmesi için gelecekte ele alınması gereken geliştirme alanlarını işaret etmektedir.

5.7. Klinik Uygulamalar İçin Çıkarımlar

Önerilen hibrit CNN–LSTM modelinin ortaya koyduğu yüksek performans, çeşitli klinik uygulama senaryoları açısından önemli bir translasyonel potansiyel sunmaktadır. Modelin mobil sağlık (mHealth) uygulamaları, akıllı stetoskop sistemleri ve teletıp platformları ile entegre edilmesi, pediatristlerin ve pratisyen hekimlerin kalp üfürümü taramasını daha hızlı, nesnel ve tutarlı bir şekilde gerçekleştirmelerine olanak sağlayabilir. Nitekim son yıllarda yapay zekâ destekli oskültasyon sistemlerinin, birinci basamak sağlık hizmetlerinde tanısal doğruluğu artırmada önemli katkılar sunduğu gösterilmiştir (Thompson vd., 2019; Clifford vd., 2022).

Güvenilir bir otomatik tarama aracının kullanımı, özellikle erken tanı süreçlerinde kritik bir rol oynamaktadır. Bu tür sistemler, konjenital kalp hastalıklarının (CHD) erken evrede tespit edilmesine katkı sağlayarak kaçırılan vakaların azaltılmasına, uzman olmayan klinisyenler için karar destek mekanizmalarının güçlendirilmesine ve pediatrik kardiyolojiye yapılacak yönlendirmelerin daha etkin biçimde önceliklendirilmesine imkân tanıyabilir. Ayrıca, sağlık hizmetlerine erişimin sınırlı olduğu bölgelerde bu tür teknolojilerin kullanımı, tarama kapasitesini artırarak sağlık hizmetlerinin eşitliğine ve bakım kalitesinin iyileştirilmesine katkı sunmaktadır (World Health Organization, 2021).

Modelin yalnızca beş saniyelik kısa FKG segmentleri ile çalışabilmesi, onu uzaktan izleme ve sürekli takip senaryoları için de son derece uygun hâle getirmektedir. Ev tabanlı izleme sistemleri, mobil uygulamalar ve düşük güç tüketimli Nesnelerin İnterneti (Internet of Things – IoT) cihazları aracılığıyla elde edilen periyodik veya sürekli kayıtların otomatik analizi, özellikle uzun dönem izlem gerektiren pediatrik hastalarda erken uyarı sistemlerinin geliştirilmesine olanak sağlayabilir. Bu yaklaşım, sürekli

fizyolojik veri akışının yapay zekâ ile analiz edilmesine dayalı dijital sağlık çözümlerinin klinik pratikteki önemini daha da artırmaktadır (Clifford vd., 2016).

5.8. Gelecek Çalışmalar İçin Öneriler

Bu çalışmadan elde edilen bulgular ve belirlenen sınırlılıklar doğrultusunda, gelecekteki araştırmalar için çeşitli geliştirme yönleri öne çıkmaktadır. Öncelikle, modelin klinik güvenilirliğini ve benimsenebilirliğini artırmak amacıyla açıklanabilir yapay zekâ (Explainable AI – XAI) yaklaşımlarının entegrasyonu kritik bir gereklilik olarak değerlendirilmektedir. Grad-CAM gibi görselleştirme teknikleri ile Mel-spektrogramlar üzerinde üfürümle ilişkili zaman–frekans bölgelerinin vurgulanması, Yerel Yorumlanabilir Model-Agnostik Açıklamalar (LIME) ve Shapley katkı değerleri (SHAP) aracılığıyla model kararlarının ayrıştırılması ve dikkat mekanizmaları ile modelin odaklandığı zaman aralıklarının ortaya konulması, klinisyenlerin model çıktısını daha şeffaf, yorumlanabilir ve güvenilir bir şekilde değerlendirmesine olanak sağlayabilir (Selvaraju vd., 2017; Ribeiro vd., 2016; Lundberg ve Lee, 2017).

Buna ek olarak, tanısal doğruluğun artırılması amacıyla çok modlu öğrenme (multimodal learning) yaklaşımlarının araştırılması önemli bir gelişim alanı olarak öne çıkmaktadır. FKG sinyallerinin EKG verileri, hasta öyküsü ve demografik bilgiler ile bütünleştirilmesi, kardiyak patolojilerin daha bağlama duyarlı ve kapsamlı bir şekilde değerlendirilmesine imkân tanıyabilir. Literatürde, çok modlu biyomedikal veri entegrasyonunun klinik karar destek sistemlerinde performansı artırdığı ve daha güvenilir klinik yorumlara olanak sağladığı gösterilmiştir (Baltrusaitis vd., 2019).

Son olarak, yalnızca ikili üfürüm tespiti ile sınırlı kalınmayıp, çok sınıflı sınıflandırma yaklaşımlarına yönelmek önemli bir araştırma perspektifi sunmaktadır. Sistolik ve diyastolik üfürümlerin ayrımı, masum ve patolojik üfürümlerin sınıflandırılması gibi daha ayrıntılı tanısal görevleri hedefleyen modeller, klinik karar destek sistemlerinin gerçek dünya kullanım değerini ve klinik faydasını önemli ölçüde artırabilir. Bu tür gelişmiş sınıflandırma yaklaşımları, özellikle pediatrik kardiyoloji alanında daha hassas ve klinik açıdan anlamlı karar süreçlerinin desteklenmesine katkı sağlayacaktır.

