



T.C.

BARTIN ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

BİLİŞİM SİSTEMLERİ VE TEKNOLOJİLERİ ANABİLİM DALI

YÜKSEK LİSANS TEZİ

**KULAK GÖRÜNTÜLERİNDEN SOFT BİYOMETRİK
ÖZELLİKLERİN MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE
SINIFLANDIRILMASI**

MERTKAN SİNOPLU

DANIŞMAN

DR. ÖĞR. ÜYESİ EYÜP BURAK CEYHAN

BARTIN-2023



T.C.

BARTIN ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

BİLİŞİM SİSTEMLERİ VE TEKNOLOJİLERİ ANABİLİM DALI

**KULAK GÖRÜNTÜLERİNDEN SOFT BİYOMETRİK ÖZELLİKLERİN
MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE SINIFLANDIRILMASI**

YÜKSEK LİSANS TEZİ

Mertkan SİNOPLU

BARTIN-2023

KABUL VE ONAY

Mertkan SİNOPLU tarafından hazırlanan “KULAK GÖRÜNTÜLERİNDEN SOFT BİYOMETRİK ÖZELLİKLERİN MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE SINIFLANDIRILMASI” başlıklı bu çalışma, 03.05.2023 tarihinde yapılan savunma sınavı sonucunda oy birliği ile başarılı bulunarak jürimiz tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.

Başkan : Doç. Dr. Yusuf Ziya OLPAK

Üye : Doç. Dr. Ramazan YILMAZ

Üye : Dr. Öğr. Üyesi Eyüp Burak CEYHAN

Bu tezin kabulü Lisansüstü Eğitimi Enstitüsü Yönetim Kurulu’nun/...../20... tarih ve 20...../.....-..... sayılı kararıyla onaylanmıştır.

Prof. Dr. H. Selma ÇELİKİYAY
Enstitü Müdürü

BEYANNAME

Bartın Üniversitesi Lisansüstü Eğitim Enstitüsü tez yazım kılavuzuna göre Dr. Öğr. Üyesi Eyüp Burak CEYHAN danışmanlığında hazırlamış olduğum “KULAK GÖRÜNTÜLERİNDEN SOFT BİYOMETRİK ÖZELLİKLERİN MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE SINIFLANDIRILMASI” başlıklı Yüksek Lisans tezimin bilimsel etik değerlere ve kurallara uygun, özgün bir çalışma olduğunu, aksinin tespit edilmesi halinde her türlü yasal yaptırımını kabul edeceğimi beyan ederim.

03.05.2023

Mertkan SİNOPLU

ÖNSÖZ

Bu çalışmada danışmanlığımı üstlenerek tezi hazırlama ve tamamlama süreçlerinde yardımlarını esirgemeyen değerli hocam Dr. Öğr. Üyesi Eyüp Burak CEYHAN'a teşekkürlerimi ve saygılarımı sunarım.

Geçirdiğim yüksek lisans sürecinde yaşadığım sorunlar hakkında yardımları esirgemeyen Doç. Dr. Ramazan YILMAZ hocama teşekkür ederim.

Yüksek lisans sürecinde destek için her zaman arkamda olan öncelikle babam Oktay SİNOPLU, annem Aylin SİNOPLU ve ablam Tuğçe SİNOPLU olmak üzere tüm aileme teşekkür ederim.



ÖZET

Yüksek Lisans Tezi

KULAK GÖRÜNTÜLERİNDEN SOFT BİYOMETRİK ÖZELLİKLERİN MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE SINIFLANDIRILMASI

Mertkan SİNOPLU

Bartın Üniversitesi

Lisansüstü Eğitim Enstitüsü

Bilişim Sistemleri ve Teknolojileri Anabilim Dalı

Tez Danışmanı: Dr. Öğr. Üyesi Eyüp Burak CEYHAN

Bartın-2023, sayfa: 79

Yapılan çalışmanın amacı, kulak görüntülerinden cinsiyet, yaş, boy, doğal saç rengi, saç tipi, göz rengi ve ten rengi gibi soft biyometrik özelliklerin makine öğrenmesi yöntemleri ile sınıflandırılmasıdır. Çalışma kapsamında etik kurul belgesi alınarak 105 kişiden kulak görüntüsü toplanmış ve bu kişilere birer anket uygulanmıştır. Buna ek olarak internet üzerinden elde edilen kulak görüntüleri ve soft biyometrik veriler yardımı ile 250 kişiye ait 300 kulak görüntüsü ile ikinci bir veri tabanı da sınıflandırma testleri için oluşturulmuştur.

Çalışmanın kulak görüntüleri ön işleme bölümünde, kulak görüntülerinin kesilmesi ve yeniden boyutlandırılması için Python programlama dili kullanılmıştır. Daha sonra alınan görüntüler üzerinde yapılan ön işlemede Matlab uygulaması tercih edilmiştir. Görüntüler öncelikle Matlab uygulamasında grayscale formata dönüştürülmüştür. Daha sonra, ortam ışık koşullarından oluşabilecek farkların minimum seviyeye indirilmesi amacıyla histogram eşitlemesi yapılmıştır. Yapılan histogram eşitlemesinden sonra, görüntüler siyah beyaz formata dönüştürülmüştür. Son olarak elde edilen siyah-beyaz formattaki görüntülerin piksel değerleri Excel dosyasına aktarılarak, her bir piksel değeri bir sütuna aktarılmıştır. Her bir soft biyometrik özellik için ayrı Excel dosyaları oluşturulmuş, tüm görüntülerin piksel değerleri Excel dosyalarına aktarıldıktan sonra sınıf değerleri son sütuna kaydedilmiştir.

Çalışma kapsamında hedeflenen soft biyometrik özelliklerin tahmini için RapidMiner ve Weka veri işleme uygulamaları tercih edilmiştir. Bu uygulamalardaki sınıflandırma algoritmalarından Derin Öğrenme, Destek Vektör Makineleri, LibSVM, Lojistik Regresyon, Doğrusal Regresyon, Rastgele Orman ve Naive Bayes algoritmaları kullanılmıştır. 10 katlamalı çapraz doğrulama yöntemi kullanılarak yapılan testler sonucunda uygulamalar arasında doğruluğu en yüksek olan iki algoritma detaylı olarak karşılaştırılmıştır.

105 kişiden oluşan ilk veri tabanı için cinsiyet sınıflandırma sonuçları incelendiğinde, en başarılı algoritmanın %84,76 ile Weka uygulamasındaki LibSVM algoritması olduğu bulunmuştur. Yaş tahmini için bakıldığında, en başarılı algoritmanın %40,00 ile Rapidminer uygulamasındaki LibSVM algoritması olduğu tespit edilmiştir. Boy tahmini konusunda en başarılı algoritmanın %44,82 sınıflandırma başarısı ile RapidMiner uygulamasındaki LibSVM algoritması olduğu görülmüştür. Göz rengi ve doğal saç rengi tahmini için yapılan sonuçlara bakıldığında, en başarılı algoritmanın sırasıyla %68,47 ve %45,91 doğruluğa sahip RapidMiner uygulamasındaki Derin Öğrenme algoritması olduğu görülmüştür. Ten rengi tahmini için yapılan testler ile en başarılı sınıflandırma algoritmasının %50 doğruluk ile Weka uygulamasındaki LibSVM algoritması olduğu görülmüştür. Son olarak saç tipi tahmini için yapılan testlerde, RapidMiner uygulamasındaki LibSVM algoritmasının %68,89 doğruluk oranı ile en başarılı algoritma olduğu sonucuna ulaşılmıştır.

300 kişiden oluşan ikinci veri tabanı için cinsiyet sınıflandırma sonuçları incelendiğinde, en başarılı algoritmanın %86,33 ile Weka uygulamasındaki LibSVM algoritması olduğu bulunmuştur. Yaş tahmini için bakıldığında, en başarılı algoritmanın %68,40 ile Rapidminer uygulamasındaki LibSVM algoritması olduğu tespit edilmiştir. Boy tahmini konusunda en başarılı algoritmanın %60,74 sınıflandırma başarısı ile Weka uygulamasındaki LibSVM algoritması olduğu görülmüştür. Göz rengi tahmini için yapılan sonuçlara bakıldığında, en başarılı algoritmanın %80,59 ile Weka uygulamasındaki LibSVM algoritması olduğu görülmüştür. Buna ek olarak doğal saç rengi için yapılan testler sonucunda, en başarılı algoritmanın %57,62 ile RapidMiner uygulamasındaki Derin Öğrenme algoritması olduğu görülmüştür. Ten rengi tahmini için yapılan testler ile en başarılı sınıflandırma algoritmasının %60,68 doğruluk ile Weka uygulamasındaki LibSVM algoritması olduğu görülmüştür. Son olarak saç tipi tahmini için yapılan testlerde, RapidMiner uygulamasındaki LibSVM algoritmasının %69,06 doğruluk oranı ile en başarılı algoritma olduğu sonucuna ulaşılmıştır.

Kulak görüntüsünden cinsiyet ve yaş tahmin eden çalışmalar az da olsa olmakla birlikte, saç tipi, doğal saç rengi, boy ve göz rengi tahmini ile ilgili herhangi bir çalışmaya rastlanmamıştır. Bu bakımdan tez çalışmamız literatürde ilk kez kulak görüntüsünden saç tipi, doğal saç rengi, boy ve göz rengi tahmin sonuçlarını içermektedir.

Anahtar Kelimeler: Soft biyometrik, kulak, görüntü işleme, yapay zekâ, derin öğrenme.



ABSTRACT

M. Sc. Thesis

CLASSIFICATION OF SOFT BIOMETRIC FEATURES FROM EAR IMAGES BY MACHINE LEARNING METHODS

Mertkan SİNOPLU

Bartın University

Graduate School

Department of Information Systems and Technologies

Thesis Advisor: Assist. Prof. Dr. Eyüp Burak CEYHAN

Bartın-2023, pp: 79

The aim of the study is to classify soft biometric features such as gender, age, height, natural hair color, hair type, eye color and skin color from ear images by machine learning methods. Within the scope of the study, the ear images were collected from 105 people by obtaining the ethics committee certificate and a questionnaire was applied to each of these people. In addition, with the help of ear images and soft biometric data obtained over the internet, a second database was created with 300 ear images of 250 people for classification tests.

In the ear images preprocessing part of the study, Python programming language was used for cutting and resizing the ear images. Matlab application was preferred for preprocessing on the images taken afterwards. The images were first converted to grayscale format in Matlab application. Then, histogram equalization was performed in order to minimize the differences that may occur from ambient light conditions. After the histogram equalization, the images were converted to black and white format. Finally, the pixel values of the black-and-white images were transferred to the Excel file, and each pixel value was transferred to a column. Separate Excel files were created for each soft biometric feature, and the class values were recorded in the last column after the pixel values of all images were transferred to Excel files.

RapidMiner and Weka data processing applications were preferred for the estimation of the soft biometric features targeted within the scope of the study. Among the classification algorithms in these applications, Deep Learning, Support Vector Machines, LibSVM, Logistic Regression, Linear Regression, Random Forest and Naive Bayes algorithms were used. As a result of the tests using the 10-fold cross validation method, the two algorithms with the highest accuracy among the applications were compared in detail.

When the gender classification results for the first database consisting of 105 people were examined, it was found that the most successful algorithm was the LibSVM algorithm in the Weka application with 84.76%. For age estimation, it was determined that the most successful algorithm was the LibSVM algorithm in Rapidminer application with 40.00%. It has been seen that the most successful algorithms for height estimation is the LibSVM algorithm in the RapidMiner application, with a classification success of 44.82%. Looking at the results for eye color and natural hair color prediction, it was seen that the most successful algorithm was the Deep Learning algorithm in the RapidMiner application, which had an accuracy of 68.47% and 45.91%, respectively. With the tests made for skin color prediction, it was seen that the most successful classification algorithm was the LibSVM algorithm in the Weka application with 50% accuracy. Finally, in the tests for hair type prediction, it was concluded that the LibSVM algorithm in RapidMiner application was the most successful algorithm with an accuracy rate of 68.89%.

When the gender classification results for the second database consisting of 300 people were examined, it was found that the most successful algorithm was the LibSVM algorithm in the Weka application with 86.33%. For age estimation, it was determined that the most successful algorithm was the LibSVM algorithm in Rapidminer application with 68.40%. It has been seen that the most successful algorithm for height estimation is the LibSVM algorithm in the Weka application with a classification success of 60.74%. Looking at the results for eye color estimation, it was seen that the most successful algorithm was the LibSVM algorithm in the Weka application with 80.59%. In addition, as a result of the tests for natural hair color, it was seen that the most successful algorithm was the Deep Learning algorithm in RapidMiner application with 57.62%. With the tests made for skin color prediction, it was seen that the most successful classification algorithm was the LibSVM algorithm in the Weka application with an accuracy of 60.68%. Finally, in the tests for hair type prediction, it was concluded that the LibSVM algorithm in RapidMiner application was

the most successful algorithm with an accuracy rate of 69.06%.

Although there are few studies estimating gender and age from ear image, no studies related to hair type, natural hair color, height and eye color estimation were found. In this respect, our thesis study includes hair type, natural hair color, height and eye color prediction results from ear image for the first time in the literature.

Keywords: Soft biometric, ear, image processing, artificial intelligence, deep learning.



İÇİNDEKİLER

KABUL VE ONAY.....	ii
BEYANNAME	iii
ÖNSÖZ	iv
ÖZET	v
ABSTRACT	viii
İÇİNDEKİLER.....	xi
ŞEKİLLER DİZİNİ.....	xiv
TABLolar DİZİNİ.....	xvi
EKLER DİZİNİ	xix
SİMGELER VE KISALTMALAR DİZİNİ.....	xx
1. GİRİŞ.....	1
2. BİYOMETRİK SİSTEMLER.....	5
2.1. Biyometrik Özellikler	6
2.1.1 Soft Biyometrik Özellikler	6
2.2. Biyometrik Sistemlerin Çalışma Prensipleri.....	7
2.3. Yaygın Kullanılan Biyometrik Özellikler.....	8
2.3.1 DNA Tanıma	8
2.3.2 Yüz Tanıma	9
2.3.3 Retina ve İris Tanıma	10
2.3.4 Ses Tanıma.....	11
2.3.5 Parmak İzi Tanıma	11
2.3.6 Damar Tanıma	12
2.3.7 Kulak Tanıma	13
3. KULAK TANIMA SİSTEMLERİ VE İŞLEMLERİ.....	15
3.1 Kulak Biyometriğinin Uygulama Alanları	16
3.2 Kulak Biyometriğinin Güçlü ve Zayıf Yönleri.....	17
3.3 Ön İşleme ve Öznitelik Çıkarımı.....	18
3.4 Eşleştirme	20
4. SINIFLANDIRMA ALGORİTMALARI	21
4.1 Destek Vektör Makineleri Algoritması.....	21
4.2 Derin Öğrenme Algoritması	22

4.3 Lojistik Regresyon Algoritması.....	23
4.4 Doğrusal Regresyon Algoritması.....	23
4.5 LibSVM Algoritması	24
4.6 Rastgele Orman Algoritması	25
4.7 Naive Bayes Algoritması	27
5. KULAK BİYOMETRİĞİ ÇALIŞMALARI	28
6. KULAK VERİ SETLERİNİN HAZIRLANMASI.....	33
6.1.1 VT1 İçin Kulak Görüntülerinin Elde Edilmesi ve Kaydedilmesi.....	33
6.1.2 VT1 İçin Anket Verilerinin Veri Tabanına Kaydedilmesi.....	34
6.2 VT2 İçin Kulak Görüntülerinin ve Soft Biyometrik Özelliklerin Elde Edilmesi ve Kaydedilmesi	34
7. KULAK GÖRÜNTÜSÜNDEN SOFT BİYOMETRİK ÖZELLİKLERİN TAHMİN EDİLMESİ	35
7.1 Kulak Görüntüleri Ön İşlemleri.....	35
7.2 Soft Biyometrik Tahmin Sisteminin Geliştirilmesi.....	37
7.3 Tahmin İşlemleri İçin Kıstaslar.....	38
7.4 VT1 İçin Tahmin İşlemleri.....	40
7.4.1 VT1 İçin Cinsiyet Tahmini.....	40
7.4.2 VT1 İçin Yaş Tahmini	42
7.4.3 VT1 İçin Boy Tahmini.....	44
7.4.4 VT1 İçin Göz Rengi Tahmini.....	46
7.4.5 VT1 İçin Doğal Saç Rengi Tahmini.....	47
7.4.6 VT1 İçin Ten Rengi Tahmini.....	49
7.4.7 VT1 İçin Saç Tipi Tahmini	52
7.5 VT2 İçin Tahmin İşlemleri.....	53
7.5.1 VT2 İçin Cinsiyet Tahmini.....	53
7.5.2 VT2 İçin Yaş Tahmini	55
7.5.3 VT2 İçin Boy Tahmini.....	56
7.5.4 VT2 İçin Göz Rengi Tahmini.....	58
7.5.5 VT2 İçin Doğal Saç Rengi Tahmini.....	60
7.5.6 VT2 İçin Ten Rengi Tahmini.....	62
7.5.7 VT2 İçin Saç Tipi Tahmini	65
8. SONUÇ VE ÖNERİLER	67
KAYNAKLAR.....	70

EKLER.....	76
ÖZGEÇMİŞ.....	78



ŞEKİLLER DİZİNİ

Şekil	Sayfa
No	No
1.1: Temel biyometrik özellikler (Ceyhan, 2012)	2
2.1: Soft biyometrik özellikler.....	7
2.2: DNA ve içeriğindeki nükleotitler (Travers ve Muskhelishvili, 2015)	8
2.3: Yüz tanıma sistemlerinin yapısı (Kortli vd., 2020).....	9
2.4: Retina ağı (Manav, 2010).....	10
2.5: Parmak izi örneği (Ceyhan, 2012).....	12
2.6: El damar ağı (Wang ve Wang, 2017).....	13
2.7: Kulak tanımda kullanılan öznitelik noktaları (Yaman vd., 2018).....	13
3.1: Dış kulak anatomisi. (Kavipriya ve Muthukumar, 2019).....	15
3.2: Öznitelik çıkarım yöntemleri. (Mutlag vd., 2020)	19
3.3: Kulak tanıma işlemi. (Hansley vd., 2018).....	20
4.1: Doğrusal ayırma formülü (Küçüksille ve Ateş, 2013)	21
4.2: Doğrusal ayırma ve destek vektörleri (Küçüksille ve Ateş, 2013)	22
4.3: Örnek bir ileri beslemeli yapay sinir ağı yapısı (Arı ve Berberler, 2017).....	22
4.4: Lojistik regresyon formülü (Dreiseitl ve Ohno-Machado, 2002)	23
4.5: Örnek bir lojistik regresyon grafiği (Panesar vd., 2019).....	23
4.6: Örnek bir doğrusal regresyon grafiği (Potok, 2021)	24
4.7: Örnek bir çoklu doğrusal regresyon formülü (Kılıç, 2013)	24
4.8: Örnek bir LibSVM mimarisi (Chen vd., 2020).....	25
4.9: Rastgele Orman algoritmasının çalışma prensibi (Gajjar ve Zaveri, 2017).....	26
4.10: Homojenlik hesaplama formülü (Özlem ve Güngör, 2012).....	26
4.11: Bayes teoremi (Solmaz vd., 2014)	27
4.12: Basit bir Naive Bayes yapısı (Haltaş ve Alkan, 2013).....	27
6.1: VT1 veri tabanındaki kulak görüntüsü örneği.....	33
6.2: VT2 veri tabanındaki kulak görüntüsü örneği.....	34
7.1: Kırpma ve yeniden boyutlandırma işleminden geçen kulak görüntüsü a) VT1 b) VT2	35
7.2: Kırpma, yeniden boyutlandırma ve grayscale formata çevrilme işleminden geçen kulak görüntüsü a) VT1 b) VT2.....	36

7.3: Kırpma, yeniden boyutlandırma, grayscale formata çevrilme ve histogram eşitleme işleminden geçen kulak görüntüsü a) VT1 b) VT2.....	36
7.4: Kırpma, yeniden boyutlandırma, grayscale formata çevrilme, histogram eşitleme ve medyan filtre işleminden geçen örnek bir VT2 veri tabanı görüntüsü.....	37
7.5: Kırpma, yeniden boyutlandırma, grayscale formata çevrilme, histogram eşitleme ve siyah-beyaz işleminden geçen kulak görüntüsü	37
7.6: Örnek bir hata matrisi.....	39
7.7: Recall hesaplama formülü.....	39
7.8: Precision hesaplama formülü	39
7.9: F1 skoru hesaplama formülü	40
7.10: Sınıflandırma başarısı hesaplama formülü	40
7.11: VT1 için yaş gruplarının veri dağılımları.....	42
7.12: VT1 için boy gruplarının veri dağılımları	44
7.13: VT1 için göz rengi gruplarının veri dağılımları	46
7.14: VT1 için doğal saç rengi gruplarının veri dağılımları.....	48
7.15: VT1 için ten rengi gruplarının veri dağılımları.....	50
7.16: VT1 için saç tipi gruplarının veri dağılımları.....	52
7.17: VT2 için yaş gruplarının veri dağılımları.....	55
7.18: VT2 için boy gruplarının veri dağılımları	57
7.19: VT2 için göz rengi gruplarının veri dağılımları	59
7.20: VT2 için doğal saç rengi gruplarının veri dağılımları.....	61
7.21: VT2 için ten rengi gruplarının veri dağılımları	63
7.22: VT2 için saç tipi gruplarının veri dağılımları.....	65