5.8.1. Gerçek Hastane Verileri İle Doğrulama

Önerilen hibrit CNN–LSTM modelinin sağlamlığının ve genellenebilirliğinin güvenilir biçimde değerlendirilebilmesi için, modelin gerçek klinik ortamlarda toplanmış harici veri kümeleri üzerinde test edilmesi kritik bir gereklilik olarak öne çıkmaktadır.

Farklı hastanelerden elde edilen FKG kayıtları, çeşitli dijital stetoskop donanımları ve değişken gürültü seviyeleri altında gerçekleştirilecek harici doğrulama çalışmaları, modelin gerçek dünya koşullarındaki performansını daha nesnel ve kapsamlı biçimde ortaya koyacaktır. Bu tür çok merkezli ve heterojen veri setleri üzerinde yapılan değerlendirmeler, modelin veri dağılımı değişimlerine karşı dayanıklılığını test etmek açısından büyük önem taşımaktadır (Kelly vd., 2019).

Yapay zekâ tabanlı sistemlerin klinik uygulamalara aktarılmasından önce harici doğrulama süreçlerinin gerçekleştirilmesi, literatürde temel bir metodolojik standart olarak kabul edilmektedir. Harici doğrulama, modelin yalnızca eğitildiği veri kümesine özgü örüntüleri öğrenmediğini, farklı klinik ortamlarda da tutarlı ve güvenilir sonuçlar üretebildiğini göstermesi açısından kritik bir rol oynamaktadır (Liu vd., 2019). Bu bağlamda, modelin farklı cihazlar, hasta popülasyonları ve kayıt koşulları altında test edilmesi, klinik güvenilirliğin sağlanması açısından vazgeçilmezdir.

Bu doğrultuda yürütülecek ileri düzey çalışmalar, hibrit CNN–LSTM çerçevesinin genellenebilirliğini, klinik geçerliliğini ve uygulama potansiyelini daha da güçlendirecektir. Özellikle çok merkezli prospektif çalışmalar ile desteklenen doğrulama süreçleri, pediatrik kalp üfürümü taramasında yapay zekâ destekli karar destek sistemlerinin güvenli, etkili ve regülasyonlara uygun biçimde klinik pratiğe entegrasyonuna önemli katkılar sağlayacaktır.

5.9. Özet

Önerilen hibrit CNN–LSTM modelinin sağlamlığının ve genellenebilirliğinin güvenilir biçimde doğrulanabilmesi için, modelin gerçek klinik ortamlarda toplanmış harici veri kümeleri üzerinde değerlendirilmesi kritik bir gereklilik olarak öne çıkmaktadır. Farklı hastanelerden elde edilen FKG kayıtları, çeşitli dijital stetoskop donanımları ve değişken gürültü seviyeleri altında gerçekleştirilecek harici doğrulama çalışmaları, modelin gerçek dünya koşullarındaki performansını daha kapsamlı, nesnel ve güvenilir biçimde ortaya koyacaktır. Bu tür heterojen veri kaynakları üzerinde yapılan değerlendirmeler, modelin veri dağılımı değişimlerine karşı dayanıklılığını test etmek açısından büyük önem taşımaktadır (Kelly vd., 2019).

Yapay zekâ tabanlı sistemlerin klinik uygulamalara aktarılmasından önce harici doğrulama süreçlerinin gerçekleştirilmesi, literatürde temel bir metodolojik standart olarak kabul edilmektedir. Harici doğrulama çalışmaları, modelin yalnızca belirli bir veri kümesine aşırı uyum sağlayan (overfitting) bir çözüm olmadığını, farklı klinik ortamlarda da kararlı ve güvenilir sonuçlar üretebildiğini göstermesi açısından kritik öneme sahiptir

(Liu vd., 2019). Bu bağlamda, modelin farklı cihazlar, hasta popülasyonları ve kayıt koşulları altında sistematik olarak test edilmesi, klinik güvenilirliğin sağlanması açısından vazgeçilmezdir.

İleri çalışmalar, hibrit CNN–LSTM modelinin açıklanabilirlik, genellenebilirlik ve klinik etkisini artıracaktır. Özellikle çok merkezli doğrulamalar, yapay zekâ destekli sistemlerin klinikte güvenli ve etkili kullanımını destekleyerek yöntemin uygulanabilir ve güvenilir olduğunu gösterecektir.



6. SONUÇ ve GELECEK ÇALIŞMALAR

6.1. Giriş

Bu bölüm, Mel-spektrogram temsilleri ve geleneksel akustik özelliklerin birlikte kullanıldığı pediatrik kalp üfürümlerinin otomatik tespiti için önerilen hibrit CNN–LSTM çerçevesinin temel amaçlarını, metodolojik yaklaşımını ve elde edilen bulguları bütüncül bir perspektifle özetleyerek tezi sonlandırmaktadır. Bu kapsamda, çalışmanın başlıca bilimsel katkıları sistematik biçimde ortaya konulmakta; elde edilen sonuçların klinik uygulamalar açısından anlamı tartışılmakta; mevcut sınırlılıklar eleştirel bir bakış açısıyla değerlendirilmektedir. Ayrıca, literatürdeki güncel gelişmeler doğrultusunda gelecekteki araştırmalar için potansiyel yönler de önerilmektedir.

Bu bölümün temel amacı, gerçekleştirilen araştırmanın özellikle pediatrik popülasyonlarda yapay zekâ destekli kardiyak oskültasyon alanına sağladığı katkıyı kapsamlı ve analitik bir çerçevede ortaya koymaktır. Bu doğrultuda, çalışmanın hem metodolojik yenilikleri hem de klinik uygulanabilirliği birlikte ele alınarak, elde edilen bulguların gelecekteki araştırmalar ve gerçek dünya klinik uygulamaları üzerindeki olası etkileri bütüncül ve sistematik bir şekilde değerlendirilmiştir.

6.2. Araştırmanın Özeti

Bu tezin temel amacı, pediatrik FKG kayıtlarından kalp üfürümlerini yüksek doğruluk ve güvenilirlik ile tespit edebilen otomatik bir sistem geliştirmektir. Çalışma, PhysioNet kapsamında sunulan CirCor DigiScope veri kümesi üzerine inşa edilmiş olup; veri ön işleme, özellik çıkarımı, model eğitimi, performans değerlendirmesi ve karşılaştırmalı analiz aşamalarını kapsayan uçtan uca bir metodolojik çerçeve sunmaktadır (Clifford vd., 2022).

Bu kapsamda, FKG kayıtları beş saniyelik sabit uzunlukta segmentlere standartlaştırılmış, sınıf dengesizliğinin etkilerini azaltmak amacıyla çeşitli veri artırma teknikleri uygulanmış ve kalp seslerindeki zaman–frekans örüntülerini daha etkin temsil edebilmek için Mel-spektrogram temsillerine dönüştürülmüştür. Önerilen model mimarisinde, spektral ve uzamsal örüntülerin öğrenilmesi amacıyla evrişimsel sinir ağları (CNN) kullanılırken, kardiyak döngüler boyunca ortaya çıkan zamansal bağımlılıkların modellenmesi için Uzun Kısa Süreli Bellek (LSTM) katmanları entegre edilmiştir.

Bu yaklaşım, literatürde zaman–frekans tabanlı özelliklerin derin öğrenme ile birlikte kullanımının kalp sesi analizinde etkili olduğunu ortaya koyan çalışmalarla uyumludur (Springer vd., 2016).

Buna paralel olarak, toplam 128 adet el ile tasarlanmış geleneksel akustik özellik çıkarılmış ve bu özellikler, klasik makine öğrenmesi modellerinin eğitilmesinde karşılaştırma amacıyla temel (baseline) yaklaşım olarak kullanılmıştır. Elde edilen deneysel bulgular, önerilen hibrit CNN–LSTM mimarisinin hem geleneksel makine öğrenmesi tabanlı modelleri hem de çalışmada değerlendirilen diğer derin öğrenme yaklaşımlarını anlamlı biçimde geride bıraktığını göstermektedir. Bu sonuç, spektral ve zamansal öğrenme stratejilerinin birlikte kullanılmasının pediatrik kalp üfürümü tespitinde önemli bir performans avantajı sağladığını açık biçimde doğrulamaktadır.

6.3. Başlıca Bulguların Özeti

Araştırmanın en önemli bulgularından biri, önerilen hibrit CNN–LSTM mimarisinin yüksek ve dengeli bir sınıflandırma performansı sergilemesidir. Model, test kümesinde %84,6 doğruluk, 0,748 kesinlik (precision), 0,800 duyarlılık (recall), 0,846 F1-skoru ve 0,879 ROC–AUC değeri elde etmiştir. Bu performans metrikleri, modelin hem pozitif vakaları doğru tespit etme hem de yanlış pozitifleri sınırlama açısından dengeli bir öğrenme gerçekleştirdiğini göstermektedir.

Elde edilen sonuçlar, önerilen mimarinin pediatrik kalp üfürümü taraması için klinik açıdan anlamlı ve güvenilir bir performans düzeyi sunduğunu ortaya koymaktadır. Özellikle duyarlılık (recall) değerinin yüksek olması, gerçek üfürüm vakalarının önemli bir kısmının doğru biçimde tespit edilebildiğini ve dolayısıyla yanlış negatif oranının sınırlı kaldığını göstermektedir. Bu durum, tarama amaçlı klinik uygulamalarda kritik öneme sahiptir, zira kaçırılan vakaların azaltılması erken tanı süreçlerinin etkinliğini doğrudan etkilemektedir (Clifford vd., 2016).