TABLÖLER DİZİNİ

Tablo No	Sayfa No
2.1: Biyometrik sistemlerin kullanım alanları (Guennouni vd., 2019)	5
2.2: Biyometrik özelliklerin sınıflandırılması	6
5.1: Literatürdeki kulak biyometriği çalışmaları.....	32
7.1: VT1 için RapidMiner uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının cinsiyet sınıflandırma sonucu.....	41
7.2: VT1 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının cinsiyet sınıflandırma sonucu.....	41
7.3: VT1 için RapidMiner ve Weka uygulamalarının cinsiyet tahmini için en başarılı algoritmalarının karşılaştırılması.	41
7.4: VT1 için RapidMiner uygulamasında çalıştırılan Derin Öğrenme ve DVM algoritmalarının yaş sınıflandırma sonucu.	42
7.5: VT1 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının yaş sınıflandırma sonucu.....	43
7.6: VT1 için RapidMiner ve Weka uygulamalarının yaş tahmini üzerindeki en başarılı algoritmalarının karşılaştırılması.....	43
7.7: VT1 için RapidMiner uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının boy sınıflandırma sonucu.	44
7.8: VT1 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının boy sınıflandırma sonucu.	45
7.9: VT1 için RapidMiner ve Weka uygulamalarının boy tahmini üzerindeki en başarılı algoritmalarının karşılaştırılması.....	45
7.10: VT1 için RapidMiner uygulamasında çalıştırılan Derin Öğrenme ve LibSVM algoritmalarının göz rengi sınıflandırma sonucu	46
7.11: VT1 için Weka uygulamasında çalıştırılan Derin Öğrenme ve LibSVM algoritmalarının göz rengi sınıflandırma sonucu.	47
7.12: VT1 için RapidMiner ve Weka uygulamalarının göz rengi tahmini için en başarılı algoritmalarının karşılaştırılması.....	47
7.13: VT1 için RapidMiner uygulamasında çalıştırılan Derin Öğrenme ve Lojistik Regresyon algoritmalarının doğal saç rengi sınıflandırma sonucu.....	48
7.14: VT1 için Weka uygulamasında çalıştırılan Derin Öğrenme ve LibSVM algoritmalarının doğal saç rengi sınıflandırma sonucu.	49

7.15: VT1 için RapidMiner ve Weka uygulamalarının doğal saç rengi tahmini üzerindeki en başarılı algoritmalarının karşılaştırılması.....	49
7.16: VT1 için RapidMiner uygulamasında çalıştırılan Derin Öğrenme ve LibSVM algoritmalarının ten rengi sınıflandırma sonucu.	50
7.17: VT1 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının ten rengi sınıflandırma sonucu.	51
7.18: VT1 için RapidMiner ve Weka uygulamalarının ten rengi tahmini üzerindeki en başarılı algoritmalarının karşılaştırılması.....	51
7.19: VT1 için RapidMiner uygulamasında çalıştırılan LibSVM ve Rastgele Orman algoritmalarının saç tipi sınıflandırma sonucu.	52
7.20: VT1 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının saç tipi sınıflandırma sonucu.....	53
7.21: VT1 için RapidMiner ve Weka uygulamalarının saç tipi tahmini için en başarılı algoritmalarının karşılaştırılması.	53
7.22: VT2 için RapidMiner uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının cinsiyet sınıflandırma sonucu.....	54
7.23: VT2 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının cinsiyet sınıflandırma sonucu.....	54
7.24: VT2 için RapidMiner ve Weka uygulamalarının cinsiyet tahmini için en başarılı algoritmalarının karşılaştırılması.	54
7.25: VT2 için RapidMiner uygulamasında çalıştırılan LibSVM ve Derin Öğrenme algoritmalarının yaş sınıflandırma sonucu.	55
7.26: VT2 için Weka uygulamasında çalıştırılan LibSVM ve Lojistik Regresyon algoritmalarının yaş sınıflandırma sonucu.	56
7.27: VT2 için RapidMiner ve Weka uygulamalarının yaş tahmini üzerindeki en başarılı algoritmalarının karşılaştırılması.....	56
7.28: VT2 için RapidMiner uygulamasında çalıştırılan Rastgele Orman ve Derin Öğrenme algoritmalarının boy sınıflandırma sonucu.	57
7.29: VT2 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının boy sınıflandırma sonucu.	58
7.30: VT2 için RapidMiner ve Weka uygulamalarının boy tahmini üzerindeki en başarılı algoritmalarının karşılaştırılması.....	58
7.31: VT2 için RapidMiner uygulamasında çalıştırılan LibSVM ve Derin Öğrenme algoritmalarının göz rengi sınıflandırma sonucu	59

7.32: VT2 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının göz rengi sınıflandırma sonucu.	60
7.33: VT2 için RapidMiner ve Weka uygulamalarının göz rengi tahmini için en başarılı algoritmalarının karşılaştırılması.....	60
7.34: VT2 için RapidMiner uygulamasında çalıştırılan Derin Öğrenme ve DVM algoritmalarının doğal saç rengi sınıflandırma sonucu.	61
7.35: VT2 için Weka uygulamasında çalıştırılan DVM ve LibSVM algoritmalarının doğal saç rengi sınıflandırma sonucu.	62
7.36: VT2 için RapidMiner ve Weka uygulamalarının doğal saç rengi tahmini üzerindeki en başarılı algoritmalarının karşılaştırılması.	62
7.37: VT2 için RapidMiner uygulamasında çalıştırılan LibSVM ve Rastgele Orman algoritmalarının ten rengi sınıflandırma sonucu.	63
7.38: VT2 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının ten rengi sınıflandırma sonucu.	64
7.39: VT2 için RapidMiner ve Weka uygulamalarının ten rengi tahmini üzerindeki en başarılı algoritmalarının karşılaştırılması.....	64
7.40: VT2 için RapidMiner uygulamasında çalıştırılan LibSVM ve Rastgele Orman algoritmalarının saç tipi sınıflandırma sonucu.	65
7.41: VT2 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının saç tipi sınıflandırma sonucu.....	66
7.42: VT2 için RapidMiner ve Weka uygulamalarının saç tipi tahmini için en başarılı algoritmalarının karşılaştırılması.	66
8.1: Soft biyometriklerin sınıflandırma sonuçları.	67
8.2: Literatürdeki yakın çalışmaların yapılan çalışma ile karşılaştırması.	68

EKLER DİZİNİ

Ek	Sayfa
<u>No</u>	<u>No</u>
EK 1. Etik Kurul Onay Raporu	76
EK 2. Çalışma Kapsamında Kullanılan Anket Formu.	77



SİMGELER VE KISALTMALAR DİZİNİ

cm	: Santimetre
mp	: Megapiksel

KISALTMALAR

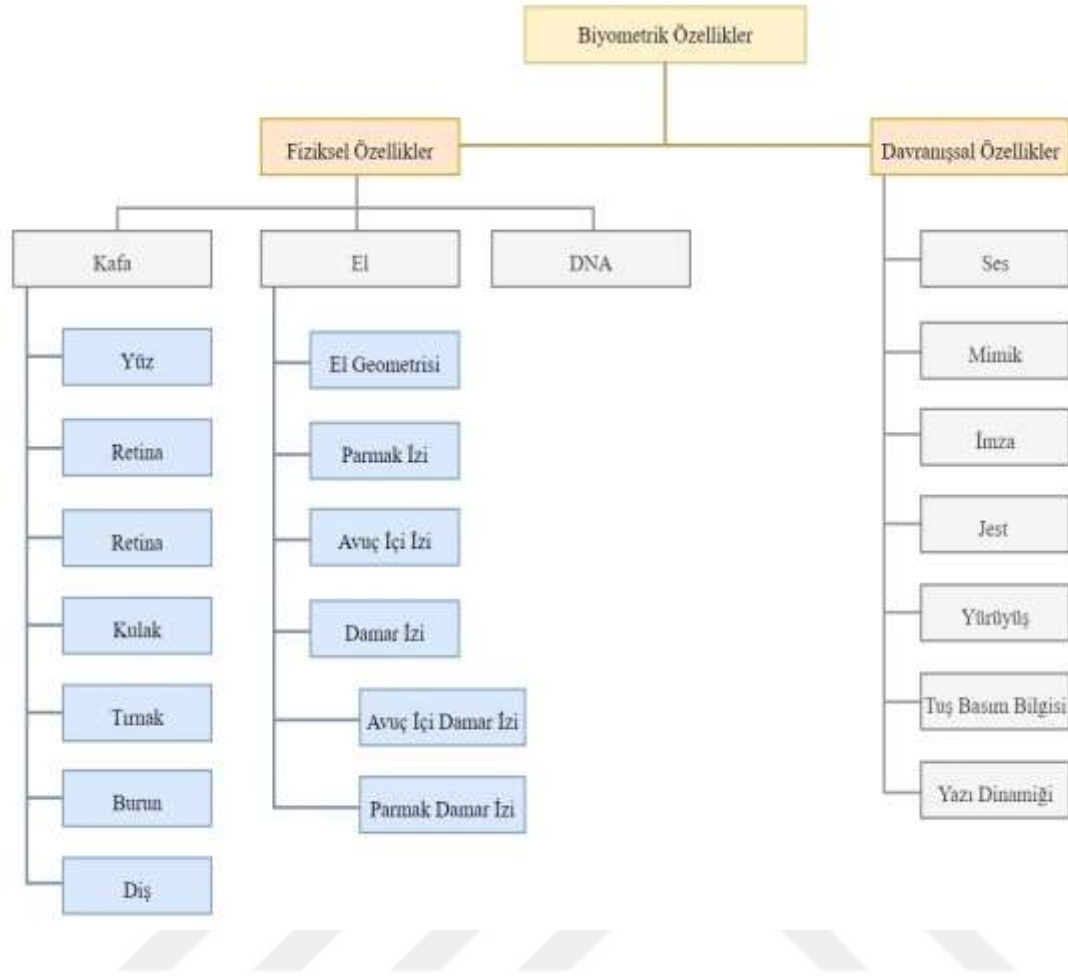
DP	: Doğru Pozitif
DN	: Doğru Negatif
DSH	: Dizinlenmiş Şekillerin Histogramı algoritması
DVM	: Destek Vektör Makineleri algoritması.
ESA	: Evrişimli Sinir Ağları.
GSKM	: Gri Seviye Kesişim Matrisi.
K-EYK	: K – En Yakın Komşu algoritması.
SYYM	: Sağlam Yerel Yönelimli Modeller.
TBA	: Temel Bileşenler Analizi algoritması.
TSA	: Tekrarlayan Sinir Ağları.
VT1	: Veri tabanı 1
VT2	: Veri tabanı 2
WPUYE	: Batı Pomerian Teknoloji Üniversitesi Kulak Veri Tabanı.
YN	: Yanlış Negatif
YP	: Yanlış Pozitif
YSA	: Yapay Sinir Ağları.

1. GİRİŞ

Teknolojinin gelişmesi ile birlikte bilgi, günümüzün en değerli kaynaklarından biri haline gelmiştir. Bilginin çok farklı alanlarda kullanılması bu değerin artmasında büyük rol oynamaktadır. Bu nedenle bilgi güvenliği günümüzde çok daha önemli bir konudur (Nieles vd., 2017). Bu verilerin güvenliğini sağlamak için şirketler büyük miktarlarda bütçe ayırmaktadır (Somepalli vd., 2020). Bu bilgilerin korunmasında geleneksel olarak şifreler kullanıldığı gibi, gelişen teknoloji sayesinde parmak izi, kulak görüntüsü, retina taraması ve DNA gibi kişilere özel biyometrik özellikler de kullanılmaya başlanmıştır. Bu özellikleri kullanmak için taşımamıza veya aklımızda tutmamıza gerek olmaması, biyometrik güvenlik yöntemlerinin anahtar ve şifre gibi geleneksel yöntemlere göre avantajlarından sadece biridir (Riaz vd., 2018). Ayrıca biyometrik bilgileri geleneksel yöntemlere göre çalmanın ve taklit etmenin daha zor olması da bu verilerin güvenlik için tercih edilmesinde rol oynayan önemli özelliklerindendir (Trikoš vd., 2019).

Biyometrik sistemlerin geliştirilmesindeki aşamalar, geleneksel yöntemlere göre farklılık göstermektedir (Sarhan vd., 2017). Biyometrik sistemlerde, geleneksel yöntemlerde kullanılan parola ve anahtar yerine retina taraması, parmak izi, kulak görüntüsü ve DNA gibi kişilere özel biyometrik özellikler kullanılır. Bu özelliklerin taklit edilebilmesi ve yeniden oluşturulabilmesi zor olduğundan geleneksel yöntemlere göre daha yüksek güvenlik sağlamaktadır (Manzoor ve Selwal, 2018).

Biyometrik sistemlerde parmak izi, retina taraması, DNA ve kulak görüntüsü gibi fiziksel biyometrik özellikler kullanıldığı gibi yürüyüş şekli, imza, ses ve el yazısı gibi davranışsal biyometrik özellikler de kullanılabilir (Boyras vd., 2019). Şekil 1.1’de temel biyometrik özellikler tablo halinde verilmiştir.



Şekil 1.1: Temel biyometrik özellikler (Ceyhan, 2012)

Şekil 1.2’de verilen biyometrik özelliklerin kullanıldığı sistemlerin artması ile birlikte, araştırmacıların bu alana olan ilgileri artmıştır. Araştırmaların artması ile biyometrik özelliklerin kendi aralarındaki ilişkiler ile ilgili çalışmalar da artmıştır. Bu çalışmalara bir örnek olarak Ceyhan vd. (2014) çalışmalarında, parmak izi ile yaş arasındaki ilişkiyi tespit etmeyi amaçlamışlardır.

Yapılan tezin amacı, literatürde varlığı kanıtlanmış kulak görüntüsü biyometrikleri ile soft biyometrik özellikler olan cinsiyet, yaş, boy, göz rengi, doğal saç rengi, ten rengi ve saç tipi arasındaki ilişkinin araştırılmasıdır. Ayrıca bu araştırma sonucunda elde edilen bilgiler ışığında, kulak görüntüsü ile ilgili soft biyometrik özelliklerin tahmin edilmesi amaçlanmıştır. Bu tezin başarılı bir şekilde geliştirilmesi ile birlikte kriminal vakalarda kullanılabilecek bir uygulama olabileceği düşünülmektedir. Kriminal vakaların yanında güvenlik alanında da diğer biyometrik tanımlamaların yetersiz olduğu durumlarda bu çalışma kullanılarak kulak görüntülerinden görüntüdeki kişiye ait soft biyometrik özellikler

tahmin edilebilir. Ayrıca bu çalışmanın MOBESE sistemlerinde diğer kişi tanımlama sistemleriyle birlikte kullanılması, sistemlerin güvenliğini artırabilir.

Tezin ikinci bölümünde, biyometrik sistemlerin temel özellikleri ve çalışma prensipleri üzerinde durulmuştur. Ayrıca biyometrik sistemler ile ilgili literatür araştırması yapılmıştır. Biyometrik sistemler ile ilgili literatür araştırmasında DNA, yüz, göz, ses, parmak izi, damar ve kulak gibi biyometrik özellikler göz önünde bulundurulmuştur. Yapılan araştırma ile birlikte bu özelliklerin kullanım alanları ve özelliklerinden bahsedilmiştir. Ek olarak tezin araştırma kapsamında kullanılan ve biyometrik veriler ile ilişkilendirilen soft biyometrik özelliklerden de bu bölümde bahsedilmiştir.

Tezin üçüncü bölümünde, kulak biyometriklerinin tanıma işlemleri ve bu işlemleri kullanan sistemler üzerinde durulmuştur. Kulak biyometriklerinin uygulama alanları üzerinde durulmuş ve bu biyometriklerin güçlü-zayıf yönlerinden bahsedilmiştir. Ayrıca bu biyometriklerin kullanılmasında kullanılan ön işleme ve öz nitelik çıkarımından ve bu özelliklerin eşleştirilmesinden bahsedilmiştir.

Tezin dördüncü bölümünde, çalışmada kullanılacak sınıflandırma algoritmaları üzerinde durulmuştur. Sınıflandırma algoritmaları ile ilgili bilgiler verilerek algoritmaların çalışma prensipleri açıklanmıştır. Ayrıca sınıflandırma algoritmalarının içerikleri ile ilgili örnek şekillere yer verilmiştir.

Tezin beşinci bölümünde, kulak görüntüsünden çıkarım yapılan çalışmalar ile ilgili literatür araştırması üzerinde durulmuştur. Bu çalışmalarda kullanılan biyometrik ve soft biyometrik özellikler verilip, bu özelliklerin ile ilgili ilişkilerin hangi yöntemler ile araştırıldığı belirtilmiştir.

Tezin altıncı bölümünde, tez için kullanılan biyometrik veri tabanlarının oluşturulması üzerinde durulmuştur. Bu bölümde, tez kapsamında farklı açılardan çekilen kulak görüntülerinin ve katılımcıların bu bağlamda doldurdukları anketin veri tabanına kayıt edilmesi işlemleri detaylı bir şekilde açıklanmıştır. Ayrıca tezde kullanılan ikinci veri tabanını için gerekli verilerin nasıl elde edildiği üzerinde durulmuştur.

Tezin yedinci bölümünde, kulak görüntüsünden soft biyometrik özelliklerin tahmini üzerinde durulmuştur. Bu bağlamda kulak verilerinin elde edilmesi, kulak görüntülerinden

ıkarımların yapılması ve bu ıkarımlar ile verilerin işlenmesi hakkında bilgiler verilmiştir. Ayrıca araştırma kapsamında ele alınan cinsiyet, yaş, boy, göz rengi, doğal saç rengi, ten rengi ve saç tipi gibi soft biyometrik özelliklerin kulak biyometrikleri ile ilgili ilişkileri ortaya koyulmuştur.

Tezin sekizinci bölümünde, yedinci bölümde açıklanan çalışma kapsamında elde edilen sonuçlar ve kazanımlar üzerinde durulmuştur. Ayrıca bu çalışmadan elde edilen bilgiler doğrultusunda, kulak biyometrikleri ile ilgi yapılacak gelecek çalışmalarda hangi konular üzerine durulmaları ve hangi konulara dikkat etmeleri konusunda tavsiyeler verilmiştir.



2. BİYOMETRİK SİSTEMLER

Biyometrik sistem olarak ele alınan sistemler, biyometrik verilerin kullanıldığı sistemler olarak açıklanabilir. Bu sistemler bireylerin fiziksel ve davranışsal özellikleri üzerinde durup kimlik belirlenmesinde kullanılmaktadır (İnel Özkiper, 2019). Bu sistemler, kopyalanması veya taklit edilmesi zor olan biyometrik özelliklerin kullanılması için geliştirilmiş sistemlerdir (Singla vd., 2021).

Biyometrik sistemlerin geleneksel sistemlere karşı avantajlarından dolayı günümüzde pek çok alanda kullanıldığı görülmektedir (Guenouni vd., 2019). Bu sistemlerin büyük bir çoğunluğunun güvenlik üzerinde durduğu söylenebilir. Biyometrik sistemlerin kullanım alanlarına örnek olarak internet bankacılığı, telefon kilit sistemleri, havaalanı işlemleri, ATM'ler, e-imza ve tez kapsamında yapılan çalışma sonucunda kullanılması planlanan bir alan olan kriminal amaçlı tespit uygulamaları örnek olarak verilebilir. Tablo 2.1'de biyometrik sistemlerin kullanıldığı alanlara örnekler verilmiştir.

Tablo 2.1: Biyometrik sistemlerin kullanım alanları (Guenouni vd., 2019)

Ticari	Devlet
ATM kullanıcı tanımlama	Suçlu tespiti
Çağrı merkezi kimlik tespiti	Hastane hasta takibi
Personel takibi	Havaalanları giriş-çıkış takibi
Sigorta şirketi kimlik tespiti	Sınır kapıları giriş-çıkış takibi
Masaüstü ve mobil bilgi güvenliği	Askeri kaynak takibi
E-ticaret işlemleri	Devlet kurumları giriş-çıkış takibi
Akıllı ev sistemleri	Kimlik uygulamaları
Kombine bilet uygulamaları	Uzaktan eğitim sınav işlemleri
Yüksek güvenli giriş çıkış takibi	Kayıp kişi tespiti

Tablo 2.1'de belirtilen biyometrik sistemlerin kullanım alanları sayısının gün geçtikçe arttığını ve gündelik yaşamda daha sık kullanıldığını söylemek mümkündür.

2.1 Biyometrik Özellikler

Biyometrik özellikler, kişilere özel ve taklit edilip kopyalanması mümkün olmayan özelliklerdir (Singla vd., 2021). Biyometrik özellikler temelde fiziksel ve davranışsal biyometrik özellikler olarak iki farklı kategoride incelenebilir. Bu özellikler Tablo 2.2’de belirtilmiştir.

Tablo 2.2: Biyometrik özelliklerin sınıflandırılması

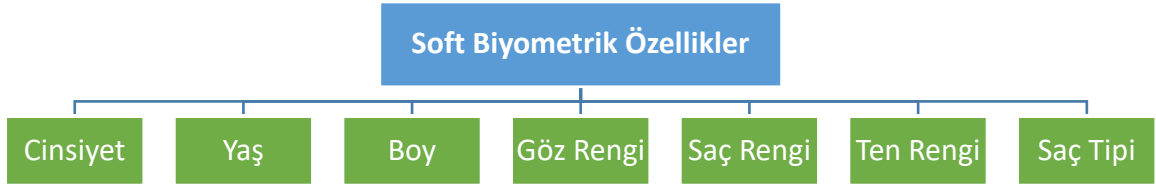
Fiziksel Biyometrik Özellikler	Davranışsal Biyometrik Özellikler
Yüz	Ses
Kulak	Yürüyüş şekli
DNA	Tuş vuruşu
El geometrisi	İmza
İris-Retina	El yazısı
Parmak izi	
Damar izleri	

Biyometrik sistemlerde kullanılan temel biyometrik özellikler Tablo 2.2’de belirtilmiştir. Biyometrik sistemler ile ilgili yapılan araştırmaların artması ile birlikte farklı biyometrik özelliklerin de bu tabloya eklenebileceği söylenebilir.

2.1.1 Soft Biyometrik Özellikler

Soft biyometrik özellikler, bireyler ile ilgili bazı bilgiler sağlayan fakat iki farklı bireyi yeterince ayırt edecek bilgiye ve kalıcılığa sahip olmayan özellikler olarak tanımlanmaktadır (Hassan vd., 2021). Bu özellikler biyometrik özelliklere göre daha az ayırt edici olsa da kişilerin tespitlerinde kullanılabilecek önemli özellikler arasında yer almaktadır.

Soft biyometrik özellikler, cinsiyete bağlı olarak değişebildiği gibi ortak özelliklerden de oluşabilmektedir. Bu özelliklerin bazılarını Şekil 2.1’de yer verilmiştir.



Şekil 2.1: Soft biyometrik özellikler

Soft biyometrik özelliklerin kendi başına ele alındığında ayırt ediciliği çok olmamasına karşın, farklı soft biyometrik özellikler birbirleri ile ilişkilendirildiğinde güçlü bağlantılar ortaya çıkabilmektedir. Örnek olarak kişiyi tanımlamak için sadece saç rengi kullanıldığında olası kişi sayısı; saç rengi, saç tipi, yaş, boy ve cinsiyet gibi soft biyometrik özellikler ile kullanıldığında sonuç kümesindeki kişi sayısı daha az olacaktır.

2.2 Biyometrik Sistemlerin Çalışma Prensipleri

Biyometrik sistemlerin çalışma prensipleri, hangi alanda uygulanacağına göre farklılık gösterebilmektedir. Sistemin uygulama alanından bağımsız olarak, biyometrik sistemlerde belirli aşamalardan bahsetmek mümkündür. Bu aşamalar temel olarak;

- İletim
- Veri toplama
- Modelleme
- Eşleştirme
- Öznitelik çıkarımı olarak verilebilir.

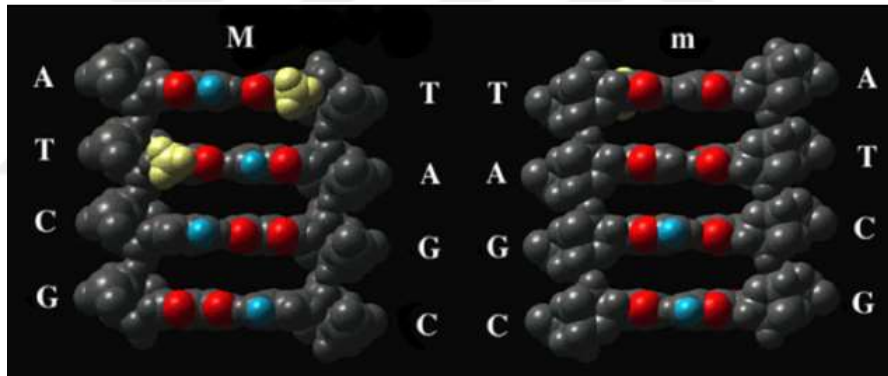
Biyometrik sistem aşamalarından veri toplamada, biyometrik sistemin kullanacağı biyometrik veriler toplanır ve kaydedilir. İletim aşaması, veri toplama aşamasından elde edilen biyometrik verilerin sisteme iletilmesidir. Bu aşamayı, iletilen verilerin anlamlı özniteliklere indirildiği öznitelik çıkarımı aşaması takip eder. Bu özniteliklerin farklı modeller olarak toplanması da modelleme aşamasında gerçekleştirilir. Bu aşamalar gerçekleştikten sonra, sisteme araştırılmak için verilen veri ile veri tabanındaki verinin karşılaştırıldığı eşleştirme aşaması yapılarak sistem gerekli çıktıyı üretir.

2.3 Yaygın Kullanılan Biyometrik Özellikler

Teknolojinin gelişmesiyle birlikte farklı biyometrik özelliklerin farklı alanlarda kullanımı artmaktadır. Bu biyometrik özelliklerden en yaygın olarak kullanılanlarının DNA, yüz, retina/iris, ses, parmak izi, damar ve kulak olduğunu söylemek mümkündür.

2.3.1 DNA Tanıma

DNA tanıma, gen ve moleköl teknolojisinin gelişmesi ile birlikte günümüzdeki en güvenilir kimlik doğrulama yaklaşımlarından biri haline gelmiştir (Namasudra vd., 2020). Bu yöntem sık olarak kriminal vakalarda kullanılmasının yanı sıra, babalık davaları gibi adli vakalarda da sıkça kullanılmaktadır (Çetli vd., 2019). Bu yöntemde hücre içerisinde bulunan ve DNA moleküllerindeki nükleotitlerin dizilimlerinden yararlanılmaktadır. DNA ve nükleotit dizilimleri Şekil 2.2’de görülebilir.



Şekil 2.2: DNA ve içeriğindeki nükleotitler (Travers ve Muskhelishvili, 2015)

DNA tanıma işleminde, kişinin biyolojik materyallerindeki DNA’dan yararlanılır. DNA moleküllerindeki nükleotitler adenin, sitozin, guain ve timin nükleobazlarından oluşur (Pramanik vd., 2018). Bu nükleobazların birlikte oluşturduğu dizilimlere göre genetik bilgiler kodlanır. Kodlanan bilgiler birlikte incelendiğinde benzersiz bir dizilim görülür. DNA tanıma işleminde de bu benzersiz dizilimlerin bilgisinden yararlanılır.

DNA tanıma sisteminin çok güvenilir bir biyometrik sistem olmasına karşın dezavantajları da vardır. Bu dezavantajlardan en büyüğü, tanıma için kullanılan biyolojik materyalin temiz ve yeni olması gerekmektedir. Bu nedenle DNA tanıma işlemleri genellikle maliyetlidir ve yüksek bütçeli laboratuvarlarda gerçekleştirilir (Smith, 2018).

2.3.2 Yüz Tanıma

Yüz tanıma sistemleri, günümüzde her alanda kullanılmakta olan güvenlik sistemi türüdür. Akıllı telefonlarda, bilgisayarlarda ve kamera içeren bütün sistemlerde, yüz tanıma sistemleri güvenlik sistemi olarak kullanılabilir (Aksoy, 2019). Yüz tanıma sistemlerinin yapısı Şekil 2.3’de verilmiştir.



Şekil 2.3: Yüz tanıma sistemlerinin yapısı. (Kortli vd., 2020)

Yüz tanıma sistemleri ile ilgili ilk çalışmalar 1960 ve 1970’li yıllarda başlamıştır. Bu yıllara yüz tanıma sistemleri için ilk aşama yılları da demek mümkündür. Bu yıllarda yüz tanıma ile ilgili çalışmalara, yüz tanıma sistemlerinin kurucusu olarak kabul edilen Woody Bledsoe öncülük etmiştir (Yang vd., 2018). Woody çalışmalarında, yüz tanımda kullanılmak için gerekli olan yüz hatları üzerinde durmuştur.

Yüz tanıma sistemlerinde, yüzün özellik çıkarımı bölgelerinden referans noktaları alınır. Bu noktaların konumları veri tabanına kaydedilerek, daha sonraki yüz tanımlarda eşleşme kontrolü için kullanılır.

Yüz tanıma sistemleri sık olarak kullanılmasının yanında, bu sistemlerin dezavantajlarının da olduğu göz önünde bulundurulmalıdır. Bu dezavantajlardan en önemlileri şunlardır (Aksoy, 2019):

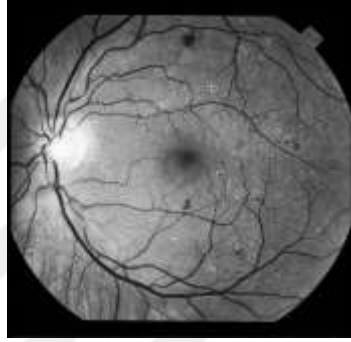
- Yüz tanıma işlemi, görüntüler üzerinden yapıldığından hassas bir sınıflandırma yapılması gerekir.
- Yüz tanıma işlemlerinde kullanılacak görüntülerin elde edilmesinde kullanılan cihaz, ortamın ışığı, poz ayarı ve renk ayarı gibi dış etmenlerin doğru ve tutarlı olmaması, sınıflandırmanın yanlış sonuçlanmasına neden olabilir.
- Yüz tanıma işleminde, yüzün referans noktaları kaydedildiğinden, yüzde kaza vb.

etkenler ile deęişimin gerekleşmesinde, yüz tanıma işleminin başarısı azalmaktadır.

- Yüz tanımda kullanılan özellik noktalarının her birey için ayrı saklanması, büyük bir veri kümesi oluşmasını doğurmaktadır. Bu nedenle yüz tanıma işleminin gerekleştirildięi veri tabanının büyümesi, sistemin doğruluk oranını arttırabilmesine karşın sistemin hızının azalmasına da yola açabilmektedir.

2.3.3 Retina ve İris Tanıma

Retina ve iris tanıma, günümüzde sıkça kullanılan ve güvenilirlięi yüksek kabul edilen biyometrik sistem türlerindendir. Retina örneęi Şekil 2.4’de görülebilir.



Şekil 2.4: Retina aęı. (Manav, 2010)

Retina, gözümüzde bulunan ve üzerinde gelen ışınların beynimize gönderilmesinde görev alan bir aę yapısıdır. Bu aę yapının görüntüsü, retina tabanlı göz tanıma sistemlerinin temelini oluşturur. Bu taramanın, aę tabakasının görüntüsü üzerinden yapılmasının sonucu olarak gözlük takılması veya hareket edilmesi durumunda sistemin yanlış sonuçlar üretebileceęi söylenebilir (Manav, 2010).

İris, gözümüzün renkli kısımlarını oluşturan bölümdür. İris tabanlı göz tanıma sistemlerinde, irisin görüntüsü üzerinden belirli referans noktaları belirlenir. Bu yöntem kulak ve parmak izi yöntemlerine benzemektedir. Bu sistemin en önemli avantajı, irisin her kişide farklı olmasıdır. Bu nedenle ayırt edici özellikler arasında yer almaktadır. Bu sistemlerde karşılaşılabilecek en büyük dezavantaj, referans görüntüsü alınırken sabit bir noktaya bakılmasının gerekliliğidir. Aksi takdirde, daha sonra karşılaştırılma yapılacak görüntünün farklı noktaya bakılarak çekilmesi ile karşılaştırmının zorlaştırılacağı öngörülebilir (Manav, 2010).

2.3.4 Ses Tanıma

Ses tanıma sistemleri, kişilerin seslerine bağlı olarak çalışan sistemlerdir. Bu sistemlerde sesin kullanıldığı farklı yöntemler kullanılmaktadır. Temelde metin bağımlı ve metin bağımsız olarak ayrılan bu yöntemler, alt dal olarak 4 farklı yöntemde ele alınabilir. Bu yöntemler şunlardır (Çakır, 2017):

- **Temel Alınan Ses Birimine Göre Konuşma Tanıma:** Bu yöntemde, seslerin kelimelere ve hecelere göre ele alınması kullanılır. Yöntemde ses tanıma için kelimeler ve heceler arası boşluklar ve süreler hesaplanır ve karşılaştırma yapılır.
- **Konuşmacıya Göre Konuşma Tanıma:** Bu yöntemde, konuşma yapan kullanıcı ile daha önce sistemde kayıtlı olan sesler karşılaştırılır. Karşılaştırma için verilen belirli kelimeler konuşmacıya okutulur ve aynı kelimelerin okunduğu kayıtlar ile karşılaştırılır.
- **Sesin Sürekliliğine Göre Konuşma Tanıma:** Bu yöntemde, okuyucuya verilen kelime dizilerinin okunmasındaki belirli kelimelerin tespiti ile gerçekleştirilir.
- **Metne Göre Konuşma Tanıma:** Bu yöntemde kullanılan tanımlama metne bağlıdır. Bu nedenle kullanılan eğitim seti metne bağlı olarak oluşturulur.

Ses tanıma sistemleri, insanların seslerine bağlı olarak çalışmasından dolayı örneğin hastalık gibi sesin üretilmesine etki eden bir durumda kullanımı zordur. Ayrıca tüm ses dosyalarının da kaydedilmesi gerektiğinden veri tabanı maliyeti de yüksektir.

2.3.5 Parmak İzi Tanıma

Parmak izi tanıma sistemleri, günümüzde en sık kullanılan güvenlik ve tanıma sistemlerinden biridir. Akıllı telefonlarda, kriminal vakalarda ve personel takip sistemlerinde parmak izi tanıma sistemleri sıkça kullanılmaktadır. Parmak izi örneği Şekil 2.5’de görülebilir.



Şekil 2.5: Parmak izi örneği. (Ceyhan, 2012)

Parmak izi taramada, parmak izinin görüntüsünde belirlenen farklı alt özellik noktaları kullanılır. Bu özellik noktaları vadiler ve tepelerden oluşmaktadır. Parmak izi görüntülerindeki bu vadilerin ve tepelerin konum noktaları kaydedilir ve daha sonraki tanıma işlemlerinde karşılaştırma için kullanılır. Bu noktalara öznitelik noktaları denilmektedir. Yeterli sayıda benzer öznitelik noktası bulunduğunda parmak izi karşılaştırması tamamlanmış olur (Xia vd., 2017).

Parmak izi sistemleri otomatik ve yarı otomatik olmak üzere iki temel kategoride incelenebilir. Otomatik sistemlerde parmak izinden öznitelik noktalarının çıkarımı otomatik olarak gerçekleşir. Yarı otomatik sistemde ise bu işlem formüller ile gerçekleşir. Parmak izi sistemlerinde genellikle otomatik sistemler kullanılmaktadır (Deshpande ve Malemath, 2022).

Parmak izi sistemleri, parmak izlerinin ayırt edici özellik olmasından dolayı yüksek doğruluğa sahip tanıma sistemleri arasında yer almaktadır. Bu nedenle yüksek güvenilirlik gereken alanlarda sıkça kullanılmaktadır.

2.3.6 Damar Tanıma

Damar tanıma sistemleri, göz tanıma sistemlerinde kullanılan retina tarama sistemleri gibi damarların görüntüleri üzerinden yapılan karşılaştırma ile kullanılır. Bu yöntemde parmak veya el gibi vücudun sabit bir bölümündeki damar haritasının görüntüsü kullanılır. El damar haritasının görüntüsünün kişilere özgü olduğu kanıtlandığından, sık kullanılsa da ayırt edici bir tanıma sistemi olduğu söylenebilir (Wang ve Wang, 2017). El damar ağı örneği Şekil 2.6'da görülebilir.

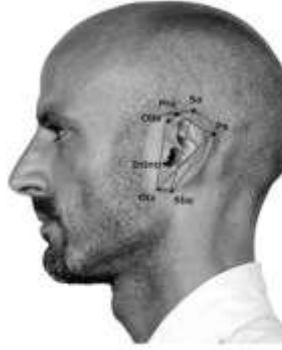


Şekil 2.6: El damar ağı (Wang ve Wang, 2017)

Damar tanıma sistemlerinin sık kullanılmamasının başlıca sebebi, kullanıcının kilo değişimlerinde damar yapısında değişim meydana gelebilmesidir. Ayrıca damar görüntüleri kızıl ötesi olarak kaydedildiğinden depolaması zordur.

2.3.7 Kulak Tanıma

Kulak tanıma sistemleri, günümüzde popülerliği artmakta olan biyometrik tanıma sistemleri arasındadır. Kulak görüntülerinin, kişiler arasında ayırt edici özellik olarak kullanıldığı farklı çalışmalar doğrultusunda kanıtlanmıştır (Karauslu vd., 2022).



Şekil 2.7: Kulak tanımada kullanılan öznitelik noktaları (Yaman vd., 2018)

Kulak tanımada, parmak izi tanımada olduğu gibi kulak görüntüsünü üzerinde bulunan öznitelik noktaları belirlenir. Bu noktaların konumu ve aralarındaki ilişki göz önünde bulundurularak sistemdeki kayıtlı noktalar ile karşılaştırma yapılır. Kulak tanımada kullanılan öznitelik noktaları Şekil 2.7’de görülebilir.

Kulak tanıma sistemlerinde farklı yöntemler kullanılsa da en çok kullanılan yöntemlerin derin öğrenme yöntemleri olduğunu söylemek mümkündür (Dodge vd., 2018). Derin

öğrenme yöntemleri arasında da en çok Evrişimli Sinir Ağları tercih edilmektedir. Evrişimli Sinir Ağları sayesinde kulak görüntülerindeki öznitelik noktalarının ayırt edilmesi gerçekleşir. Bu ayırt edici öznitelik noktalarının birbirleriyle ilişkilendirilmesi ve veri tabanında daha önceden elde edilen veriler ile karşılaştırılması ile kulak tanıma işlemi gerçekleştirilir (Karauslu vd., 2022).

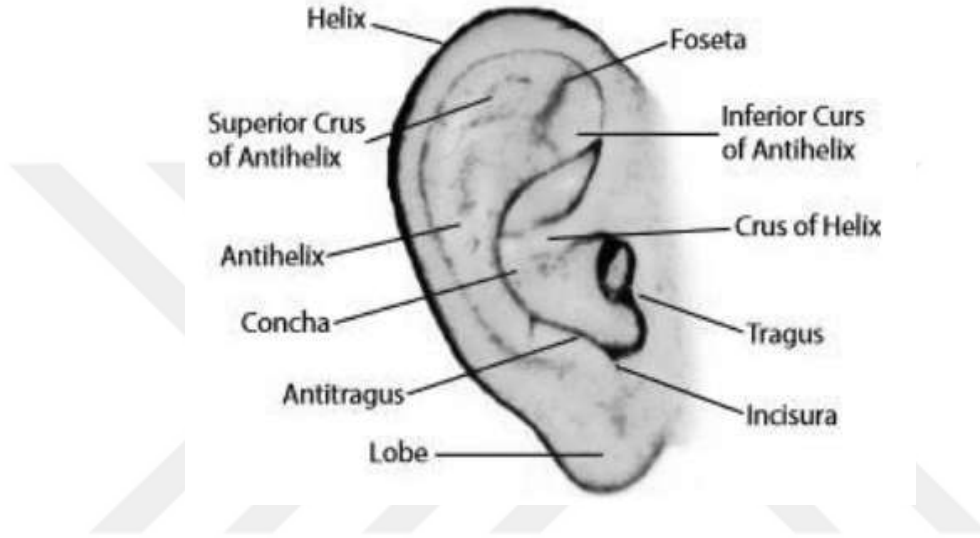
Kulak tanıma sistemlerinin güvenilir sistemler arasının olmasının yanında, dezavantajları da vardır. Bu dezavantajlar şunlardır:

- Veri tabanının büyüklüğü, sistemin doğruluğunu büyük ölçüde etkileyebilmektedir.
- Kulak tanımada kullanılacak kulak görüntülerinin çekimi özenli olarak yapılmalıdır. Çekim yapılan ortamdaki ışık, poz ayarı, renk ayarı ve uzaklık gibi etkenlerin değişmesi, kulak tanıma işlemlerinde farklı sonuçlar alınabilmesini doğurabilmektedir.

Bu dezavantajlar göz önünde bulundurulduğunda, kontrollü koşullarda kulak tanıma sistemlerinin yüksek doğruluk oranlarında çalışabileceği öngörülebilir.

3- KULAK TANIMA SİSTEMLERİ VE İŞLEMLERİ

Kulak biyometriği, günümüzde yaygınlaşmaya başlayan biyometrik özellikler arasında yer almaktadır. Kulak biyometriği ile ilgili yapılan çalışmaların artması ile kulak tanıma sistemlerinin günlük hayatta kullanımının yaygınlaştığı görülmektedir (Galdámez vd., 2017).



Şekil 3.1: Dış kulak anatomisi. (Kavipriya ve Muthukumar, 2019)

Kulak tanıma sistemlerinde dış kulak anatomisi kullanılmaktadır. Bu bölümler Şekil 3.1’de helix, foseta, antihelix, superior crus of antihelix, crus of helix, inferior crus of helix, concha, antitragus, tragus ve incisura olmak üzere gösterilmiştir. Bu bölümler kulak tanıma sistemlerinde öznitelik noktalarının belirlenmesinde kullanılmaktadır (Kavipriya ve Muthukumar, 2019).

Kulak tanıma sistemlerinde temel olarak, kulağın dış görüntüsünde bulunan öznitelik noktaları kullanılır. Bu öznitelik noktalarının konumu ve aralarındaki uzaklıklar gibi özellikler kullanılarak karşılaştırma yapılır. Bu karşılaştırma sonucunda eşleşmesi en yüksek olan kulak görüntüsü seçilir ve kulak tanıma tamamlanmış olur (Kavipriya ve Muthukumar, 2019).

3.1 Kulak Biyometriğinin Uygulama Alanları

Kulak biyometriği, gelişen teknoloji ile günümüzde pek çok alanda kullanılmaya başlanmıştır. Kulak biyometriğinin kişilere özel olması ve taklidinin zor olması sebebiyle güvenilir bir tanıma yöntemi olarak görülmektedir.

Kulak biyometriğinin kullanıldığı ve kullanılması uygun alanlarına örnek olarak;

- Sistemlerin kimlik doğrulama işlemleri,
- Akıllı ev sistemleri,
- Bina giriş ve yoklama işlemleri,
- Havaalanı gibi seyahat yerlerinde kontrol işlemi,
- Kriminal vakalarda kimlik tespiti,
- Kimlik kartı, pasaport ve sürücü belgesi gibi devlet uygulamaları,
- Doğrulama gerektiren finansal işlemler

gibi çeşitli alanlarda kulak biyometriğinin kullanımının uygulanabileceği söylenebilir.

Kulak biyometriği tek başına kullanılabildiği gibi farklı biyometrik tanıma sistemlerine yardımcı olarak da kullanılabilmektedir. Örnek olarak kriminal vakalarda suçlunun kimliğinin tanınmasına kullanılacak yüz tanıma sisteminin başarısı yetersiz olabilir. Böyle bir durumda, araştırmanın konusu olan kulak biyometriğinden soft biyometrik özelliklerin çıkarımı kullanılarak, yüz tanıma sistemine yardımcı olarak kriminal vakalarda suçlunun kimliğini tespit edilmesi sağlanabilir.

Bir önceki örnekteki gibi; kimlik doğrulama sistemlerinde, bina giriş ve yoklama işlemlerinde, havaalanı gibi seyahat yerlerinde ses ve yüz tanıma gibi biyometrik sistemlere yardımcı olarak kulak tanıma sistemlerinin kullanılmasının kimlik tespiti ve tanıma gibi işlemlerin doğruluğunu arttıracığı söylenebilir.

3.2 Kulak Biyometriğinin Güçlü ve Zayıf Yönleri

Kulak biyometriğinin kullanıldığı sistemlerin de diğer biyometrik sistemler gibi zayıf ve güçlü yönleri bulunmaktadır. Bu zayıf ve güçlü yönlerin bilinmesi, sistemin kullanılacağı alan açısından son derece önem arz etmektedir. Kulak biyometrisinin kullanıldığı sistemlerde güçlü yönlerin ağır basmasının, sistemin güvenilirliğini ve doğruluğunu artıracakı öngörülebilmektedir.

Kulak biyometriğinin, biyometrik sistemlerde kullanılmasında göz önünde bulundurulacak güçlü yönleri (Meng vd.,2021; Wang vd., 2021);

- 8 ile 70 yaş arasında kulak şeklinin değişmemesi; dolayısıyla kulak biyometrikleri ile tanımanın bu yaşlar arasında mümkün olması,
- Kulağın önüne gelen saçın kızılötesi çekim ile göz ardı edilebilmesi,
- Topluma açık alanlarda kullanıldığında parmak izi gibi biyometrik sistemlere göre temassız gerçekleştirilebilmesi ve bu nedenle daha hijyenik olması,
- Kulak görüntülerindeki arka planın kulak tanımadan bağımsız olması ve dolayısıyla farklı ortamlardan elde edilen görüntüler ile kolaylıkla çalışabilmesi,
- Cep telefonu kamerası gibi gündelik kullanılan araçlar yardımı ile elde edilebilir olması,
- Kulak şeklinin mimikler ile çok az değişime uğraması gibi önemli noktalarda belirtilebilir.