Bu bulgular, önerilen hibrit yaklaşımın pediatrik tarama senaryolarında erken tespit, risk sınıflandırması ve uygun uzman yönlendirmesi süreçlerini destekleyebilecek güçlü bir potansiyele sahip olduğunu göstermektedir. Dolayısıyla, modelin klinik karar destek sistemlerine entegre edilmesi, özellikle birinci basamak sağlık hizmetlerinde tanısal doğruluğun artırılmasına katkı sağlayabilir.

6.3.1. Spektro-Zamansal Modellemenin Önemi

Önerilen hibrit CNN–LSTM modeli, yalnızca CNN veya yalnızca LSTM tabanlı mimarilere kıyasla daha üstün ve tutarlı bir performans sergilemiştir. Bu bulgu, kalp sesi

sınıflandırma probleminin hem frekans alanındaki özelliklerin hem de ardışık kardiyak döngüler boyunca ortaya çıkan zamansal bağımlılıkların birlikte modellenmesini gerektirdiğini açık biçimde ortaya koymaktadır.

Bu durum, FKG sinyallerinin doğası gereği zamana bağlı, periyodik ve çok ölçekli örüntüler içermesinden kaynaklanmaktadır. Yalnızca statik spektral temsillere dayalı yaklaşımlar, bu dinamik yapının tamamını yakalamakta yetersiz kalabilmektedir. Buna karşılık, evrimsel sinir ağları (CNN) ile spektral ve uzamsal özelliklerin çıkarılması ve Uzun Kısa Süreli Bellek (LSTM) ağları ile zamansal bağımlılıkların modellenmesinin birleştirilmesi, kalp seslerindeki karmaşık örüntülerin daha bütüncül ve ayırt edici biçimde temsil edilmesine olanak sağlamaktadır (Yaseen vd., 2018; Kiranyaz vd., 2021).

Elde edilen sonuçlar, biyomedikal ses işleme literatüründe son yıllarda raporlanan çalışmalarla uyumlu olup, hibrit derin öğrenme mimarilerinin spektral ve zamansal bilgiyi entegre ederek daha güçlü temsil öğrenme kapasitesi sunduğunu doğrulamaktadır. Bu yaklaşım, özellikle pediatrik kalp üfürümü tespitinde daha yüksek doğruluk ve genellenebilirlik sağlayan modellerin geliştirilmesi açısından önemli bir metodolojik avantaj sunmaktadır.

6.3.2. Geleneksel Makine Öğreniminin İçgörü Sağlaması Ancak Optimal Performans Sunmaması

Geleneksel makine öğrenmesi modelleri, uygulanan özellik seçimi stratejileri sayesinde belirli bir performans artışı sağlamış ve karşılaştırmalı analizler için anlamlı bir referans (baseline) noktası oluşturmuştur. Bununla birlikte, en iyi performansı gösteren geleneksel model olan Destek Vektör Makineleri (SVM)'nin yaklaşık $F1 = 0,716$ seviyesinde doygunluğa ulaşması, yalnızca el ile tasarlanmış özelliklere dayanan yaklaşımların, özellikle farklı akustik imzalara ve değişken zamansal yapılarla sahip kalp üfürümleri söz konusu olduğunda, FKG sinyallerinin karmaşıklığını tam anlamıyla temsil etmekte yetersiz kaldığını göstermektedir (Guyon ve Elisseeff, 2003).

Bu bulgu, derin öğrenme yöntemlerinin sunduğu otomatik ve hiyerarşik özellik öğrenme kapasitesinin önemini açıkça ortaya koymaktadır. Derin öğrenme tabanlı mimariler, ham veya dönüştürülmüş sinyaller üzerinden çok katmanlı temsiller öğrenerek, el ile tasarlanmış özelliklerin yakalayamadığı ince spektral ve zamansal örüntüleri ortaya çıkarabilmektedir. Bu durum, özellikle karmaşık ve gürültüye duyarlı biyomedikal sinyallerin analizinde derin öğrenme yaklaşımlarının neden daha üstün performans sergilediğini açıklayan temel faktörlerden biri olarak değerlendirilmektedir (LeCun vd., 2015).

6.3.3. Sağlam Ön İşlemenin Model Genellemesine Katkısı

Bu çalışmada uygulanan veri artırma stratejileri, hasta düzeyinde veri bölme yaklaşımı ve sabit uzunluklu segment normalizasyonu, aşırı öğrenmenin (overfitting) önlenmesi ve klinik açıdan anlamlı bir genelleme performansı elde edilmesi açısından kritik bir rol oynamıştır. Özellikle hasta düzeyinde gerçekleştirilen veri bölme işlemi, aynı hastaya ait örneklerin eğitim ve test kümeleri arasında paylaşılmasını engelleyerek veri sızıntısını (data leakage) önlemiş ve elde edilen performans ölçütlerinin daha gerçekçi, tarafsız ve klinik açıdan güvenilir olmasını sağlamıştır (Kaufman vd., 2012).

Buna ek olarak, uygulanan veri artırma teknikleri sınıf dengesizliğinin olumsuz etkilerini azaltarak modelin farklı akustik varyasyonlara karşı daha dayanıklı hâle gelmesine katkıda bulunmuştur. Veri artırma yaklaşımları, özellikle sınırlı veri kümelerinde eğitilen derin öğrenme modellerinde genelleme kapasitesini artıran önemli bir düzenleme (regularization) mekanizması olarak değerlendirilmektedir (Shorten ve Khoshgoftaar, 2019). Bu sayede model, eğitim verisine aşırı uyum sağlamak yerine daha genellenebilir temsiller öğrenmiş ve görülmemiş veriler üzerinde daha kararlı bir performans sergilemiştir.

6.4. Bilimsel ve Klinik Katkıları

Bu tez, hem yapay zekâ temelli biyomedikal sinyal işleme literatürüne hem de klinik kardiyak değerlendirme alanına önemli katkılar sunmaktadır. Metodolojik açıdan, pediatrik FKG sinyallerine özgü olarak tasarlanmış hibrit bir CNN–LSTM mimarisi geliştirilmiş; Mel-spektrogram temsilleri ile el ile tasarlanmış geleneksel akustik özellikler birleştirilerek zengin ve çok boyutlu bir özellik uzayı oluşturulmuştur. Ayrıca, yeniden üretilebilirliği yüksek ve bilimsel açıdan güvenilir sonuçlar sağlayan sistematik bir ön işleme ve deneysel değerlendirme hattı tanımlanmıştır.

Klinik açıdan değerlendirildiğinde, geliştirilen sistem pediatrikler, pratisyen hekimler ve uzman olmayan sağlık çalışanları için erken kalp üfürümü tespitinde etkili bir karar destek aracı olma potansiyeline sahiptir. Bu tür yapay zekâ destekli sistemlerin, özellikle birinci basamak sağlık hizmetlerinde tanısal doğruluğu artırdığı ve klinik iş yükünü azalttığı literatürde de vurgulanmaktadır (Topol, 2019). Ayrıca, önerilen yaklaşım uzman bakımına erişimdeki coğrafi engellerin azaltılmasına katkı sağlayarak teletıp uygulamalarını güçlendirebilir ve kaynakların sınırlı olduğu ortamlarda tarama doğruluğunu artırabilir. Bunun yanı sıra, objektif ve standartlaştırılmış oskültasyon

ölçütleri sunması sayesinde klinisyenlerin öznel değerlendirmelere dayalı karar süreçlerine olan bağımlılığını azaltma potansiyeli taşımaktadır.

Dijital sağlık perspektifinden bakıldığında ise önerilen mimari ve geliştirilen iş hattı; dijital stetoskop sistemlerine, yapay zekâ destekli mobil sağlık uygulamalarına ve uzaktan hasta izleme platformlarına entegre edilebilecek niteliktedir. Bu tür entegrasyonlar, kalp sesi analizinin daha geniş ölçekte uygulanmasını mümkün kılarken, erken tanı ve sürekli izlem süreçlerinin güçlendirilmesine katkı sağlayabilir. Bu yönüyle çalışma, yapay zekâ tabanlı kardiyak değerlendirme sistemlerinin klinik pratiğe entegrasyonu açısından önemli bir adım niteliği taşımaktadır.

6.5. Çalışmanın Sınırlılıkları

Elde edilen güçlü sonuçlara rağmen, bu çalışmanın bazı önemli sınırlılıkları bulunmaktadır. Öncelikle, kullanılan veri kümesi yalnızca PhysioNet kapsamında sunulan CirCor DigiScope veri tabanı ile sınırlıdır. Bu veri kümesi görece geniş ve çeşitli olmakla birlikte, farklı kayıt cihazları, klinik ortamlar ve hasta demografilerinden kaynaklanan tüm varyasyonları tam olarak yansıtmayabilir. Bu durum, modelin farklı veri dağılımlarına sahip gerçek klinik ortamlardaki performansının değerlendirilmesi için harici doğrulama çalışmalarının gerekliliğini ortaya koymaktadır (Clifford vd., 2022).

Buna ek olarak, mevcut çalışma yalnızca ikili üfürüm tespitine odaklanmaktadır. Oysa klinik pratikte masum ve patolojik üfürümlerin ayrımı, sistolik ve diyastolik üfürümlerin sınıflandırılması ve üfürüm şiddetinin derecelendirilmesi gibi daha ayrıntılı tanısal görevler büyük önem taşımaktadır. Bu nedenle, gelecekte geliştirilecek çok sınıflı ve daha ince ayrımlara odaklanan modellerin, klinik karar destek sistemlerinin tanısal değerini ve klinik kullanılabilirliğini önemli ölçüde artırması beklenmektedir.

Değerlendirme sürecinin tek bir tabakalı eğitim–doğrulama–test bölünmesine dayanması da bir diğer sınırlılık olarak değerlendirilebilir. Bu yaklaşım literatürde yaygın olarak kullanılmakla birlikte, model performansının istatistiksel güvenilirliğini daha da güçlendirmek amacıyla gelecekte K-katlı çapraz doğrulama, bootstrap yeniden örnekleme ve McNemar testi gibi istatistiksel anlamlılık analizlerinin uygulanması önerilmektedir (Dietterich, 1998).