Kulak biyometriğinin yukarıda belirtilen güçlü yönlerine karşın zayıf yönleri de bulunmaktadır. Bu zayıf yönler ise (Meng vd.,2021; Wang vd., 2021);

- Kulak tanıma kullanılacak fotoğrafların çekildiği ortamdaki ışık, pozlama mesafesi, renk ayarı gibi dış faktörlerinin önemli rol oynaması,
- Kulak tanıma işlemi görüntü üzerinden yapıldığından görüntü kalitesinin önem arz etmesi,
- Kulak tanıma işlemlerinde kullanılmak üzere depolanmış kulak görüntülerinin veri tabanında büyük bir yer kaplaması ve bu nedenle büyük veri tabanları içeren sistemlerde çalışma hızlarında veri sayısı ile ters orantılı olarak yavaşlama yaşanması olarak belirtilebilir. Kulak biyometriklerinin kullanılması planlanan sistemlerde bu

güçlü ve zayıf yönlerin doğru olarak ele alınmasının, sistemin başarılı bir şekilde çalışması açısından önemli bir role sahip olduğunu söylemek mümkündür.

3.3 Ön İşleme ve Öznitelik Çıkarımı

Kulak biyometriğinin kullanıldığı tanıma sistemlerinde kulak görüntülerinde yararlanılmaktadır. Bu kulak görüntülerinin tanıma işlemlerinde kullanılabilmesi için ön işlemeye tabii tutulması gerekmektedir (Dodge vd., 2018).

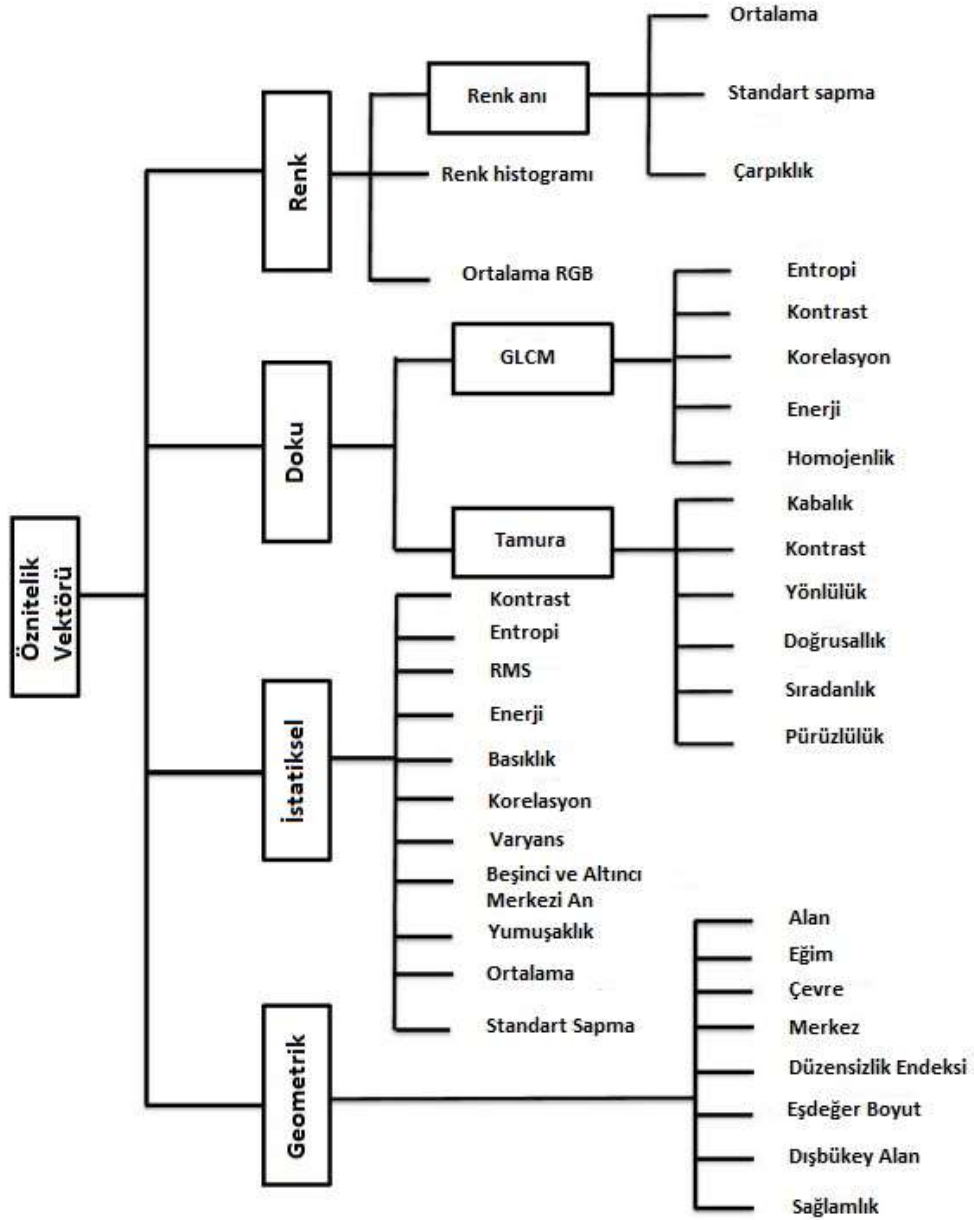
Ön işlemenin temel amacı, kulak tanımada kullanılacak görüntüdeki özniteliklerin daha doğru elde edilebilmesidir. Kulak tanıma kapsamındaki öznitelikler, dış kulakta bulunan farklı bölgeler olarak tanımlanmaktadır. Bu bölgelerin doğru ve kararlı bir şekilde belirlenmesi, tanıma işleminin doğruluğunun artmasını sağlayacaktır.

Kulak tanıma işlemlerinde ön işleme için farklı iyileştirmeler kullanılmaktadır. Bu iyileştirmelerden bazıları şunlardır (Kaçar, 2019):

- Bulanıklık giderme,
- Gürültü temizleme,
- Kontrast ayarı,
- Yumuşatma,
- Görüntü yoğunluğu ayarı,
- Çözünürlük artırma,
- Histogram eşitleme.

Yukarıda belirtilen iyileştirmelerden bir veya birkaçının ön işlemede kullanılmasıyla elde edilen görüntülerin öznitelik çıkarımı için daha uygun bir hale getirilmesi mümkündür.

Kulak tanımada kullanılan özniteliklerin çıkarımında farklı yöntemlerden yararlanılmaktadır. Bu yöntemlerde temel olarak öznitelik vektörleri kullanılmaktadır. Bu vektörlerin nasıl elde edildiği de yöntemlerin temel ayrımını oluşturmaktadır.

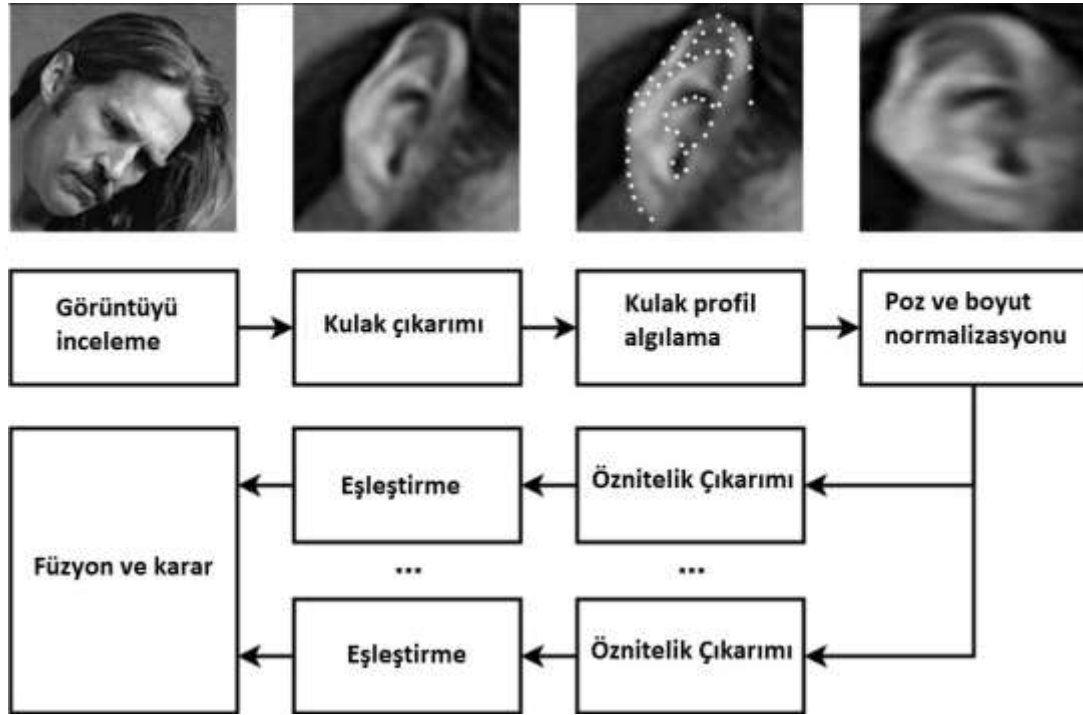


Şekil 3.2: Öznitelik çıkarım yöntemleri. (Mutlag vd., 2020)

Öznitelik çıkarımında kullanılan temel yöntemler Şekil 3.2’de verilmiştir. Bu yöntemler temel olarak geometrik, istatiksel, doku ve renk olmak üzere dört ana başlıkta incelenebilmektedir. Bu başlıklarına altında kullanılan yöntemlerde dikkate alınan hususların benzerlik gösterdiğini görmek mümkündür. Elde edilen veriler göz önünde bulundurularak bu yöntemlerden bir veya birden fazlasının kullanılmasıyla öznitelik çıkarımı işlemleri gerçekleştirilmektedir.

3.4 Eşleştirme

Kulak tanıma işleminde kullanılacak görüntüler, ön işleme ve öz nitelik çıkarımı işlemlerinden sonra eşleştirmeye tabii tutulmaktadır. Bu işlem öz nitelikler arasındaki bağlantılardan elde edilen verilerin veri tabanındaki veriler ile karşılaştırılması yoluyla gerçekleştirilmektedir (Yaman vd., 2018).



Şekil 3.3: Kulak tanıma işlemi. (Hansley vd., 2018)

Kulak tanıma işleminin temel adımları Şekil 3.3’de verilmiştir. Şekilde görüldüğü gibi ön işleme ve öz nitelik çıkarımı sonrasında eşleme işlemi gerçekleştirilmektedir. Eşleştirme işleminin sonucuna göre de diğer eşleştirmeler işe birleştirme ve karar verme gerçekleştirilmektedir. Karar verme sonucunda da doğruluk oranları göz önünde bulundurularak bir sınıflandırma yapılır ve kulak tanıma işlemi gerçekleşmiş olur.

4- SINIFLANDIRMA ALGORİTMALARI

Sınıflandırma algoritmaları, herhangi bir veri seti içerisindeki herhangi bir veri sınıfının, veri sınıfı bilinmeyen yeni bir veri ile sınıflandırılma gerçekleştirilmesi için kullanılan algoritmalarlardır. Bu algoritmalar arasında en sık kullanılanları arasında Destek Vektör Makineleri, Derin Öğrenme, Lojistik Regresyon, Doğrusal Regresyon, LibSVM, Rastgele Orman ve Naive Bayes algoritmaları yer almaktadır.

4.1 Destek Vektör Makineleri Algoritması

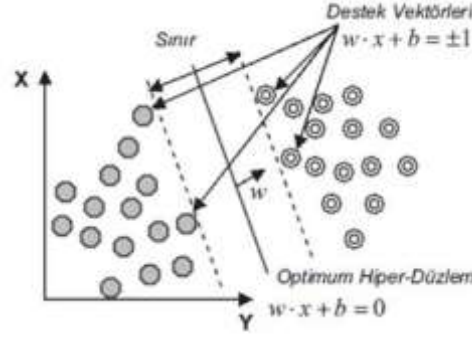
Destek Vektör Makineleri algoritması, risk azaltma yöntemine dayalı çalışan bir sınıflandırma algoritması türüdür. Bu algoritma türü, sınıflandırma ve veri bilgilerindeki örüntülerin tanınması için Vladimir Vapnik ve Alexey Chervonenkis tarafından geliştirilmiştir (Cortes ve Vapnik (1995)). Destek Vektör Makinesi algoritması herhangi bir dağılım içeren fonksiyona gerek duymadan çalışabildiği için, bu algoritma bağımsız öğrenme algoritmaları arasında yer almaktadır (Ayhan ve Erdoğan, 2014).

Destek Vektör Makineleri algoritmasında, iki boyutlu uzayda bulunan veriler doğrusal olarak ayrılmaktadır. Yapılan doğrusal ayrılma sonrasında elde edilen iki farklı veri kümesinin, ayrılan doğruya paralel olarak doğruya en yakın doğrular destek vektörleri olarak adlandırılmaktadır (Küçüksille ve Ateş, 2013). Doğrusal ayrılma formülü Şekil 4.1’de verilmiştir.

$$f(\vec{x}) = w^T \vec{x} + b$$

Şekil 4.1: Doğrusal ayırma formülü (Küçüksille ve Ateş, 2013)

Şekil 4.1’de doğrusal ayrılma formülü verilmiştir. Formülde bulunan W^T ibaresi karar fonksiyonunda bulunan normal, b ifadesi doğrunun eğilim verisini ve x ise destek vektörleri üzerinde bulunan noktaları temsil etmektedir (Küçüksille ve Ateş, 2013). Doğrusal ayırma ve destek vektörleri Şekil 4.2’de örnek olarak verilmiştir.

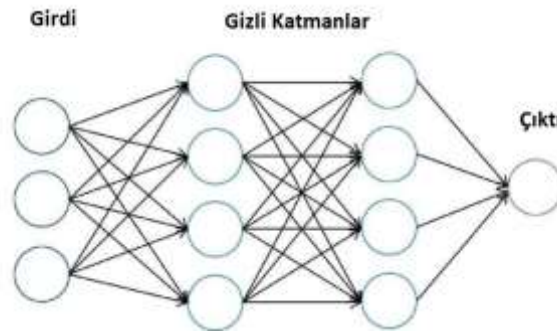


Şekil 4.2: Doğrusal ayırma ve destek vektörleri (Küçüksille ve Ateş, 2013).

4.2 Derin Öğrenme Algoritması

Derin öğrenme algoritması, insan beyninin taklit edilmesi ile gerçekleştirilen bir makine öğrenmesi sınıflandırma algoritmasıdır. Derin öğrenme algoritmaları içerisinde, yapay sinir ağlarının kullanımı sıkça gerçekleşmektedir. Ancak yapay sinir ağlarından yararlanmadan da derin öğrenme algoritmaları sınıflandırma yapabilmektedir (Arunadevi vd., 2018).

Çalışmada kullanılan derin öğrenme algoritması RapidMiner uygulaması üzerinde H2O tarafından geliştirilen ileri beslemeli yapay sinir ağları mantığı ile çalışmaktadır. Bu sinir ağlarının normal sinir ağlarından farklı ise düğümler arasındaki bağlantıların girdi katmanından çıktı katmanına doğru tek yönlü ilerlemesidir. Bu algoritma genellikle derin sinir ağları ile birlikte kullanılmaktadır. Derin öğrenme algoritmasının derin sinir ağları içerisindeki Olasılıksal Dereceli Azalma ile birlikte kullanılması sayesinde model eğitiminin daha başarılı bir şekilde gerçekleşmesi sağlanabilmektedir (Deng vd. 2013). Derin Öğrenme algoritmasında kullanılan ileri beslemeli yapay sinir ağları yapısına ait bir örnek Şekil 4.3’de verilmiştir.



Şekil 4.3: Örnek bir ileri beslemeli yapay sinir ağı yapısı (Arı ve Berberler, 2017)

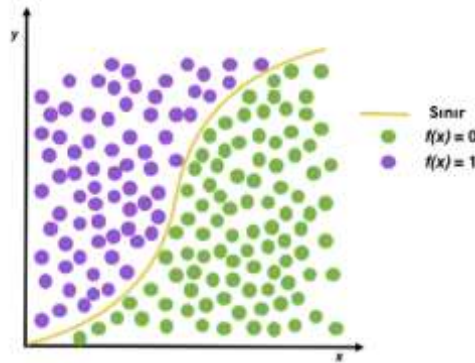
4.3 Lojistik Regresyon Algoritması

Lojistik Regresyon algoritması, herhangi bir varsayım gerektirmeden regresyon modelleri kurulmasını sağlayan bir algoritmadır. Ayrıca bu algoritma, algoritmaya verilen değişkenliklerin sürekli veya süreksiz olmasını önemsemeyen bir algoritmadır. Lojistik regresyona benzer algoritmalar bulunmasına karşın, kaç adet grup olduğu gibi farklı değişkenlerden ötürü diğer regresyon algoritmalarından farklıdır (Şenel ve Alatl, 2014). Bu algoritmada veri ayrımı çizgisi için en küçük kareler yöntemi ve sigmoid fonksiyonu kullanılmaktadır. Lojistik regresyon için kullanılan formül Şekil 4.4’de verilmiştir. Bu formül ile herhangi iki sınıftan biri için, verinin o sınıfa ait olma olasılığı hesaplanmaktadır (Dreiseitl ve Ohno-Machado, 2002).

$$P(1|x, \alpha) = \frac{1}{1 + e^{-(\alpha \cdot x)}}$$

Şekil 4.4: Lojistik regresyon formülü (Dreiseitl ve Ohno-Machado, 2002).

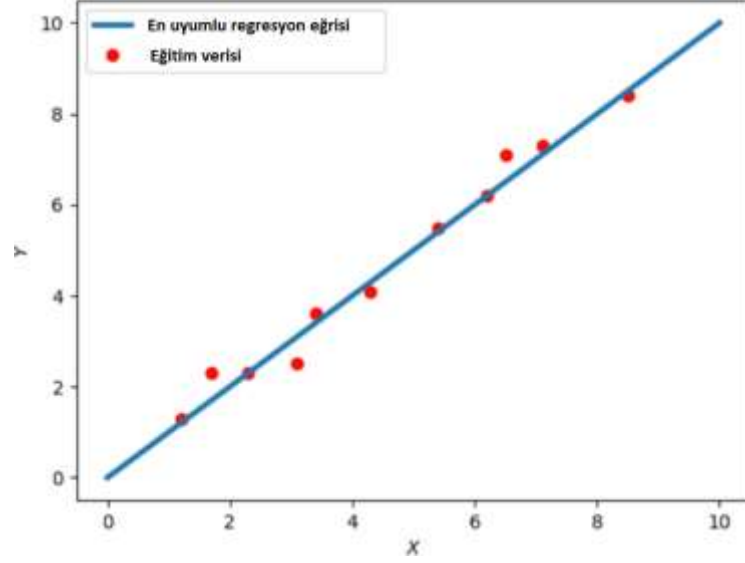
Lojistik Regresyon algoritmasının sonuçları veri tablosu olarak verilebildiği gibi aynı zamanda grafik olarak da verilebilmektedir. Lojistik regresyon algoritmasına ve sigmoid fonksiyonuna ait bir örnek Şekil 4.5’de verilmiştir.



Şekil 4.5: Örnek bir lojistik regresyon grafiği (Panesar vd., 2019).

4.4 Doğrusal Regresyon Algoritması

Doğrusal regresyon, her ne kadar regresyon algoritması olsa da lojistik algoritmalarından farklılıklar göstermektedir. Doğrusal regresyon algoritması içerisinde yapılan analizin amacı, belirli bir veri seti kullanılarak yeni verilerin değişkenlerinin tahmin edilmesini sağlamaktır (Kılıç, 2013). Örnek bir doğrusal regresyon grafiği Şekil 4.6’da görülmektedir.



Şekil 4.6: Örnek bir doğrusal regresyon grafiği (Potok, 2021).

Doğrusal regresyondaki değişkenlerin durumu, doğrusal regresyon modelini belirlemektedir. Doğrusal regresyondaki bağımsız değişkenler bir adet ise bu regresyon basit doğrusal regresyon olarak adlandırılmaktadır. Aynı değişkenlerin birden fazla olması durumunda ise çoklu doğrusal regresyon modeli kullanılmaktadır. Bu algoritmada veri ayrımı çizgisi için maksimum olabilirlik yöntemi kullanılmaktadır. Çoklu doğrusal regresyon modeli için bir örnek formül Şekil 4.7’de verilmiştir. Verilen formülde β değeri değişken katsayılarını belirtirken ε değeri ise regresyon sırasında oluşabilecek tesadüfi hata değerini belirtmektedir. (Kılıç, 2013).

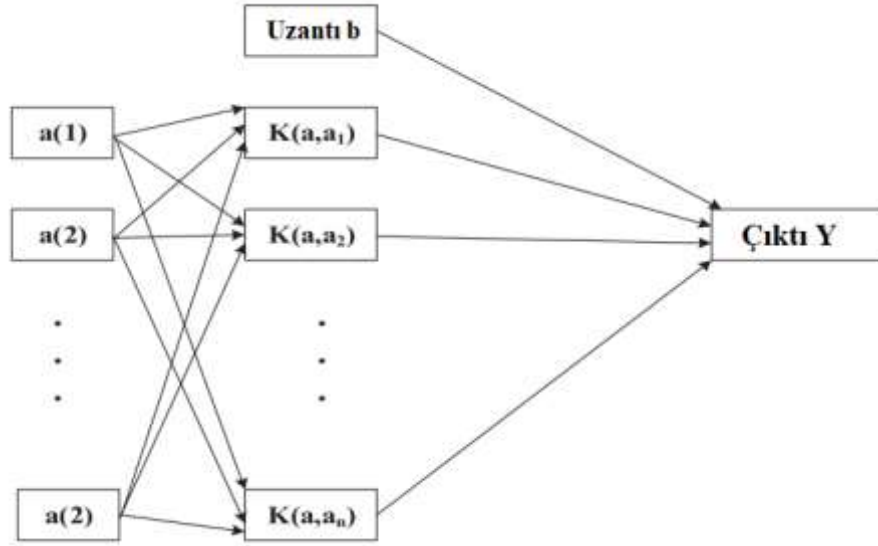
$$Y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_{ip} + \varepsilon_i$$

Şekil 4.7: Örnek bir çoklu doğrusal regresyon formülü (Kılıç, 2013).

4.5 LibSVM Algoritması

LibSVM kütüphane algoritması, Chang ve Lin (2011) tarafından doktora döneminde oluşturulmuş, sınıflandırma algoritmalarından biri olan Destek Vektör Makineleri algoritması için kütüphaneler bulunduran bir kütüphanedir. Ayrıca bu kütüphane sınıflandırma algoritması olarak da kullanılabilir. Bu algoritmanın Destek Vektör Makineleri algoritmasından en büyük farkı, çoklu veri türlerinin sınıflandırılması için farklı işlemlere gerek duymamasıdır (Bektaş vd., 2016). LibSVM kütüphanesi içerisinde C-SVM, nu-SVM ve tek sınıf DVM algoritmaları kullanılmaktadır.

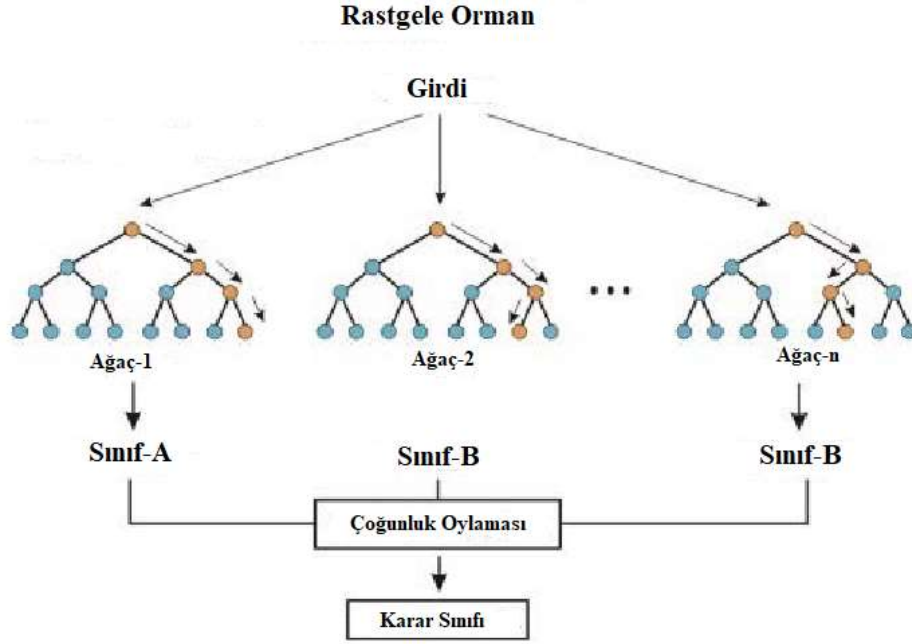
LibSVM algoritması, çoğu zaman iki farklı sınıf ile sınıflandırma için, normalizasyon sonrasındaki negatif ve pozitif değerlerin bir hiper düzlemde olabildiğince ayırtırmak üzerine çalışmaktadır (Chen vd., 2020). Ayrıca bu algoritma, birden çok girdi ve tek bir çıktı alınması prensibinde çalışan bir mimariye sahiptir. Örnek LibSVM mimarisi Şekil 4.8’de görülebilmektedir.



Şekil 4.8: Örnek bir LibSVM mimarisi (Chen vd., 2020).

4.6 Rastgele Orman Algoritması

Rastgele Orman algoritması, karar ağacı algoritmasının türlerinden biridir. Bu algortmada girdi olarak verilen değerlerden çıktı elde edilirken ağaç yapısı temel alınarak yeni ağaçlar üretilir. Ayrıca algoritma ile kullanılacak veri seti parçalara bölünüp, her bir veri seti ile düğümler oluşturulur. Daha sonra düğümlerden çıkan ağaçlar için rastgele bir şekilde özellik seçimi uygulanır ve ağaçlar geliştirilmeye devam edilir. Bu algoritmanın karar ağacı algoritmasından en büyük farkı ise budama işleminin yapılmamasıdır (Özlem ve Güngör, 2012). Rastgele Orman algoritmasının çalışma prensibi Şekil 4.9’da görülebilir.



Şekil 4.9: Rastgele Orman algoritmasının çalışma prensibi (Gajjar ve Zaveri, 2017).

Rastgele orman algoritmasında, sınıflandırma başarısını belirleyen önemli üç değişken vardır. Bu üç değişken şunlardır (Biau ve Scornet, 2016):

- Her bir ağaçta bulunan veri düğümlerinin sayısı.
- Dallanma için gidilebilecek yön sayısı.
- Dallanmayan düğümün altındaki veri sayısı.

Yukarıda belirtilen üç değişken göz önünde bulundurularak, veri seti yapısına uygun olacak şekilde bir model geliştirilebilmektedir. Ayrıca bu işlemler ile oluşan her bir düğümün homojenliği ölçülerek doğru bir sınıflandırma amaçlanır. Homojenlik hesaplanma formülü Şekil 4.10’da görülebilir. Burada T değeri eğitime sokulan veriyi, C_i verinin ait olduğu sınıfı ve $f(C_i, T)/T$ ibaresi ise seçilen her bir örneğin belirlenen sınıfın bir üyesi olması olasılığını belirtmektedir (Özlem ve Güngör, 2012).

$$\sum \sum_{j \neq i} (f(C_i, T)/|T|)(f(C_j, T)/|T|)$$

Şekil 4.10: Homojenlik hesaplama formülü (Özlem ve Güngör, 2012).

4.7 Naive Bayes Algoritması

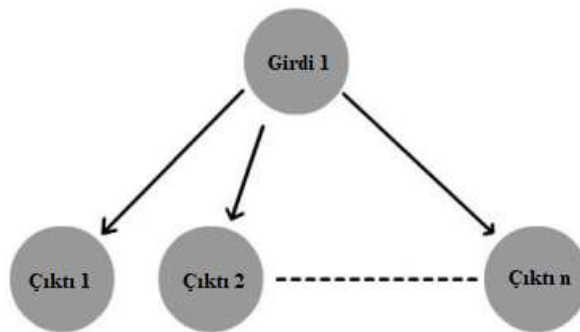
Naive Bayes algoritması, Bayes teoremine dayanan bir sınıflandırma algoritmasıdır. Bu algoritmada, özellikler bir vektör şeklinde algoritmaya verilir ve Bayes teoremi sayesinde Naive Bayes sınıflandırıcısı hesaplanır. Bu hesaplamadaki temel girdiler, istenilen sınıfın doğru veya yanlış tahmin edilme olasılıklarıdır (Arpacı ve Kalıpsız, 2018).

Bayes teoremi, sınıflandırmada tahmin edilecek sınıfın olasılığı ile sınıfın içerisindeki bilinen bir özelliğin olasılığının hesaplanması kullanılarak yeni bir özelliğin sınıflandırılması için kullanılır. Yani bu teoremden, bilinen bir özellik için sınıf tahmin olasılığı ve bilinen bir sınıf için bilinmeyen bir özelliğin tahmin olasılığından yararlanılır. Bu teorem Şekil 4.11’de verilmiştir. Teoremden sınıf olarak C değeri ve özellik olarak x değeri kullanılmıştır (Solmaz vd., 2014).

$$P(C|x) = P(C) \cdot \frac{P(x|C)}{P(x)}$$

Şekil 4.11: Bayes teoremi (Solmaz vd., 2014).

Bayes algoritması grafiksel bir model olarak da gösterilebilmektedir. Model olarak gösterimde her bir düğüm, algoritma içerisindeki bir değişkeni temsil etmektedir. Ayrıca bu düğümler arasındaki bağlantılar ise algoritmanın değişkenlerinin arasındaki ilişkileri temsil etmektedir. Bu ilişkiler sayesinde değişkenlerin birbirlerine bağımlılıkları tespit edilebilmektedir (Haltaş ve Alkan, 2013). Basit bir Naive Bayes yapısı Şekil 4.12’de görülebilmektedir.



Şekil 4.12: Basit bir Naive Bayes yapısı (Haltaş ve Alkan, 2013).

5. KULAK BİYOMETRİĞİ ÇALIŞMALARI

Çalışmanın bu bölümünde, kulak görüntüleri yardımı ile kulak tanıma, kimlik tespiti veya soft biyometrik özelliklerin tahminini ve sınıflandırmasını yapan çalışmalar incelenmiştir.

Fırtına ve Silahtaroglu (2014) çalışmalarında, kulaktan kişi ve kimlik tespiti için bir uygulama geliştirilmesi üzerinde durmuşlardır. Çalışmada kulak görüntüleri üzerinden kenar saptama yöntemi kullanarak kulağa ait düğüm noktalarının tespitini sağlamışlardır. Daha sonra OpenCV algoritmasındaki Haar Cascade algoritması ile kimlik tespiti yapmışlardır. Yapılan çalışma sonucunda %78,4 doğruluk oranı ile kimlik tespiti gerçekleştirmişlerdir.

Oraved vd. (2016) çalışmalarında, mobil cihazlar üzerinde kulak tanıma yapan bir uygulama geliştirmişlerdir. Çalışmalarında veri seti için kendi oluşturdukları veri seti ve WPUTE veri tabanını kullanmışlardır. Çalışmalarında kulak tanımlama için Haar Cascade ve Öklid Uzaklığı algoritmalarından yararlanmışlardır. Yaptıkları çalışma sonucunda en başarılı algoritmanın kendi veri setlerinde %90 ve WPUTE veri setinde %65 doğruluk ile tanıma yapabileceği sonucunu ortaya koymuşlardır.

Boodoo-Jahanger ve Baichoo (2013) çalışmalarında, Lokal İkili Örüntüler algoritmasını kullanarak kulak görüntülerinden kulak tanıma yapmayı amaçlamışlardır. Araştırmada veri tabanı olarak IIT Delhi veri tabanını kullanmışlardır. Kullandıkları algoritma ile kulak görüntülerini parçalara bölerek her bir parça içerisindeki öznitelik noktalarını hesaplamışlardır. Yaptıkları çalışma sonucunda kulak görüntülerinden tanıma için %93'lük bir doğruluk elde etmişlerdir.

Ganapathi vd. (2018) çalışmalarında, üç boyutlu kulaktan kimlik tanıma için geometrik istatistik temelli bir sistem oluşturmuşlardır. Çalışmalarında Notre Dame Üniversitesi Halka Açık Veri Tabanı'nı kullanmışlardır. Sistemlerinde kulağın üç boyutlu uzaydaki öznitelik noktalarını bularak bu noktaları iki boyutlu düzleme indirip kimlik tanıma uygulaması gerçekleştirmişlerdir. Bu noktaların bulunmasında da Yinelemeli En Yakın Nokta algoritmasından faydalanmışlardır. Yaptıkları çalışma sonucunda, üç boyutlu kulak verilerinden %98,60 doğruluk oranı ile kimlik tespiti gerçekleştirmişlerdir.

Yaman vd. (2018) çalışmalarında, kulak görüntülerinden yaş ve cinsiyet tahmini gerçekleştirmişlerdir. Çalışmada algoritma olarak ESA'dan yararlanmışlardır. Ayrıca veri tabanı olarak Multi-Pie veri tabanını kullanmışlardır. Sonuç olarak yaş için %52 ve cinsiyet için %94 oranında doğru sınıflandırma gerçekleştirmişlerdir.

Gnanasivam ve Muttan (2013) çalışmalarında, kulak biyometriklerinden cinsiyet tahmini gerçekleştirmişlerdir. Çalışmada algoritma olarak Bayes, k-EYK ve YSA algoritmalarından yararlanmışlardır. Çalışma sonucunda k-EYK algoritması ile cinsiyet tahmininde %90,42'lik bir doğruluk elde etmişlerdir.

Khorsandi ve Abdel-Mottaleb (2013) çalışmalarında, iki boyutlu kulak görüntülerinden cinsiyet tahmini üzerinde durmuşlardır. Çalışmada kulak görüntülerinden öznelik çıkarımı için Gabor filtrelerinden yararlanmışlardır. Ayrıca çalışmada veri tabanı olarak Notre Dame Üniversitesi Halka Açık Veri Tabanı'nı kullanmışlardır. Çalışmada sınıflandırma için Yumuşatılmış L0 Norm algoritmasından yararlanmışlardır. Çalışma sonucunda %89,49 doğruluk ile cinsiyet sınıflandırması yapmışlardır.

Hassaballah vd. (2020) çalışmalarında, kulaktan görüntülerinden tanıma yapmak için SYYM'den yararlanmışlardır. Çalışmada veri tabanı olarak IIT Delhi – 1 IIT Delhi – 2, AMI ve AWE veri tabanlarından faydalanmışlardır. Çalışma sonucunda Delhi – 2 veri tabanı üzerinde %98 oranında doğrulukla tespit gerçekleştirebilmişlerdir.

Omara vd. (2016) çalışmalarında, kulak görüntülerinden kulak tanıma yapmak için geometrik öznelik çıkarımı yapmışlardır. Kulak öznelik çıkarımı için kendi geliştirdikleri algorithmada maksimum ve minimum olarak Kulak Yükseklik Çizgisi verilerinden yararlanmışlardır. Çalışmalarını USTB ve IIT Delhi veri tabanları ile gerçekleştirmişlerdir. Sonuç olarak USTB ve IIT Delhi veri tabanları üzerinde sırasıyla %98,33 ve %99,60 doğruluk oranı ile kulak tanıma gerçekleştirmişlerdir.

Karasulu vd. (2022) çalışmalarında, kulak görüntülerinden cinsiyet tahmini için derin öğrenme temelli melez bir model tasarlamışlardır. Çalışmalarında ESA ve TSA yapılarını bir arada kullanarak cinsiyet tahmini doğruluk oranlarını yükseltmeyi amaçlamışlardır. Çalışma kapsamında veri tabanı olarak EarVN ve WPUT veri tabanlarını kullanmışlardır. Çalışma sonucunda EarVN için %85,16, WPUT için %87,61 cinsiyet tahmini doğruluk oranı

elde etmişlerdir.

Rahman vd. (2014) çalışmalarında, kulak tanıma için geometrik özellikleri kullanmışlardır. Kulak görüntülerinden öznitelik çıkarımı için Canny Kenar Tespit algoritmasından yararlanmışlardır. Algoritma testlerini 50 adet kulak görüntüsü üzerinde gerçekleştirmişlerdir. Çalışma sonucunda %96,8 oranında doğruluk ile kulak tanıma gerçekleştirmişlerdir.

Jacob ve Raju (2014) çalışmalarında, iki boyutlu kulak görüntülerinden kulak tanıma için doku özelliklerini kullanan bir yaklaşım üzerinde durmuşlardır. Çalışmada doku özelliklerinin kulak tanıma kullanılması alanında kullanılan Gri Seviye Kesişim Matrisi, Lokal İkili Örüntü ve Gabor filtrelerinin verimliliklerini test etmişlerdir. Sonuç olarak geliştirilen yöntem ile %94,12 doğruluk oranı ile kulak tanıma gerçekleştirmişlerdir.

Anwar vd. (2015) çalışmalarında, kulak tanıma için geometrik özellikleri kullanmışlardır. Çalışmada geometrik özellik olarak şekil, ortalama, medyan ve Öklid Uzaklığı gibi özelliklerden yararlanmışlardır. Son olarak sınıflandırma için k-EYK algoritmasından yararlanmışlardır. Yapılan testler sonucunda %98 oranında bir doğrulukla kulak tanıma gerçekleştirmişlerdir.

Ying vd. (2018) çalışmalarında, ESA kullanarak kulak tanıma gerçekleştirmişlerdir. Kulak tanıma için ESA algoritmalarını farklı açılarda çekilmiş kulak görüntüleri ile test etmişlerdir. Çalışma sonucunda farklı açılardaki kulak görüntülerinde %90 üzerindeki oranlarda kulak tanıma Başarıyla elde etmişlerdir.

Lei vd. (2013) çalışmalarında, üç boyutlu kulak görüntülerinden cinsiyet tahmini üzerinde durmuşlardır. Çalışmada kulak görüntülerinden öznitelik çıkarımı için DSH ve DVM algoritmalarından yararlanmışlardır. Çalışmanın veri tabanı olarak Notre Dame Üniversitesi Halka Açık Veri Tabanı'nı kullanmışlardır. Çalışma sonucunda, üç boyutlu kulak görüntülerinden cinsiyet sınıflandırması için %92,94 doğruluk oranı elde etmişlerdir.

Nguyen-Quoc vd. (2020) çalışmalarında, kulak görüntülerinden cinsiyet tahmini üzerinde durmuşlardır. Çalışma kapsamında veri tabanı olarak EarVN1.0 veri tabanını kullanmışlardır. Sınıflandırma algoritmalarında ise k-EYK, DVM ve Rastgele Orman

algoritmalarını kullanmışlardır. Sonuç olarak %80,36 doğruluk oranı ile cinsiyet tahmini gerçekleştirmişlerdir.

Dodge vd. (2018) çalışmalarında, Derin Sinir Ağları yardımı ile kulak tanıma işlemi gerçekleştirmişlerdir. Çalışmada öğrenme ve sınıflandırma algoritması olarak Derin Sinir Ağları algoritmasını kullanmışlardır. Çalışma sonucunda test sonuçları arasındaki en yüksek doğruluk olarak %93,48 ile kulak tanıma işlemini Derin Sinir Ağları yöntemiyle tamamlamışlardır.

Zhang ve Whang (2011) çalışmalarında, yüz profil ve kulak görüntülerinden cinsiyet tahmini yapmışlardır. Cinsiyet tahmini yapılırken, özneliklerin tespiti için Özellik Çantası yaklaşımından yararlanmışlardır. Çalışmada yapılacak sınıflandırma için DVM algoritmasını kullanmışlardır. Sonuç olarak %97,65 doğruluk oranı ile cinsiyet sınıflandırması gerçekleştirilmiştir.

Khalidi vd. (2021) çalışmalarında, denetimsiz derin aktif öğrenme kullanarak kulak tanıma işlemi gerçekleştirmeyi hedeflemişlerdir. Çalışma için Matematiksel Görüntü Analizi, Beijing Teknoloji Üniversitesi ve Açıklamalı Kulak Ağları adlı üç farklı veri tabanı kullanıp, sınıflandırma için ise ESA algoritmasından yararlanmışlardır. Sonuç olarak farklı veri tabanlarında %100, %98,33 ve %51,25 doğruluk oranı ile kulak tanıma gerçekleştirmişlerdir.

Alshazly vd. (2020) çalışmalarında, sınır olmadan kulak tanıma için Derin ESA algoritmalarını kullanmayı test etmişlerdir. Test için EarVN1.0 veri tabanından yararlanmışlardır. Algoritma testlerini ESA algoritması ile gerçekleştirmişlerdir. Çalışmada kullanılan veri seti için ESA algoritmasını ince ayar ile düzenledikten sonra, en yüksek doğruluk olarak %95,85 ile kulak tanıma işlemini gerçekleştirmişlerdir.

Literatür taramasındaki çalışmaların tahmin sınıfı, kullanılan algoritma, veri seti ve doğruluklarına ait bilgiler Tablo 5.1’de verilmiştir.

Tablo 5.1: Literatürdeki kulak biyometriği çalışmaları.

Referans	Çalışma Odağı	Kullanılan Algoritmalar	Veri seti boyutu	Başarı
Fırtına vd.(2014)	Kimlik tespiti	Haar Cascade	273	%78,4
Oravec vd. (2016)	Kulak tanıma	Haar Cascade – Öklid Uzaklığı	20	%90
Oravec vd. (2016)	Kulak tanıma	Haar Cascade – Öklid Uzaklığı	60	%65
Boodoo-Jahanger ve Baichoo (2013)	Kulak tanıma	Lokal İkili Görüntüler	375	%93
Ganapathi vd (2018)	Kimlik tespiti	Yinelemeli En Yakın Nokta	1779	%98,60
Yaman vd. (2018)	Yaş ve Cinsiyet tahmini	ESA	270	%52 ve %94
Gnanasivam ve Muttan (2013)	Cinsiyet tahmini	Bayes - k-EYK - YSA	342	%90,42
Khorsandi ve Abdel-Mottaleb (2013)	Cinsiyet tahmini	Gabor filtreleri	2340	%89,49
Hassaballah vd. (2020)	Kulak tanıma	SYVM	793	%98
Omara vd. (2016)	Kulak tanıma	Kendi geliştirilen	180	%98,33
Omara vd. (2016)	Kulak tanıma	Kendi geliştirilen	375	%99,60
Karasulu vd. (2022)	Cinsiyet tahmini	ESA - YSA	28412	%85,16
Karasulu vd. (2022)	Cinsiyet tahmini	ESA - YSA	2071	%87,61
Rahman vd. (2014)	Kulak tanıma	Canny Kenar Tespit	50	%96,8
Jacob ve Raju (2014)	Kulak tanıma	GSKM - Lokal İkili Örüntü - Gabor filtreleri	442	%94,12
Anwar vd. (2015)	Kulak tanıma	k-EYK	150	%98
Ying vd. (2018)	Kulak tanıma	TBA - DVM	Belirtilmemiş	%90
Lei vd. (2013)	Cinsiyet tahmini	DSH - DVM	942	%92,94
Nguyen-Quoc vd. (2020)	Cinsiyet tahmini	k-EYK – DVM – Rastgele Orman	28412	%80,36
Dodge vd. (2018)	Kulak tanıma	Derin Sinir Ağları	1804	%93,48
Zhang ve Whang (2011)	Cinsiyet tahmini	DVM	942	%97,65
Khaldi vd. (2021)	Kulak tanıma	ESA	700	%100
Khaldi vd. (2021)	Kulak tanıma	ESA	308	%98,33
Khaldi vd. (2021)	Kulak tanıma	ESA	1000	%51,25
Alshazly vd. (2020)	Kulak tanıma	ESA	28412	%95,85

6. KULAK VERİ SETLERİNİN HAZIRLANMASI

Çalışmada yapılacak sınıflandırma iki farklı veri tabanı oluşturulmuştur. Oluşturulan ilk veri tabanı (VT1) için kulak görüntülerinden ve anket verilerinden yararlanılmıştır. Çalışma kapsamında Bartın ili içerisinde bulunan 105 kişiden kulak görüntüsü ve bu görüntüler ile ilişkilendirilmek üzere anket verileri toplanmıştır.

Çalışmada ikinci olarak kullanılan veri tabanı (VT2) için, internet üzerinden ünlü kişilere ait kulak görüntüleri ve soft biyometrik özellikler elde edilmiştir. Çalışma kapsamında 250 ayrı ünlü kişiye ait soft biyometrik özellikler ve 300 farklı kulak görüntüsü sınıflandırma için elde edilmiştir.