Son olarak, hibrit CNN–LSTM mimarisi yüksek doğruluk sağlamasına rağmen, modelin klinik açıdan açıklanabilirliği yeterince ele alınmamıştır. Modelin kararlarını hangi zaman–frekans bölgelerine dayandığının açık ve görsel biçimde ortaya konulması, klinisyenlerin sisteme duyacağı güven açısından kritik öneme sahiptir. Bu nedenle, Grad-CAM, LIME ve SHAP gibi açıklanabilir yapay zekâ yöntemlerinin

gelecekteki çalışmalarda modele entegre edilmesi, modelin klinik kabul edilebilirliğini artıracaktır (Selvaraju vd., 2017; Ribeiro vd., 2016; Lundberg ve Lee, 2017).

6.6. Gelecek Çalışmalar için Öneriler

Gelecek çalışmalarda, bu araştırmanın çok sınıflı üfürüm sınıflandırmasını kapsayacak şekilde genişletilmesi önerilmektedir. Üfürüm tipinin, masum veya patolojik niteliğinin ve şiddet derecesinin ayrı ayrı modellenmesi, klinik karar destek sistemleri açısından daha ayrıntılı, bağlama duyarlı ve klinik olarak anlamlı çıktılar sunabilir.

Ayrıca, açıklanabilir yapay zekâ (Explainable AI – XAI) yaklaşımlarının modele entegre edilmesi, klinisyen güveninin artırılması açısından kritik bir araştırma yönü olarak öne çıkmaktadır. Grad-CAM, SHAP ve dikkat tabanlı görselleştirme teknikleri, Mel-spektrogramlar üzerinde üfürüme özgü zaman–frekans bölgelerinin vurgulanmasını sağlayarak model kararlarının daha şeffaf ve yorumlanabilir hâle getirilmesine katkı sunabilir (Selvaraju vd., 2017; Lundberg ve Lee, 2017).

Modelin gerçek klinik veriler üzerinde doğrulanması da önemli bir gelecek araştırma alanıdır. Gürültülü klinik ortamlar, farklı stetoskop donanımları ve değişken kayıt koşullarını içeren hastane veri setleri üzerinde gerçekleştirilecek harici doğrulama çalışmaları, modelin genellenebilirliğini ve klinik güvenilirliğini daha güçlü biçimde değerlendirmeye olanak tanıyacaktır (Kelly vd., 2019).

Buna ek olarak, FKG sinyallerinin EKG verileri, hasta demografik bilgileri ve klinik notlar ile bütünleştirildiği çok modlu öğrenme (multimodal learning) yaklaşımları, tanısal doğruluğun artırılması açısından önemli bir potansiyel sunmaktadır. Literatürde çok modlu biyomedikal veri entegrasyonunun klinik karar destek sistemlerinde performansı artırdığı gösterilmiştir (Baltrusaitis vd., 2019).

Mobil ve gömülü sistemler perspektifinden değerlendirildiğinde ise modelin; model sıkıştırma (model compression), kuantizasyon (quantization) ve bilgi damıtma (knowledge distillation) gibi tekniklerle optimize edilmesi, düşük güç tüketimli dijital stetoskoplar, akıllı telefonlar ve IoT tabanlı tıbbi cihazlar üzerinde çalıştırılmasını mümkün kılabilir. Bu doğrultuda, gerçek zamanlı sinyal analizi, etkileşimli üfürüm uyarıları ve klinik karar destek sistemleri ile entegrasyon gibi yetenekler, gelecekteki önemli geliştirme alanları arasında yer almaktadır (Han vd., 2016).

6.7. Nihai Sonuç

Bu tez, Mel-spektrogram temsilleri ve geleneksel akustik özelliklerin birlikte kullanıldığı hibrit bir CNN–LSTM çerçevesinin, pediatrik kalp üfürümlerinin otomatik

tespit performansını anlamlı ölçüde iyileştirebileceğini ortaya koymuştur. Önerilen model, hem geleneksel makine öğrenmesi algoritmalarını hem de çalışmada değerlendirilen gelişmiş derin CNN mimarilerini geride bırakarak güçlü, dengeli ve klinik açıdan anlamlı bir sınıflandırma performansı sergilemiştir.

FKG sinyallerinin spektral ve zamansal bileşenlerini eş zamanlı olarak modelleyebilme yeteneği, yapay zekâ tabanlı kardiyak oskültasyon alanına önemli bir metodolojik katkı sunmaktadır. Elde edilen bulgular, dijital stetoskop sistemlerinin, teletıp çözümlerinin ve konjenital kalp hastalıklarının erken taranmasına yönelik otomatik sistemlerin geliştirilmesi açısından önemli çıkarımlar sağlamaktadır. Bu tür yaklaşımların, özellikle birinci basamak sağlık hizmetlerinde tanısal doğruluğu artırma ve klinik karar destek süreçlerini güçlendirme potansiyeline sahip olduğu literatürde de vurgulanmaktadır (Topol, 2019).

Her ne kadar bazı sınırlılıklar mevcut olsa da, bu çalışma otomatik kardiyak ses analizi alanında tanısal doğruluğu, model sağlamlığını ve açıklanabilirliği artırmaya yönelik gelecekteki araştırmalar için güçlü bir temel sunmaktadır. Önerilen yaklaşım, yalnızca araştırma düzeyinde değil, aynı zamanda klinik uygulamalara entegre edilebilir, ölçeklenebilir ve güvenilir bir çözüm geliştirme yönünde önemli bir adım niteliğindedir.

Sunulan çerçeve, yapay zekâ destekli pediatrik kardiyak bakımın geliştirilmesine yönelik küresel araştırma çabalarına katkıda bulunmayı; erken tanıyı iyileştirmeyi, tanısal eşitsizlikleri azaltmayı ve klinisyenlerin daha erişilebilir, standartlaştırılmış ve yüksek kaliteli kardiyovasküler sağlık hizmeti sunmalarını desteklemeyi amaçlamaktadır.

KAYNAKÇA

- Addison, P. S. (2005). Wavelet transforms and the ECG: A review. *Physiological Measurement*, 26(5), R155–R199. doi:10.1088/0967-3334/26/5/R01
- Baltrušaitis, T., Ahuja, C. ve Morency, L.-P. (2019). Multimodal machine learning: A survey and taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(2), 423–443. doi:10.1109/TPAMI.2018.2798607
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. doi:10.1023/A:1010933404324
- Chen, T. ve Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. doi:10.1145/2939672.2939785
- Chen, W., Sun, Q., Chen, X., Xie, G., Wu, H. ve Xu, C. (2021). Deep learning methods for heart sounds classification: A systematic review. *Entropy*, 23(6), 667. doi:10.3390/e23060667
- Chen, W., Zhou, Z., Bao, J., Wang, C., Chen, H., Xu, C., ... Wu, H. (2023). Classifying heart-sound signals based on CNN trained on MelSpectrum and Log-MelSpectrum features. *Bioengineering*, 10(6), 645. doi:10.3390/bioengineering10060645
- Clifford, G. D., Liu, C., Moody, B., Springer, D., Silva, I., Li, Q. ve Mark, R. G. (2016). Classification of normal/abnormal heart sound recordings: The PhysioNet/Computing in Cardiology Challenge 2016. *Computing in Cardiology*, 43, 609–612. doi:10.22489/CinC.2016.179-154
- Cortes, C. ve Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20, 273–297. doi:10.1007/BF00994018
- Cover, T. M. ve Hart, P. E. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21–27. doi:10.1109/TIT.1967.1053964
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874. doi:10.1016/j.patrec.2005.10.010
- Goodfellow, I., Bengio, Y. ve Courville, A. (2016). *Deep learning*. MIT Press.
- Guyon, I. ve Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3, 1157–1182.
- Han, S., Mao, H. ve Dally, W. J. (2016). Deep compression: Compressing deep neural networks with pruning, trained quantization and Huffman coding. *4th International Conference on Learning Representations (ICLR)*.

- Hastie, T., Tibshirani, R. ve Friedman, J. (2009). *The elements of statistical learning* (2. bs.). Springer.
- He, K., Zhang, X., Ren, S. ve Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. doi:10.1109/CVPR.2016.90
- Hochreiter, S. ve Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. doi:10.1162/neco.1997.9.8.1735
- Kaufman, S., Rosset, S. ve Perlich, C. (2012). Leakage in data mining: Formulation, detection, and avoidance. *ACM Transactions on Knowledge Discovery from Data*, 6(4), Article 15. doi:10.1145/2020408.2020496
- Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G. ve King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, 17, Article 195. doi:10.1186/s12916-019-1426-2
- Kiranyaz, S., Ince, T. ve Gabbouj, M. (2016). Real-time patient-specific ECG classification by 1-D convolutional neural networks. *IEEE Transactions on Biomedical Engineering*, 63(3), 664–675. doi:10.1109/TBME.2015.2468589
- LeCun, Y., Bengio, Y. ve Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444. doi:10.1038/nature14539
- Leatham, A. (1991). *Lecture notes on cardiology*. Blackwell.
- Liu, C., Springer, D., Li, Q., Moody, B., Abad Juan, R., Chorro, F. J., ... Clifford, G. D. (2016). An open access database for the evaluation of heart sound algorithms. *Physiological Measurement*, 37(12), 2181–2213. doi:10.1088/0967-3334/37/12/2181
- Liu, X., Faes, L., Kale, A. U., Wagner, S. K., Fu, D. J., Bruynseels, A., ... Denniston, A. K. (2019). A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: A systematic review and meta-analysis. *The Lancet Digital Health*, 1(6), e271–e297. doi:10.1016/S2589-7500(19)30123-2
- Logan, B. (2000). Mel frequency cepstral coefficients for music modeling. *Proceedings of the International Symposium on Music Information Retrieval (ISMIR)*.
- Lundberg, S. M. ve Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765–4774.