6.1.1. VT1 İçin Kulak Görüntülerinin Elde Edilmesi ve Kaydedilmesi

Çalışmada kullanılan VT1 için kulak görüntülerinin elde edilmesinde 20.4 mp çözünürlüğe sahip Sony A58 model fotoğraf makinesi kullanılmıştır. Fotoğraf makinesi ile çekilen görüntüler yarı kontrollü ortamda elde edilmiştir. Fotoğraf çekilen ortamlarda ışık ve arka plan değişiklik gösterebilmektedir. Görüntüler elde edilirken üzerinde durulan temel husus, çekim yapılan kişinin kameraya uzaklığı ve açısıdır. Veri tabanında kullanılacak görüntüler için bu uzaklık 50 cm olarak belirlenmiştir. Ayrıca kamera açısı da kulağa karşıdan bakacak şekilde 90 derece olarak ayarlanmıştır. Araştırmada kullanılan VT1 veri tabanına ait örnek bir işlenmemiş görüntü Şekil 6.1’de görülebilir.



Şekil 6.1: VT1 veri tabanındaki kulak görüntüsü örneği.

6.1.2. VT1 İin Anket Verilerinin Veri Tabanına Kaydedilmesi

alıřma kapsamında VT1 iin kulak grntleri ile elde edilen anket verileri, kulak grntleri ile iliřkilendirilebilmek adına numaralandırılmıştır. Numaralandırılmanın kullanılmasındaki başlıca sebep, grntye ve bilgilere ait kiřinin kimliėinin alınmamasıdır. Yapılan numaralandırma gn/ay/yıl ve gnn kaıncı grnts olduėu řeklinde yapılmıştır. rnek olarak 20 Eyll 2022 tarihinde ilk yapılan kiřinin numarası “20092201” olarak verilmiştir. Bu sayede kulak grnts ve anket verisi iliřkilendirmiştir.

Anket ile elde edilen veriler numaralandırıldıktan sonra Microsoft Excel uygulamasında sınıflandırılmıştır. Bu sınıflar cinsiyet, yař, boy, doėal sa rengi, sa tipi, gz rengi ve ten rengi olmak zere yedi adet olarak belirlenmiştir. Grnt iřleme yntemleri ile elde edilen veriler de yine aynı dosya zerinde tahmin iřlemlerinde kullanılmak zere depolanmıştır.

6.2. VT2 İin Kulak Grntlerinin ve Soft Biyometrik zelliklerin Elde Edilmesi ve Kaydedilmesi

alıřmada kullanılan VT2 iin kulak grntlleri, internet zerinden fotoėrafın ekilme tarihi gz nnde bulundurularak elde edilmiştir. Elde edilen grntler farklı znrlklere ve farklı ışıklandırma ortamlarına sahiplerdir. VT2’deki grntler VT1’e gre daha ok farklılık gsterdiėini sylemek mmkndr. Arařtırmada kullanılan VT2 veritabanına ait bir grnt řekil 6.2’de verilmiştir. Ayrıca kulak grntlerinin sahibi olan kiřilerin soft biyometrik zelliklerine de internet zerinden eriřilmiştir.



řekil 6.2: VT2 veri tabanındaki kulak grnts rneėi.

7. KULAK GÖRÜNTÜSÜNDEN SOFT BİYOMETRİK ÖZELLİKLERİN TAHMİN EDİLMESİ

Literatürde kulak görüntülerinden soft biyometrik özelliklerin çıkarımı için farklı yöntemler bulunmaktadır. Bu çalışmada, soft biyometrik özelliklerin tahmini için kulak görüntüleri ön işlemlerden geçirilerek, görüntü üzerindeki piksellerin konumlarından yararlanılmıştır. Bir sonraki bölümde kulak görüntülerinden çıkarım yapılması için yapılan görüntü işleme yöntemleri açıklanmaktadır.

7.1. Kulak Görüntüleri Ön İşlemleri

Çalışma kapsamında VT1 veri tabanı için Sony A58 fotoğraf makinesi ile elde edilen görüntüler 6048 x 4032 piksel çözünürlüğündedir. Her iki veri tabanı için ilk olarak görüntüler sadece kulak kısmının görüleceği şekilde kırılıp, görüntü çözünürlüğü 64 x 64 piksel olacak şekilde yeniden ölçeklendirilmiştir. Bu aşamalardan geçen örnek bir görüntü Şekil 7.1’de görülebilir.



Şekil 7.1: Kırpma ve yeniden boyutlandırma işleminden geçen kulak görüntüsü a) VT1 b) VT2

Görüntüleri kırpma ve yeniden boyutlandırma işlemi Python programlama dili ile gerçekleştirilmiştir. Bu işlemlerden sonraki ön işleme işlemlerinde Matlab uygulaması kullanılmıştır.

Kırılmış ve yeniden boyutlandırılmış görüntüler, Matlab uygulamasında ilk olarak grayscale formata çevrilmiştir. Yani bütün renkler grinin tonu olacak şekilde ayarlanmıştır. Bu aşamadaki görüntü Şekil 7.2’de görülebilir.



Şekil 7.2: Kırpma, yeniden boyutlandırma ve grayscale formata çevrilme işleminden geçen kulak görüntüsü a) VT1 b) VT2

Grayscale formata çevrilme işlemine tabi tutulan görüntüler daha sonra histogram eşitlemesine tabi tutulmuştur. Bu sayede ortam ışıklandırmasından kaynaklanabilecek farklılıkların azaltılması amaçlanmıştır. Histogram eşitlemesinde kullanılan referans görüntü, veritabanındaki ilk görüntü olarak seçilmiştir. Önceki ön işleme işlemlerine ve histogram eşitleme işlemine tabi tutulan örnek bir görüntü Şekil 7.3’de görülebilir.



Şekil 7.3: Kırpma, yeniden boyutlandırma, grayscale formata çevrilme ve histogram eşitleme işleminden geçen kulak görüntüsü a) VT1 b) VT2

İnternet üzerinden elde edilen kulak görüntülerinin farklı çözünürlükte olmasından kaynaklanabilecek bozukluklar için VT2 veri tabanındaki görüntülere, histogram eşitleme işleminden sonra medyan filtre uygulanmıştır. Bu sayede görüntüde oluşabilecek kirliliklerin en aza indirilmesi amaçlanmıştır. Medyan filtre uygulanmış kulak görüntüsü örneği Şekil 7.4’de verilmiştir.



Şekil 7.4: Kırpma, yeniden boyutlandırma, grayscale formata çevrilme, histogram eşitleme ve medyan filtre işleminden geçen örnek bir VT2 veri tabanı görüntüsü

Görüntüler, önceki ön işleme işlemlerinden geçirildikten sonra, görüntüdeki piksellerin konumlarının siyah ve beyaz ikili sistemde alınabilmesi için görüntüler siyah – beyaza çevrilmiştir. Bu sayede beyaza yakın olan gri pikseller beyaz, siyaha yakın olan gri pikseller de siyah olarak dönüştürülmüştür. Ayrıca bu sayede piksel verileri sıfır ve bir olarak elde edilebilmektedir. Tüm ön işlemlerden geçmiş olan örnek bir kulak görüntüsü Şekil 7.5’de görülebilir. İki veri tabanı için uygulanan işlemler arasındaki tek fark VT2 veri tabanına ait kulak görüntülerine histogram eşitleme işleminden sonra medyan filtre uygulanmasıdır. Bu filtrenin uygulanma sebebi, VT2’nin internet üzerinden elde edilen kulak görüntülerinden oluşmasından ötürü oluşabilecek gereksiz görüntü bozukluklarını giderilmesidir.



Şekil 7.5: Kırpma, yeniden boyutlandırma, grayscale formata çevrilme, histogram eşitleme ve siyah-beyaz işleminden geçen kulak görüntüsü a) VT1 b) VT2

Görüntülere ait pikseller, tüm ön işlemlere tabi tutulduktan sonra, siyah veya beyaz olmalarına göre sıfır ve bir değerlerini alabilmektedir. Ancak bu piksel değerlerinin sınıflandırmada kullanılabilmesi için tek yönlü bir vektör olarak alınması gerekmektedir. Normal haliyle iki boyutlu olan görüntü piksel verileri (en ve boy) tek bir boyuta çekilmek üzere, her satır bir önceki satırın sonuna gelecek şekilde düzenlenmiştir. Daha sonra her bir piksel verisi bir kolona gelecek şekilde, anket form verilerinin tutulduğu Excel dosyasına kaydedilmiştir.

7.2. Soft Biyometrik Tahmin Sisteminin Geliştirilmesi

Ön işleminden geçirilen görüntüler ve düzenlenen form verileri, sınıflandırma ve tahmin için RapidMiner ve Weka uygulamalarında analiz edilmiştir. Bu uygulamaların seçilme nedenleri, daha önce soft biyometrik tahmin çalışmalarında tercih edilen veri işleme uygulamaları olmalarıdır. Örnek olarak Gillani vd. (2020) çalışmalarında yürüyüş analizi kullanarak insan yaşını ve cinsiyetini tahmin etmeye çalışmışlardır. Çalışmada kullanılan verilerin sınıflandırılması için Weka uygulamasından yararlanmışlardır. Rapidminer kullanılan bir çalışmaya örnek olarak da Pahwa ve Aggarwal (2016) çalışmalarında, ses analizinden cinsiyet tahmini için Rapidminer uygulamasından yararlanmışlardır.

Bu uygulamaların sonuçlarının karşılaştırılması için benzer sınıflandırma algoritmaları kullanılmasına karşın, bazı tahmin sınıflarında ikiden fazla veri çeşidi olabileceği için sadece ikili sınıflandırma yapabilen algoritmalar kullanılmamıştır.

Yapılan tahmin işlemlerinde kullanılan algoritmaların testleri ve doğruluk sınıfları için tüm tahmin işlemlerinde 10 katlamalı çapraz doğrulama kullanılmıştır. Bu sayede test ve eğitim verisinin değişkenliğinden kaynaklanabilecek yanlış doğruluk oranlarının önüne geçilmesi amaçlanmıştır. Ayrıca ikiden fazla sınıf tahmini yapılan testlerde uygulamalar polinomsal sınıftan ikili sınıfa çevirme uygulayarak DVM gibi algoritmaların ikiden fazla sınıf bulunan veri setlerini sınıflandırmada kullanılmasını sağlamıştır.

7.3. Tahmin İşlemleri İçin Kıstaslar

Anket verileri ile elde edilen bilgilerin görüntülerden tahmin edilme işlemlerinde, tahmin sınıflarının tahmin doğrulukları en yüksek doğruluğa sahip ilk iki algoritma olarak bir sonraki bölümde detaylı bir şekilde açıklanmıştır. Seçilen iki algoritma, DVM, Derin Öğrenme, Lojistik Regresyon, LibSVM, Doğrusal Regresyon, Rastgele Orman ve Naive Bayes algoritmaları arasından en başarılı olan algoritmalarlardır. Ayrıca çalışmanın doğruluğu açısından tahmin gruplarından %10'un altında olan gruplar ilgili algoritma testlerine dâhil edilmemiştir. Bunun sebebi %10'un altında olan gruplar, algoritma testlerinde tahmin edilememektedir.

Algoritma testleri sonucunda yapılan tahmin işlemlerinin sonuçları, sınıflandırma başarısı dışında yanlış pozitif, doğru pozitif, recall, precision ve F1 skoru gibi kıstaslara detaylı olarak verilmiştir. Bu kıstaslar sınıflandırma işlemleri sonucunda elde edilen hata matrisine örnek bir matris üzerinden hesaplanarak açıklanmıştır. Örnek hata matrisi Şekil 7.6’da verilmiştir.

	Erkek olarak sınıflandırılmış	Kadın olarak sınıflandırılmış
Gerçek Erkek	49	10
Gerçek Kadın	6	40

Şekil 7.6: Örnek bir hata matrisi

Örnek olarak verilen hata matrisinden “erkek” cinsiyet grubu için elde edilebilecek sonuç kıstasları şunlardır (Grandini vd., 2020):

- **Doğru Pozitif:** Algoritmanın tahmin edilmek istenilen sınıfı **doğru** bir şekilde tahmin ettiğini gösteren bir değerdir. Örnek hata matrisi için bu değer 49’dur.
- **Yanlış Pozitif:** Algoritmanın tahmin edilmek istenilen sınıf dışındaki tahminlerinde, tahmin edilmek istenilen sınıfı tahmin ettiğini gösteren bir değerdir. Örnek hata matrisi için bu değer 6’dır.
- **Doğru Negatif:** Algoritmanın tahmin edilmek istenilmeyen sınıfları **doğru** olarak tahmin ettiğini gösteren bir değerdir. Örnek hata matrisi için bu değer 40’dır.
- **Yanlış Negatif:** Algoritmanın tahmin edilmek istenilen sınıfa ait değerleri farklı bir sınıf olarak tahmin ettiğini gösteren bir değerdir. Örnek hata matrisi için bu değer 10’dur.
- **Recall:** Doğru pozitiflerin, doğru pozitifler ve yanlış negatif ile toplamına oranı ile hesaplanır. Recall hesaplama formülü Şekil 7.7’de görülebilir. Örnek hata matrisi için bu değer 0,831’dir.

$$Recall = \frac{DP}{DP + YN}$$

Şekil 7.7: Recall hesaplama formülü.

- **Precision:** Doğru pozitiflerin tüm tanımlamalara göre oranını belirtir. Yani doğru pozitifin, yanlış ve doğru pozitiflerin toplamına oranı ile hesaplanır. Precision hesaplama formülü Şekil 7.8’de görülebilir. Örnek hata matrisi için bu değer 0,891’dir.

$$Precision = \frac{DP}{DP + YP}$$

Şekil 7.8: Precision hesaplama formülü.

- **F1 Skoru:** Recall ve precision değerlerinin harmonik ortalamasıdır. F1 skoru hesaplama formülü Şekil 7.9’da görülebilir. Örnek hata matrisi için bu değer 0,860’dır.

$$F1 = 2 * \frac{precision * recall}{precision + recall}$$

Şekil 7.9: F1 skoru hesaplama formülü.

- **Sınıflandırma Başarısı:** Algoritmanın sınıflandırma başarısını gösteren değerdir. Sınıflandırma başarısının hesaplama formülü şekil 7.10’da görülebilir. Örnek hata matrisi için bu değer %84,76’dır.

$$Sınıflandırma Başarısı = \frac{DP + DN}{DP + DN + YP + YN}$$

Şekil 7.10: Sınıflandırma başarısı hesaplama formülü.

7.4. VT1 İçin Tahmin İşlemleri

7.4.1. VT1 İçin Cinsiyet Tahmini

Çalışma kapsamında elde edilen form verilerindeki cinsiyet sınıfı verilerine bakıldığında, 53 erkek ve 52 kadın katılımcı olduğu görülmüştür. Bu durum, sınıf tahmini için doğru bir dağılım olabileceğinden tahmin işlemine yardımcı olabileceği öngörülmüştür.

Cinsiyet tahmini için RapidMiner uygulamasında yapılan testler sonucunda en başarılı iki algoritmanın sonuçları Tablo 7.1’de verilmiştir. Bu algoritmalar LibSVM ve DVM olup, doğruluk oranları sırasıyla %84,18 ve %80 olarak tespit edilmiştir.

Tablo 7.1: VT1 için RapidMiner uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının cinsiyet sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
LibSVM	10	45	7	43	0,865	0,818	0,841	Kadın	%84,18
	7	43	10	45	0,811	0,860	0,835	Erkek	
DVM	12	43	9	41	0,827	0,782	0,804	Kadın	%80,00
	9	41	12	43	0,811	0,860	0,835	Erkek	

Cinsiyet tahmini için aynı veri seti kullanılarak Weka uygulaması üzerinde yapılan testler sonucunda en başarılı algoritmalar %84,76 sınıflandırma başarısı ile LibSVM algoritması ve %77,14 sınıflandırma başarısı ile DVM algoritması olduğu görülmüştür. Bu algoritmaların detaylı sonuçları Tablo 7.2’de verilmiştir.

Tablo 7.2: VT1 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının cinsiyet sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
LibSVM	10	46	6	43	0,885	0,821	0,852	Kadın	%84,76
	6	43	10	46	0,811	0,878	0,843	Erkek	
DVM	12	40	12	41	0,769	0,769	0,769	Kadın	%77,14
	12	41	12	40	0,774	0,771	0,773	Erkek	

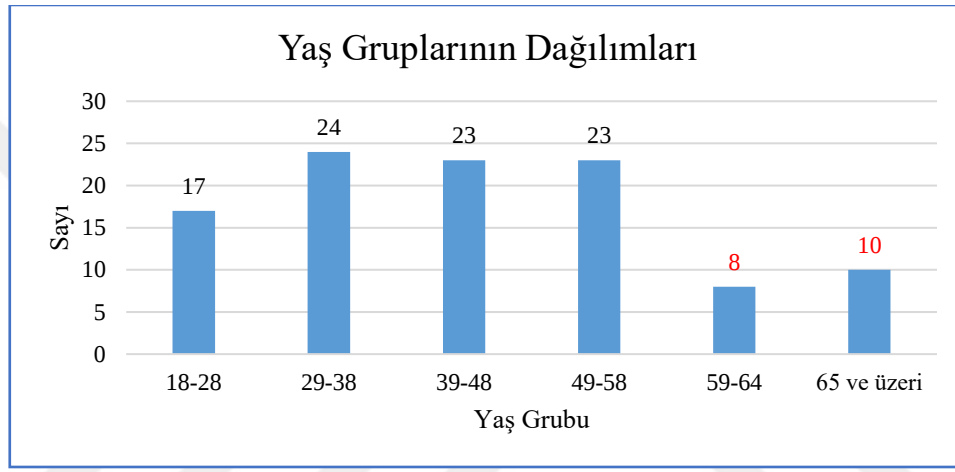
İki farklı uygulamada sınıflandırma algoritmaları ile yapılan testler sonucunda, cinsiyet tahmininde en başarılı algoritmanın %84,76 sınıflandırma başarısı ile Weka uygulamasında bulunan LibSVM algoritması olduğu görülmüştür. RapidMiner uygulaması üzerindeki en başarılı algoritma ise %84,18 doğruluk oranı ile LibSVM algoritması olmuştur. Her iki uygulamanın cinsiyet tahmini için en başarılı algoritmalarının karşılaştırılması Tablo 7.3’de görülebilir.

Tablo 7.3: VT1 için RapidMiner ve Weka uygulamalarının cinsiyet tahmini için en başarılı algoritmalarının karşılaştırılması.

Uygulama	Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	S. Başarısı
RapidMiner	LibSVM	10	45	7	43	0,865	0,818	0,841	Kadın	%84,18
		7	43	10	45	0,811	0,860	0,835	Erkek	
Weka	LibSVM	10	46	6	43	0,885	0,821	0,852	Kadın	%84,76
		6	43	10	46	0,811	0,878	0,843	Erkek	

7.4.2. VT1 İçin Yaş Tahmini

Çalışma kapsamında toplanılan anket verilerine bakıldığında, yaş sınıfında altı farklı yaş aralığının bulunduğu görülmektedir. Bu yaş aralıkları arasından “59-65” ve “65 ve üzeri” grubundaki katılımcıların sayısının %10’dan az olduğu görülmüştür. Bu nedenle yaş tahmini yapılırken en çok katılımcının bulunduğu dört grup olan “18-28”, “29-38”, “39-48” ve “49-58” yaş grupları testlere dâhil edilmiştir. Bu grupların sayı dağılımları Şekil 7.11’de verilmiştir.



Şekil 7.11: VT1 için yaş gruplarının veri dağılımları

RapidMiner uygulamasında yapılan algoritma testleri sonucunda, en başarılı algoritmaların sırasıyla %40,00 ile LibSVM ve %33,33 ile DVM olduğu görülmüştür. Sınıflandırma sonuçlarının detayları Tablo 7.4’de verilmiştir.

Tablo 7.4: VT1 için RapidMiner uygulamasında çalıştırılan Derin Öğrenme ve DVM algoritmalarının yaş sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
LibSVM	3	2	14	33	0,125	0,400	0,191	18-28	%40,00
	18	11	13	24	0,458	0,379	0,415	29-38	
	9	6	17	29	0,261	0,400	0,316	39-48	
	21	16	7	19	0,696	0,432	0,413	49-58	
DVM	6	2	14	13	0,125	0,250	0,167	18-28	%33,33
	18	8	16	16	0,333	0,308	0,320	29-38	
	12	3	20	22	0,130	0,200	0,158	39-48	
	21	16	7	18	0,696	0,432	0,533	49-58	

Yaş tahmini için aynı veri seti kullanılarak Weka uygulamasında yapılan algoritma testlerinde en yüksek doğruluk oranlarına sahip algoritmaların %32,56 sınıflandırma başarısı ile LibSVM ve %29,07 sınıflandırma başarısı ile DVM algoritmaları olduğu görülmüştür. Bu algoritmalara ait detaylı test sonuçları Tablo 7.5’de verilmiştir.

Tablo 7.5: VT1 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının yaş sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
LibSVM	4	1	15	27	0,063	0,200	0,095	18-28	%32,56
	21	11	13	17	0,458	0,344	0,393	29-38	
	16	5	18	23	0,217	0,238	0,227	39-48	
	17	11	12	17	0,478	0,393	0,431	49-58	
DVM	11	3	13	22	0,188	0,214	0,200	18-28	%29,07
	21	8	16	20	0,333	0,276	0,302	29-38	
	12	3	20	22	0,130	0,200	0,158	39-48	
	17	11	12	17	0,478	0,393	0,431	49-58	

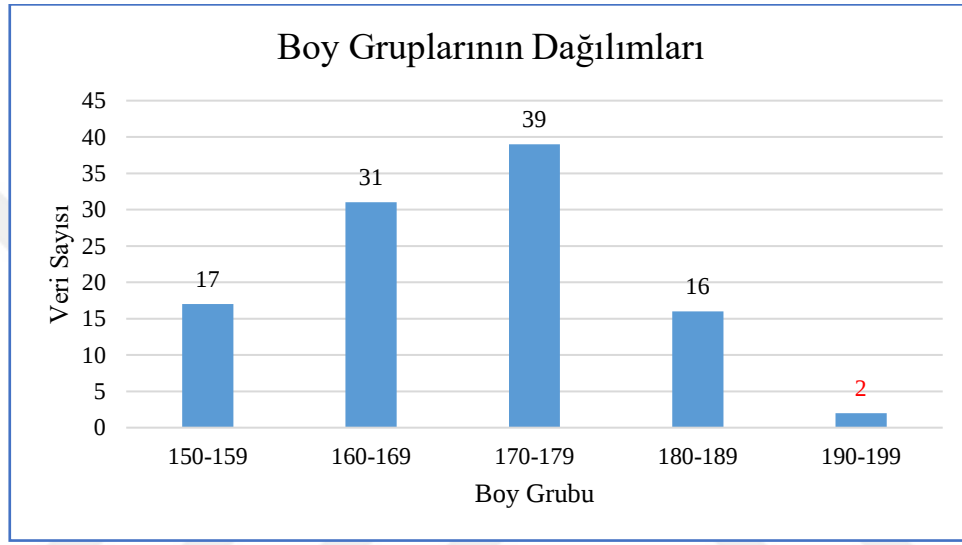
İki farklı uygulamada sınıflandırma algoritmaları ile yapılan testler sonucunda, yaş tahmininde doğruluk oranı en yüksek olan algoritmanın %40,00 doğruluk oranı ile RapidMiner uygulamasında bulunan LibSVM algoritması olduğu görülmüştür. Ayrıca Weka uygulamasındaki en başarılı algoritmanın %32,56 ile LibSVM algoritması oldukları görülmüştür. Yaş sınıflandırması açısından her iki uygulamanın en başarılı algoritmalarının karşılaştırılması Tablo 7.6’da görülebilir.