- Mangione, S. ve Nieman, L. Z. (1997). Cardiac auscultatory skills of internal medicine and family practice trainees: A comparison of diagnostic proficiency. *JAMA*, 278(9), 717–722. doi:10.1001/jama.1997.03550090041030
- McCrindle, B. W., Shaffer, K. M., Kan, J. S., Zahka, K. G., Rowe, S. A. ve Kidd, L. (1996). Cardinal clinical signs in the differentiation of heart murmurs in children. *Archives of Pediatrics & Adolescent Medicine*, 150(2), 169–174. doi:10.1001/archpedi.1996.02170270051007
- McGee, S. (2018). *Evidence-based physical diagnosis (4. bs.)*. Elsevier.
- Nigam, V. ve Priemer, R. (2005). Heart sound segmentation. *IEEE Transactions on Biomedical Engineering*, 52(10), 1677–1684. doi:10.1109/TBME.2005.855722
- Oliveira, J., Renna, F., Costa, P., Nogueira, M., Oliveira, A. C., Elola, A., ... Coimbra, M. (2022). *The CirCor DigiScope phonocardiogram dataset (Version 1.0.3)*. PhysioNet. doi:10.13026/8k54-kc42
- Potes, C., Parvaneh, S., Rahman, A. ve Conroy, B. (2016). Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds. *Computing in Cardiology*, 43.
- Rabiner, L. ve Schafer, R. (2011). *Digital speech processing*. Pearson.
- Reyna, M. A., Kiarashi, Y., Elola, A., Oliveira, J., Renna, F., Gu, A., ... Clifford, G. D. (2023). Heart murmur detection from phonocardiogram recordings: The George B. Moody PhysioNet Challenge 2022. *PLOS Digital Health*, 2(9), e0000324. doi:10.1371/journal.pdig.0000324
- Ribeiro, M. T., Singh, S. ve Guestrin, C. (2016). “Why should I trust you?”: Explaining the predictions of any classifier. *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*, 97–101. doi:10.18653/v1/N16-3020
- Rubin, J., Abreu, R., Ganguli, A., Nelaturi, S., Matei, I. ve Sricharan, K. (2016). Classifying heart sound recordings using deep convolutional neural networks and mel-frequency cepstral coefficients. *Computing in Cardiology*, 43. doi:10.22489/CinC.2016.236-175
- Saito, T. ve Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10(3), e0118432. doi:10.1371/journal.pone.0118432
- Schmidt, S. E., Holst-Hansen, C., Graff, C., Toft, E. ve Struijk, J. J. (2010). Segmentation of heart sound recordings by a duration-dependent hidden Markov model. *Physiological Measurement*, 31(4), 513–529. doi:10.1088/0967-3334/31/4/004

- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D. ve Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 618–626. doi:10.1109/ICCV.2017.74
- Shorten, C., Khoshgoftaar, T. M. ve Furht, B. (2021). Text data augmentation for deep learning. *Journal of Big Data*, 8, Article 101. doi:10.1186/s40537-021-00492-0
- Springer, D. B., Tarassenko, L. ve Clifford, G. D. (2016). Logistic regression-HSMM-based heart sound segmentation. *IEEE Transactions on Biomedical Engineering*, 63(4), 822–832. doi:10.1109/TBME.2015.2475278
- Stevens, S. S., Volkman, J. ve Newman, E. B. (1937). A scale for the measurement of the psychological magnitude pitch. *The Journal of the Acoustical Society of America*, 8(3), 185–190. doi:10.1121/1.1915893
- Thompson, W. R., Reinisch, A. J., Unterberger, M. J. ve Schriebl, A. J. (2019). Artificial intelligence-assisted auscultation of heart murmurs: Validation by virtual clinical trial. *Pediatric Cardiology*, 40(3), 623–629. doi:10.1007/s00246-018-2036-z
- Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44–56. doi:10.1038/s41591-018-0300-7
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.
- Yaseen, Son, G.-Y. ve Kwon, S. (2018). Classification of heart sound signal using multiple features. *Applied Sciences*, 8(12), 2344. doi:10.3390/app8122344

ÖZGEÇMİŞ

Asmaa ALTAIB, lise eğitimini 2011 yılında Yemen'in Taiz şehrinde bulunan Zaid Almoshiki Lisesi'nde tamamlamıştır. Ardından Taiz Üniversitesi, Al-Saeed Mühendislik ve Bilgi Teknolojileri Fakültesi'nde Bilgi Teknolojileri Mühendisliği (Bilgisayar Mühendisliği) alanında lisans eğitimini sürdürmüş ve 2018 yılında mezun olmuştur.

Mezuniyetinin ardından bilgi teknolojileri alanında çeşitli görevlerde bulunmuş, sistem yönetimi ve teknoloji altyapılarıyla ilgili farklı sorumluluklar üstlenmiştir. Hâlen Suudi Arabistan'da bulunan Wareed Laboratuvarları'nda Yapay Zekâ Mühendisi olarak çalışmaktadır. Bunun yanında CBS şirketinde Yapay Zekâ Uzmanı olarak görev yapmaktadır. Araştırma ilgi alanları arasında yapay zekâ, derin öğrenme ve yapay zekânın tıbbi uygulamaları yer almaktadır.

Yayınlar:

Al-Taib, A. A., Türe, H. ve Aslan, M. E. (2025). Hybrid CNN–LSTM framework for automated detection of pediatric heart murmurs. *2025 Medical Technologies Congress (TIPTEKNO)*, IEEE, ss. 1–4.