Tablo 7.6: VT1 için RapidMiner ve Weka uygulamalarının yaş tahmini üzerindeki en başarılı algoritmalarının karşılaştırılması.

Uygulama	Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	S. Başarısı
RapidMiner	LibSVM	3	2	14	33	0,125	0,400	0,191	18-28	%40,00
		18	11	13	24	0,458	0,379	0,415	29-38	
		9	6	17	29	0,261	0,400	0,316	39-48	
		21	16	7	19	0,696	0,432	0,413	49-58	
Weka		4	1	15	27	0,063	0,200	0,095	18-28	%32,56
		21	11	13	17	0,458	0,344	0,393	29-38	
		16	5	18	23	0,217	0,238	0,227	39-48	
		17	11	12	17	0,478	0,393	0,431	49-58	

7.4.3. VT1 İçin Boy Tahmini

Araştırma kapsamında toplanılan anket verileri incelendiğinde, boy sınıfında “190-199” grubunda %10’dan az sayıda katılımcı bulunduğu görülmüştür. Bu nedenle algoritma testlerine bu gruplar dâhil edilmemiştir ve algoritma testlerine “150-159”, “160-169”, “170-179” ve “180-189” boy grupları ile testler gerçekleştirilmiştir. Bu grupların dağılımları Şekil 7.12’de verilmiştir.



Şekil 7.12: VT1 için boy gruplarının veri dağılımları

Boy tahmini için yapılan algoritma testleri sonucunda, RapidMiner uygulaması üzerindeki en başarılı algoritmaların %44,82 ile LibSVM ve %43,00 ile DVM algoritmaları olduğu görülmüştür. Algoritmaların detaylı sonuçları Tablo 7.7’de verilmiştir.

Tablo 7.7: VT1 için RapidMiner uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının boy sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
LibSVM	8	2	15	44	0,118	0,200	0,148	150-159	%44,82
	17	16	15	30	0,516	0,485	0,500	160-169	
	29	25	14	21	0,641	0,463	0,538	170-179	
	3	3	13	43	0,188	0,500	0,191	180-189	
DVM	11	2	15	42	0,118	0,154	0,134	150-159	%43,00
	18	15	16	29	0,484	0,455	0,469	160-169	
	29	26	13	18	0,667	0,473	0,554	170-179	
	1	1	15	43	0,063	0,500	0,112	180-189	

Boy tahmini için Weka uygulamasındaki sınıflandırma algoritmaları ile yapılan testlerin sonucunda, en başarılı algoritmaların %41,75 ile LibSVM ve %37,86 ile DVM algoritmaları oldukları ortaya çıkmıştır. Bu algoritmaların detaylı sonuçları Tablo 7.8’de görülebilmektedir.

Tablo 7.8: VT1 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının boy sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
LibSVM	8	3	14	40	0,176	0,273	0,214	150-159	%41,75
	22	14	17	29	0,452	0,389	0,418	160-169	
	26	23	16	20	0,590	0,469	0,523	170-179	
	4	3	13	40	0,188	0,429	0,261	180-189	
DVM	8	2	15	37	0,118	0,200	0,148	150-159	%37,86
	22	14	17	25	0,452	0,389	0,418	160-169	
	24	19	20	21	0,487	0,442	0,463	170-179	
	10	4	12	35	0,250	0,286	0,267	180-189	

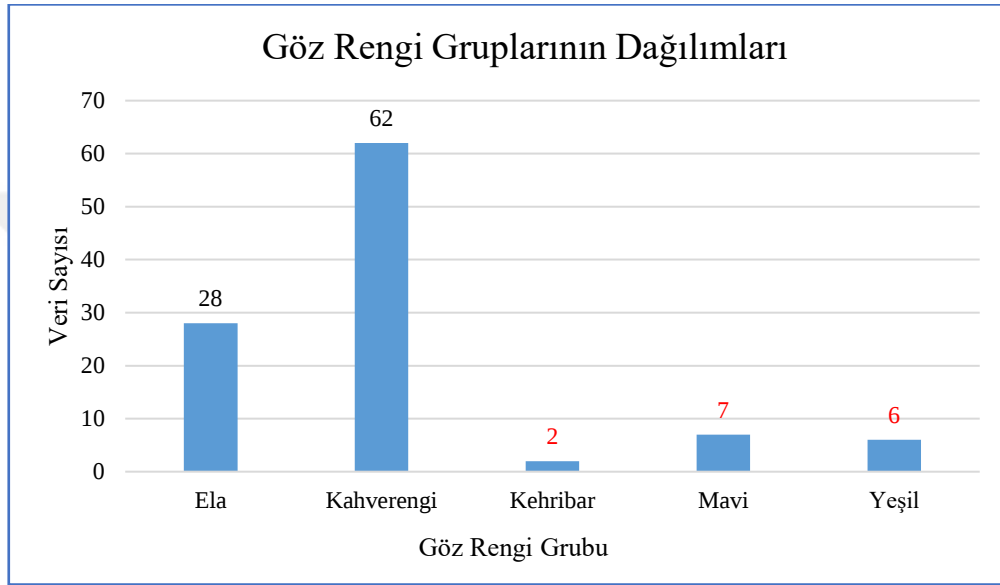
Çalışma kapsamında kullanılan iki farklı uygulama karşılaştırıldığında, en başarılı algoritmaların %44,82 ile RapidMiner uygulamasında bulunan LibSVM algoritması olduğu görülmüştür. Weka uygulamasında kullanılan algoritmalar arasında ise %41,75’lik bir sınıflandırma başarısı ile LibSVM algoritması en başarılı algoritma olmuştur. İki farklı uygulamanın, yapılan testler sonucunda boy sınıflandırması açısından en başarılı algoritmalarının karşılaştırılması Tablo 7.9’da verilmiştir.

Tablo 7.9: VT1 için RapidMiner ve Weka uygulamalarının boy tahmini üzerindeki en başarılı algoritmalarının karşılaştırılması.

Uygulama	Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	S. Başarısı
RapidMiner	LibSVM	8	2	15	44	0,118	0,200	0,148	150-159	%44,82
		17	16	15	30	0,516	0,485	0,500	160-169	
		29	25	14	21	0,641	0,463	0,538	170-179	
		3	3	13	43	0,188	0,500	0,191	180-189	
Weka		8	3	14	40	0,176	0,273	0,214	150-159	%41,75
		22	14	17	29	0,452	0,389	0,418	160-169	
		26	23	16	20	0,590	0,469	0,523	170-179	
		4	3	13	40	0,188	0,429	0,261	180-189	

7.4.4. VT1 İçin Göz Rengi Tahmini

Yapılan çalışma sonucunda anket verilerindeki göz rengi sınıfı incelenmiştir. İnceleme sonucunda “Mavi”, “Yeşil” ve “Kehribar” göz rengine sahip çok az sayıda katılımcı olduğu görülmüştür. Bu nedenle sınıflandırma algoritmaları bu grupları tahmin etmemektedir. Bu nedenle göz rengi tahmininde sadece “Kahverengi” ve “Ela” göz renklerinin tahmini yapılmıştır. Göz rengi gruplarının dağılımları Şekil 7.13’de verilmiştir.



Şekil 7.13: VT1 için göz rengi gruplarının veri dağılımları

Göz rengi ile ilgili RapidMiner uygulaması üzerinde bulunan sınıflandırma algoritmaları ile yapılan testler sonucunda, en yüksek doğruluğa sahip algoritmaların %68,47 ile Derin Öğrenme ve %65,14 ile LibSVM algoritmaları olduğu tespit edilmiştir. Bu algoritmaların detaylı sonuçları Tablo 7.10’da verilmiştir.

Tablo 7.10: VT1 için RapidMiner uygulamasında çalıştırılan Derin Öğrenme ve LibSVM algoritmalarının göz rengi sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
Derin Öğrenme	23	57	5	4	0,920	0,716	0,805	Kahverengi	%68,47
	5	4	23	57	0,148	0,444	0,222	Ela	
LibSVM	26	57	5	1	0,920	0,687	0,787	Kahverengi	%65,14
	5	1	26	57	0,037	0,100	0,054	Ela	

RapidMiner uygulamasında kullanılan aynı veri seti ile Weka uygulamasındaki sınıflandırma algoritmaları test edildiğinde, sınıflandırma başarısı en yüksek algoritmaların %67,41 ile LibSVM ve %66,29 ile Lojistik Regresyon algoritmaları olduğu görülmüştür. Bu algoritmaların detaylı sonuç bilgileri Tablo 7.11’de gösterilmektedir.

Tablo 7.11: Weka uygulamasında çalıştırılan LibSVM ve Lojistik Regresyon algoritmalarının göz rengi sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
LibSVM	19	52	10	8	0,839	0,732	0,782	Kahverengi	%67,41
	10	8	19	52	0,296	0,444	0,356	Ela	
Lojistik Regresyon	21	53	9	6	0,885	0,716	0,779	Kahverengi	%66,29
	9	6	21	53	0,222	0,400	0,286	Ela	

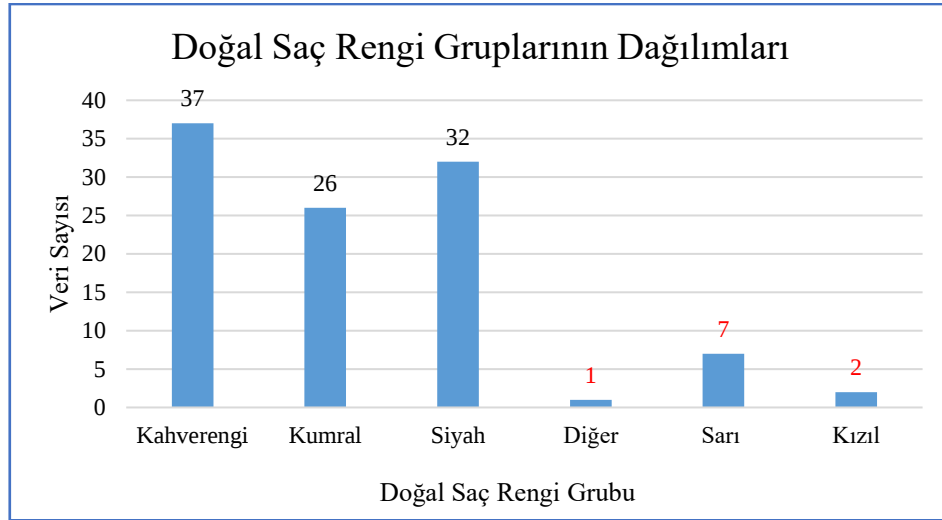
Göz rengi tahmini için her iki uygulamadaki algoritma testi sonuçlarına bakıldığında, RapidMiner uygulamasındaki Derin Öğrenme algoritmasının %68,47 doğruluk ile en başarılı algoritma olduğu belirlenmiştir. Weka uygulamasında yapılan testlerde ise göz rengi tahmini için en başarılı algoritmanın %67,41 doğruluk ile LibSVM olduğu görülmüştür. Her iki uygulamada da test edilen LibSVM algoritmasının iki uygulama bazında karşılaştırılması Tablo 7.22’de verilmiştir.

Tablo 7.12: VT1 için RapidMiner ve Weka uygulamalarının göz rengi tahmini için en başarılı algoritmalarının karşılaştırılması.

Uygulama	Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	S. Başarısı
RapidMiner	Derin Öğrenme	23	57	5	4	0,920	0,716	0,805	Kahverengi	%68,47
		5	4	23	57	0,148	0,444	0,222	Ela	
Weka	LibSVM	19	52	10	8	0,839	0,732	0,782	Kahverengi	%67,41
		10	8	19	52	0,296	0,444	0,356	Ela	

7.4.5. VT1 İçin Doğal Saç Rengi Tahmini

Araştırma kapsamında elde edilen anket verilerinin doğal saç rengi sınıfı veri dağılımı incelendiğinde, “diğer”, “kızıl” ve “sarı” doğal saç rengine sahip kişilerin %10’dan az sayıda olduğu görülmüştür. Bu durum göz önünde bulundurularak “diğer”, “sarı” ve “kızıl” saç rengine ait değerler algoritma testlerine dâhil edilmeyip, “kahverengi”, “kumral” ve “siyah” saç rengi grupları kullanılmıştır. Saç rengi gruplarının dağılımları Şekil 7.14’de verilmiştir.



Şekil 7.14: VT1 için doğal saç rengi gruplarının veri dağılımları

Çalışmada tahmin edilmek istenen doğal saç rengi için RapidMiner uygulaması üzerindeki sınıflandırma algoritmaları test edildiğinde, en başarılı algoritmaların %45,91 ile Derin Öğrenme ve %43,27 ile Lojistik Regresyon algoritmaları olduğu belirlenmiştir. Belirlenen algoritmalara ait detaylı sonuç verileri Tablo 7.13’de gösterilmiştir.

Tablo 7.13: VT1 için RapidMiner uygulamasında çalıştırılan Derin Öğrenme ve Lojistik Regresyon algoritmalarının doğal saç rengi sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
Derin Öğrenme	22	19	18	28	0,514	0,463	0,487	Kahverengi	%45,91
	18	16	17	31	0,485	0,471	0,478	Kumral	
	15	12	20	35	0,375	0,444	0,407	Siyah	
Lojistik Regresyon	22	19	18	25	0,514	0,463	0,487	Kahverengi	%43,27
	17	11	22	33	0,333	0,393	0,361	Kumral	
	19	14	18	30	0,438	0,424	0,431	Siyah	

RapidMiner uygulamasında yapılan algoritma testlerinde kullanılan aynı veri seti ile Weka uygulamasında yapılan algoritma testleri sonucunda, doğal saç rengi için en başarılı algoritmaların %42,16 ile DVM ve %40,20 ile LibSVM algoritmaları olduğu belirlenmiştir. DVM ve LibSVM algoritmalarına ait detaylı sonuç verileri Tablo 7.14’de verilmiştir.

Tablo 7.14: VT1 için Weka uygulamasında çalıştırılan DVM ve LibSVM algoritmalarının doğal saç rengi sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
DVM	33	24	13	19	0,649	0,421	0,511	Kahverengi	%42,16
	11	7	26	36	0,212	0,389	0,275	Kumral	
	15	12	20	31	0,375	0,444	0,407	Siyah	
LibSVM	25	18	19	23	0,486	0,419	0,450	Kahverengi	%40,20
	16	9	24	32	0,273	0,360	0,310	Kumral	
	20	14	18	27	0,438	0,412	0,424	Siyah	

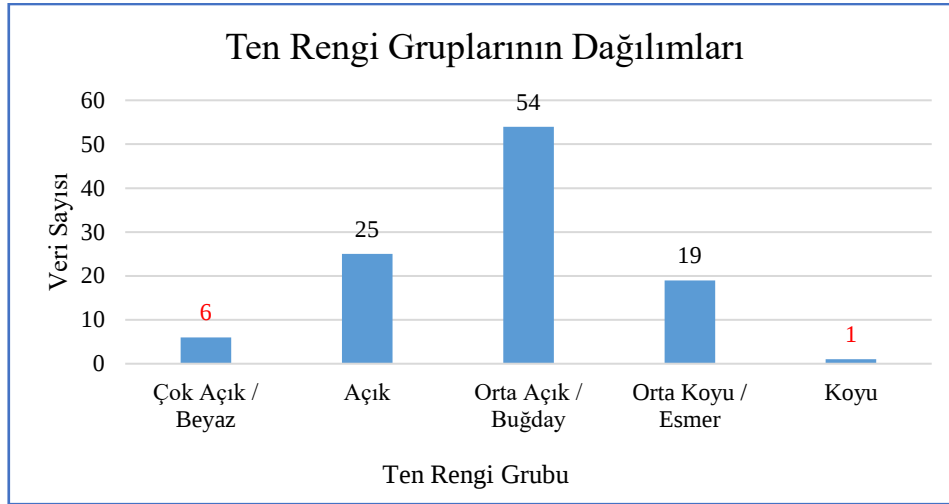
Doğal saç rengi tahmini için her iki uygulamadaki sınıflandırma algoritmalarının doğruluklarına bakıldığında, en yüksek doğruluğa sahip algoritmanın %45,91 ile RapidMiner uygulamasındaki Derin Öğrenme algoritması olduğu sonucuna ulaşılmıştır. Ayrıca Weka programındaki en başarılı algoritmanın %42,16 ile DVM algoritması olduğu görülmüştür. Her iki uygulamanın en başarılı algoritmalarının karşılaştırması Tablo 7.15’de görülmektedir.

Tablo 7.15: VT1 için RapidMiner ve Weka uygulamalarının doğal saç rengi tahmini üzerindeki en başarılı algoritmalarının karşılaştırılması.

Uygulama	Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	S. Başarısı
RapidMiner	Derin Öğrenme	22	19	18	28	0,514	0,463	0,487	Kahverengi	%45,91
		18	16	17	31	0,485	0,471	0,478	Kumral	
		15	12	20	35	0,375	0,444	0,407	Siyah	
Weka	DVM	33	24	13	19	0,649	0,421	0,511	Kahverengi	%42,16
		11	7	26	36	0,212	0,389	0,275	Kumral	
		15	12	20	31	0,375	0,444	0,407	Siyah	

7.4.6. VT1 İçin Ten Rengi Tahmini

Araştırma kapsamında katılımcılardan elde edilen anket verilerindeki ten rengi sınıfına bakıldığında, “Çok Açık / Beyaz” ve “Koyu / Kahverengi” grubuna ait %10’dan az sayıda veri olduğu görülmüştür. Bu nedenle sadece "Açık" "Orta Açık / Buğday" ve "Orta Koyu / Esmer" ten rengi grupları ile algoritma testleri gerçekleştirilmiştir. Ten rengi gruplarının veri dağılımları Şekil 7.15’de verilmiştir.



Şekil 7.15: VT1 için ten rengi gruplarının veri dağılımları

RapidMiner uygulaması içerisinde bulunan sınıflandırma algoritmaları ile ten rengi tahmini testi yapılması sonucunda, en başarılı algoritmaların %45,89 ile Derin Öğrenme ve %44,89 ile LibSVM algoritmaları olduğu görülmüştür. Bu algoritmaların detaylı sonuç bilgileri Tablo 7.16’da verilmiştir.

Tablo 7.16: VT1 için RapidMiner uygulamasında çalıştırılan Derin Öğrenme ve LibSVM algoritmalarının ten rengi sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	S. Başarısı
Derin Öğrenme	14	4	21	41	0,160	0,222	0,186	Açık	%45,89
	28	34	20	11	0,630	0,548	0,586	Orta Açık / Buğday	
	11	7	12	38	0,368	0,389	0,378	Orta Koyu / Esmer	
LibSVM	12	6	19	38	0,240	0,333	0,279	Açık	%44,89
	34	36	18	8	0,667	0,514	0,581	Orta Açık / Buğday	
	8	2	17	42	0,105	0,520	0,105	Orta Koyu / Esmer	

RapidMiner uygulamasında gerçekleştirilen ten rengi tahmini için kullanılan aynı veri seti ile Weka uygulamasında yapılan ten rengi tahmini için algoritmalar test edildiğinde, en başarılı algoritmaların %50,00 ile LibSVM ve %45,92 ile DVM algoritmaları olduğu sonucuna ulaşılmıştır. Bu sonuçların detaylı verileri Tablo 7.17’de verilmiştir.

Tablo 7.17: VT1 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının ten rengi sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	S. Başarısı
LibSVM	10	5	20	44	0,200	0,333	0,250	Açık	%50,00
	36	43	11	6	0,796	0,544	0,647	Orta Açık / Buğday	
	3	1	18	48	0,053	0,250	0,087	Orta Koyu / Esmer	
DVM	14	6	19	39	0,240	0,300	0,240	Açık	%45,92
	32	36	18	9	0,667	0,529	0,590	Orta Açık / Buğday	
	7	3	16	42	0,158	0,300	0,207	Orta Koyu / Esmer	

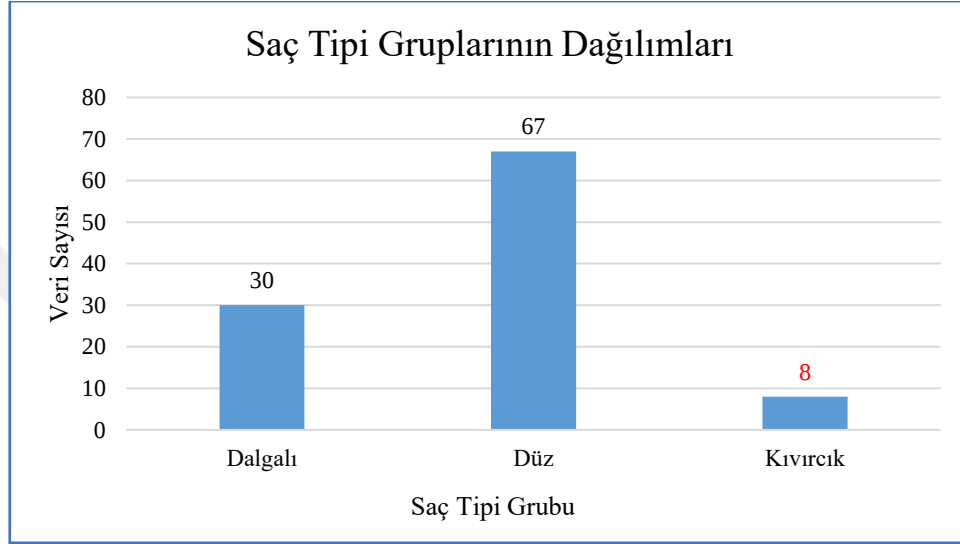
Algoritma testlerinin gerçekleştirildiği iki farklı uygulamanın algoritma testi doğruluklarına bakıldığında, en başarılı sınıflandırma algoritmasının %50,00 sınıflandırma başarıları ile Weka uygulamasında gerçekleştirilen LibSVM algoritması olduğu görülmüştür. RapidMiner uygulamasında ise en başarılı algoritma %45,89 başarı ile Derin öğrenme algoritması olmuştur. İki uygulamanın en başarılı algoritmalarının karşılaştırılması Tablo 7.18’de verilmiştir.

Tablo 7.18: VT1 için RapidMiner ve Weka uygulamalarının ten rengi tahmini üzerindeki en başarılı algoritmalarının karşılaştırılması.

Uygulama	Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	S. Başarısı
RapidMiner	Derin Öğrenme	14	4	21	41	0,160	0,222	0,186	Açık	%45,89
		28	34	20	11	0,630	0,548	0,586	Orta A. / B.	
		11	7	12	38	0,368	0,389	0,378	Orta K. / E.	
Weka	LibSVM	10	5	20	44	0,200	0,333	0,250	Açık	%50,00
		36	43	11	6	0,796	0,544	0,647	Orta A. / B.	
		3	1	18	48	0,053	0,250	0,087	Orta K. / E.	

7.4.7. VT1 İçin Saç Tipi Tahmini

Çalışmada kullanılan anket verilerinin saç tipi sınıfındaki veri dağılımları incelendiğinde, “kıvrıkcık” saç tipi grubuna ait az sayıda veri olduğu görülmüştür. Bu nedenle algoritma testlerine “kıvrıkcık” grubuna ait veriler dâhil edilmemiştir. Saç tipi gruplarının veri dağılımları Şekil 7.16’da verilmiştir.



Şekil 7.16: VT1 için saç tipi gruplarının veri dağılımları

RapidMiner uygulaması üzerindeki sınıflandırma algoritmalarının saç tipi tahmini üzerinden testleri gerçekleştirildiğinde, en başarılı algoritmaların %68,89 ile LibSVM ve %67,89 ile Rastgele Orman algoritmaları oldukları sonucuna ulaşılmıştır. Algoritmaların detaylı sonuçları Tablo 7.19’da verilmiştir.

Tablo 7.19: VT1 için RapidMiner uygulamasında çalıştırılan LibSVM ve Rastgele Orman algoritmalarının saç tipi sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
LibSVM	5	5	25	62	0,167	0,500	0,250	Dalgalı	%68,89
	25	62	5	5	0,925	0,713	0,805	Düz	
Rastgele Orman	2	1	29	1	0,033	0,333	0,060	Dalgalı	%67,89
	29	65	2	65	0,970	0,692	0,739	Düz	

Weka uygulaması üzerindeki algoritmalar ile yapılacak saç tipi sınıflandırması için yapılan algoritma testlerinde, en yüksek sınıflandırma başarısına sahip algoritmaların %68,04 ile LibSVM ve %63,92 ile DVM algoritmaları olduğu görülmüştür. Saç tipi için Weka

uygulamasında test edilen algoritmaların detaylı sonuçları Tablo 7.20’de verilmiştir.

Tablo 7.20: VT1 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının saç tipi sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
LibSVM	6	5	25	61	0,167	0,455	0,244	Dalgalı	%68,04
	25	61	6	5	0,910	0,709	0,797	Düz	
DVM	12	7	23	55	0,233	0,368	0,286	Dalgalı	%63,92
	23	55	12	7	0,821	0,705	0,759	Düz	

Her iki uygulamada yapılan algoritma testleri sonucunda, saç tipi tahmini için en başarılı algoritmanın %68,89 başarı oranı ile RapidMiner uygulamasındaki LibSVM algoritması olduğu sonucuna ulaşılmıştır. Weka uygulamasında ise aynı algoritmanın %66.67 başarı ile sınıflandırma yaptığı görülmüştür. Her iki uygulamada da en başarılı sonuca sahip LibSVM algoritmasının karşılaştırması Tablo 7.21’de verilmiştir.

Tablo 7.21: VT1 için RapidMiner ve Weka uygulamalarının saç tipi tahmini için en başarılı algoritmalarının karşılaştırılması.

Uygulama	Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	S. Başarısı
RapidMiner	LibSVM	5	5	25	62	0,167	0,500	0,250	Dalgalı	%68,89
		25	62	5	5	0,925	0,713	0,805	Düz	
Weka		6	5	25	61	0,167	0,455	0,244	Dalgalı	%68,04
		25	61	6	5	0,910	0,709	0,797	Düz	

7.5. VT2 İçin Tahmin İşlemleri

7.5.1. VT2 İçin Cinsiyet Tahmini

Çalışma kapsamında VT2 için elde edilen cinsiyet sınıfı verilerine bakıldığında, 231 erkek ve 69 kadına ait kulak görüntüsü olduğu görülmüştür.

Cinsiyet tahmini için RapidMiner uygulamasında yapılan testler sonucunda en başarılı iki algoritmanın sonuçları Tablo 7.22’de verilmiştir. Bu algoritmalar LibSVM ve DVM olup, doğruluk oranları sırasıyla %86 ve %85,33 olarak tespit edilmiştir.

Tablo 7.22: VT2 için RapidMiner uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının cinsiyet sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
LibSVM	0	27	42	231	0,391	1	0,562	Kadın	%86,00
	42	231	0	27	1	0,846	0,835	Erkek	
DVM	10	35	34	221	0,507	0,778	0,614	Kadın	%85,33
	34	221	10	35	0,957	0,867	0,910	Erkek	

Cinsiyet tahmini için aynı veri seti kullanılarak Weka uygulaması üzerinde yapılan testler sonucunda en başarılı algoritmalar %86,33 sınıflandırma başarısı ile LibSVM ve DVM algoritmaları olduğu görülmüştür. Bu algoritmaların detaylı sonuçları Tablo 7.23’de verilmiştir.

Tablo 7.23: VT2 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının cinsiyet sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
LibSVM	16	44	25	215	0,638	0,733	0,682	Kadın	%86,33
	25	215	16	44	0,931	0,896	0,913	Erkek	
DVM	25	53	16	206	0,768	0,679	0,721	Kadın	
	16	206	25	53	0,892	0,928	0,909	Erkek	

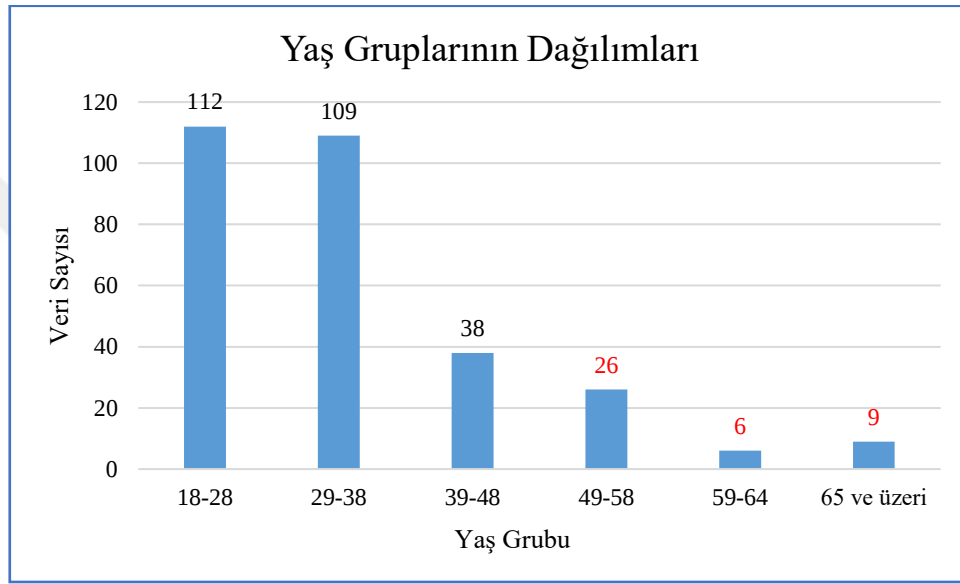
İki farklı uygulamada sınıflandırma algoritmaları ile yapılan testler sonucunda, cinsiyet tahmininde en başarılı algoritmanın %86,33 sınıflandırma başarısı ile Weka uygulamasında bulunan LibSVM ve DVM algoritmaları olduğu görülmüştür. RapidMiner uygulaması üzerindeki en başarılı algoritma ise %86 doğruluk oranı ile LibSVM algoritması olmuştur. Her iki uygulamanın cinsiyet tahmini için en başarılı algoritmalarının karşılaştırılması Tablo 7.24’de görülebilir.

Tablo 7.24: VT2 için RapidMiner ve Weka uygulamalarının cinsiyet tahmini için en başarılı algoritmalarının karşılaştırılması.

Uygulama	Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	S. Başarısı
RapidMiner	LibSVM	0	27	42	231	0,391	1	0,562	Kadın	%86,00
		42	231	0	27	1	0,846	0,835	Erkek	
Weka	LibSVM	16	44	25	215	0,638	0,733	0,682	Kadın	%86,33
		25	215	16	44	0,931	0,896	0,913	Erkek	
	DVM	25	53	16	206	0,768	0,679	0,721	Kadın	
		16	206	25	53	0,892	0,928	0,909	Erkek	

7.5.2. VT2 İçin Yaş Tahmini

Çalışma kapsamında toplanılan verilere bakıldığında, yaş sınıfında altı farklı yaş aralığının bulunduğu görülmektedir. Bu yaş aralıkları arasından “49-58”, “59-65” ve “65 ve üzeri” grubunda %10’dan az veri olduğu görülmüştür. Bu nedenle yaş tahmini yapılırken en çok verinin bulunduğu üç grup olan “18-28”, “29-38” ve “39-48” yaş grupları testlere dâhil edilmiştir. Yaş gruplarının dağılımları Şekil 7.17’de verilmiştir.



Şekil 7.17: VT2 için yaş gruplarının veri dağılımları

RapidMiner uygulamasında yapılan algoritma testleri sonucunda, en başarılı algoritmaların sırasıyla %68,40 LibSVM ve %68,37 ile Derin Öğrenme olduğu görülmüştür. Sınıflandırma sonuçlarının detayları Tablo 7.25’de verilmiştir.

Tablo 7.25: VT2 için RapidMiner uygulamasında çalıştırılan LibSVM ve Derin Öğrenme algoritmalarının yaş sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
LibSVM	41	87	24	90	0,861	0,633	0,730	18-28	%68,40
	33	73	37	103	0,664	0,689	0,676	29-38	
	8	17	21	160	0,447	0,680	0,540	39-48	
Derin Öğrenme	41	83	28	94	0,748	0,669	0,706	18-28	%68,37
	32	76	34	101	0,691	0,704	0,697	29-38	
	9	18	20	159	0,474	0,667	0,554	39-48	

Yaş tahmini için aynı veri seti kullanılarak Weka uygulamasında yapılan algoritma testlerinde en yüksek doğruluk oranlarına sahip algoritmaların %61,78 sınıflandırma başarısı ile LibSVM ve %60,23 sınıflandırma başarısıyla Lojistik Regresyon algoritmaları olduğu görülmüştür. Bu algoritmalara ait detaylı test sonuçları Tablo 7.26’da verilmiştir.

Tablo 7.26: VT2 için Weka uygulamasında çalıştırılan LibSVM ve Lojistik Regresyon algoritmalarının yaş sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
LibSVM	47	76	35	84	0,685	0,618	0,650	18-28	%61,78
	42	71	39	89	0,645	0,628	0,637	29-38	
	10	13	25	147	0,342	0,565	0,426	39-48	
Lojistik Regresyon	47	76	35	80	0,685	0,618	0,650	18-28	%60,23
	45	69	41	87	0,627	0,605	0,616	29-38	
	11	11	27	145	0,289	0,500	0,594	39-48	

İki farklı uygulamada sınıflandırma algoritmaları ile yapılan testler sonucunda, yaş tahmininde doğruluk oranı en yüksek olan algoritmanın %68,40 doğruluk oranı ile RapidMiner uygulamasında bulunan LibSVM algoritması olduğu görülmüştür. Ayrıca Weka uygulamasındaki en başarılı algoritmanın %61,78 ile LibSVM algoritması olduğu görülmüştür. Yaş sınıflandırması açısından her iki uygulamanın en başarılı algoritmalarının karşılaştırılması Tablo 7.27’de görülebilir.

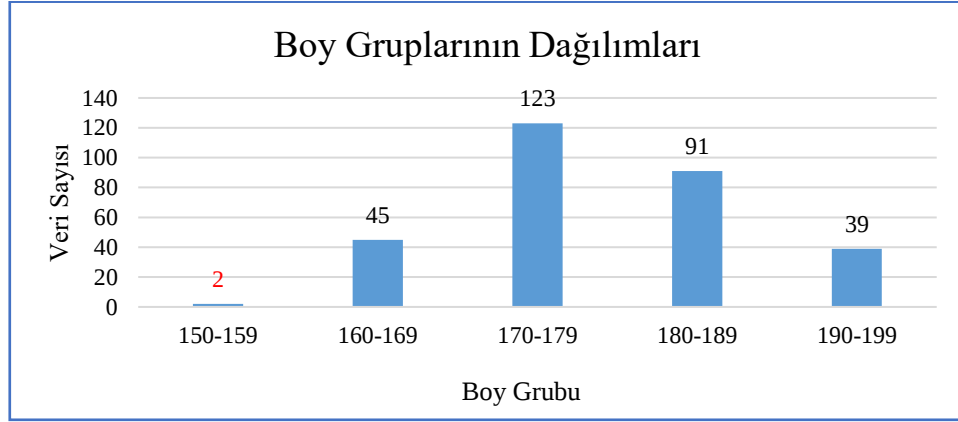
Tablo 7.27: VT2 için RapidMiner ve Weka uygulamalarının yaş tahmini üzerindeki en başarılı algoritmalarının karşılaştırılması.

Uygulama	Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	S. Başarısı
RapidMiner	LibSVM	41	87	24	90	0,861	0,633	0,730	18-28	%68,40
		33	73	37	103	0,664	0,689	0,676	29-38	
		8	17	21	160	0,447	0,680	0,540	39-48	
Weka	LibSVM	47	76	35	84	0,685	0,618	0,650	18-28	%61,78
		42	71	39	89	0,645	0,628	0,637	29-38	
		10	13	25	147	0,342	0,565	0,426	39-48	

7.5.3. VT2 İçin Boy Tahmini

Araştırma kapsamında toplanılan veriler incelendiğinde, boy sınıfında “150-159” grubunda %10’dan az sayıda veri bulunduğu görülmüştür. Bu nedenle algoritma testlerine bu gruplar dâhil edilmemiştir ve algoritma testlerine “160-169”, “170-179”, “180-189” ve “190-199”

boy grupları ile testler gerçekleştirilmiştir. Boy gruplarının dağılımları Şekil 7.14’de verilmiştir.



Şekil 7.18: VT2 için boy gruplarının veri dağılımları

Boy tahmini için yapılan algoritma testleri sonucunda, RapidMiner uygulaması üzerindeki en başarılı algoritmaların %60,06 ile Rastgele Orman ve %58,10 ile Derin Öğrenme algoritmaları olduğu görülmüştür. Algoritmaların detaylı sonuçları Tablo 7.28’de verilmiştir.

Tablo 7.28: VT2 için RapidMiner uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının boy sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
Rastgele Orman	0	18	27	161	0,400	1	0,571	160-169	%60,06
	95	108	14	71	0,885	0,532	0,665	170-179	
	24	33	58	136	0,363	0,579	0,446	180-189	
	0	20	20	159	0,500	1	0,667	190-199	
Derin Öğrenme	13	23	22	150	0,511	0,639	0,568	160-169	%58,10
	55	77	45	96	0,631	0,583	0,606	170-179	
	47	53	38	120	0,582	0,530	0,555	180-189	
	10	20	20	153	0,500	0,667	0,572	190-199	

Boy tahmini için Weka uygulamasındaki sınıflandırma algoritmaları ile yapılan testlerin sonucunda, en başarılı algoritmaların %60,74 ile LibSVM ve %56,04 ile DVM algoritmaları oldukları ortaya çıkmıştır. Bu algoritmaların detaylı sonuçları Tablo 7.29’da görülebilmektedir.

Tablo 7.29: VT2 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının boy sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
LibSVM	1	20	25	161	0,444	0,952	0,606	160-169	%60,74
	75	96	26	85	0,787	0,561	0,655	170-179	
	38	49	42	132	0,538	0,563	0,551	180-189	
	3	16	24	165	0,400	0,842	0,542	190-199	
DVM	13	25	20	142	0,556	0,658	0,602	160-169	%56,04
	64	78	44	89	0,639	0,549	0,591	170-179	
	43	43	48	124	0,473	0,500	0,486	180-189	
	11	21	19	146	0,525	0,656	0,583	190-199	

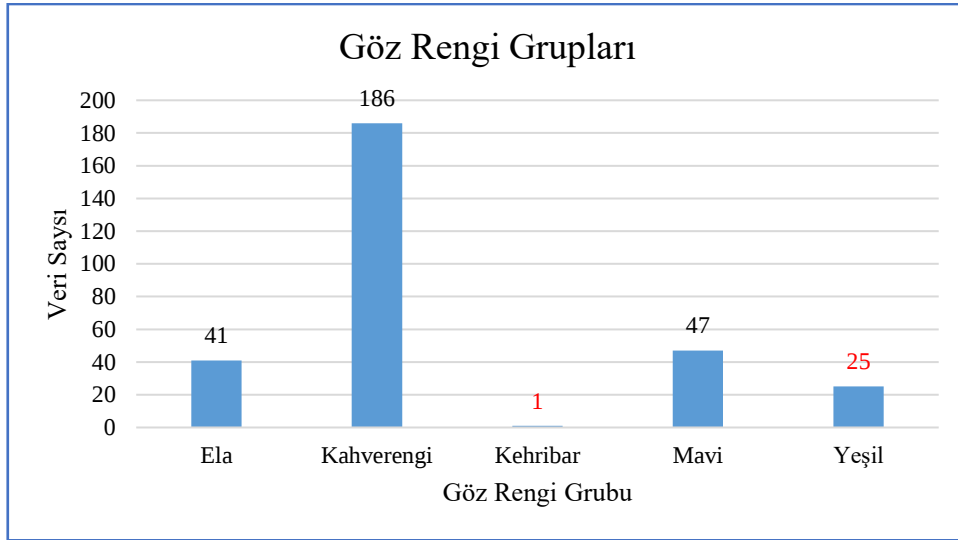
Çalışma kapsamında kullanılan iki farklı uygulama karşılaştırıldığında, en başarılı algoritmaların %60,74 ile Weka uygulamasında bulunan LibSVM algoritması olduğu görülmüştür. RapidMiner uygulamasında kullanılan algoritmalar arasında ise %60,06’lık bir sınıflandırma başarısı ile Rastgele Orman algoritması en başarılı algoritma olmuştur. İki farklı uygulamanın, yapılan testler sonucunda boy sınıflandırması açısından en başarılı algoritmalarının karşılaştırılması Tablo 7.30’da verilmiştir.

Tablo 7.30: VT2 için RapidMiner ve Weka uygulamalarının boy tahmini üzerindeki en başarılı algoritmalarının karşılaştırılması.

Uygulama	Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	S. Başarısı
RapidMiner	Rastgele Orman	0	18	27	161	0,400	1	0,571	160-169	%60,16
		95	108	14	71	0,885	0,532	0,665	170-179	
		24	33	58	136	0,363	0,579	0,446	180-189	
		0	20	20	159	0,500	1	0,667	190-199	
Weka	LibSVM	1	20	25	161	0,444	0,952	0,606	160-169	%60,74
		75	96	26	85	0,787	0,561	0,655	170-179	
		38	49	42	132	0,538	0,563	0,551	180-189	
		3	16	24	165	0,400	0,842	0,542	190-199	

7.5.4. VT2 İçin Göz Rengi Tahmini

Yapılan çalışma sonucunda elde edilen verilerdeki göz rengi sınıfı incelenmiştir. İnceleme sonucunda “Yeşil” ve “Kehribar” göz rengine sahip çok %10’dan daha az veri olduğu görülmüştür. Bu nedenle sınıflandırma algoritmaları bu grupları tahmin etmemektedir. Bu nedenle göz rengi tahmininde sadece “Kahverengi”, “Mavi” ve “Ela” göz renklerinin tahmini yapılmıştır. Göz rengi gruplarının dağılımları Şekil 7.15’de verilmiştir.



Şekil 7.19: VT2 için yaş gruplarının veri dağılımları

Göz rengi ile ilgili RapidMiner uygulaması üzerinde bulunan sınıflandırma algoritmaları ile yapılan testler sonucunda, en yüksek doğruluğa sahip algoritmaların %79,17 ile LibSVM ve %75,15 ile Derin Öğrenme algoritmaları olduğu tespit edilmiştir. Bu algoritmaların detaylı sonuçları Tablo 7.31’de verilmiştir.

Tablo 7.31: VT2 için RapidMiner uygulamasında çalıştırılan LibSVM ve Derin Öğrenme algoritmalarının göz rengi sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
LibSVM	57	186	0	30	1	0,765	0,867	Kahverengi	%79,17
	0	16	31	200	0,340	1	0,508	Mavi	
	0	14	26	202	0,350	1	0,519	Ela	
Derin Öğrenme	51	173	13	32	0,930	0,772	0,844	Kahverengi	%75.15
	9	17	30	188	0,368	0,654	0,471	Mavi	
	8	15	25	190	0,375	0,652	0,476	Ela	

RapidMiner uygulamasında kullanılan aynı veri seti ile Weka uygulamasındaki sınıflandırma algoritmaları test edildiğinde, sınıflandırma başarısı en yüksek algoritmaların %80,59 ile LibSVM ve DVM algoritmaları olduğu görülmüştür. Bu algoritmaların detaylı sonuç bilgileri Tablo 7.32’de gösterilmektedir.

Tablo 7.32: VT2 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının göz rengi sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
LibSVM	48	183	3	27	0,984	0,792	0,878	Kahverengi	%80,59
	2	19	28	201	0,404	0,905	0,559	Mavi	
	3	18	22	202	0,450	0,857	0,590	Ela	
DVM	53	186	0	32	1	0,778	0,875	Kahverengi	
	0	18	29	202	0,383	1	0,554	Mavi	
	0	16	24	204	0,400	1	0,571	Ela	

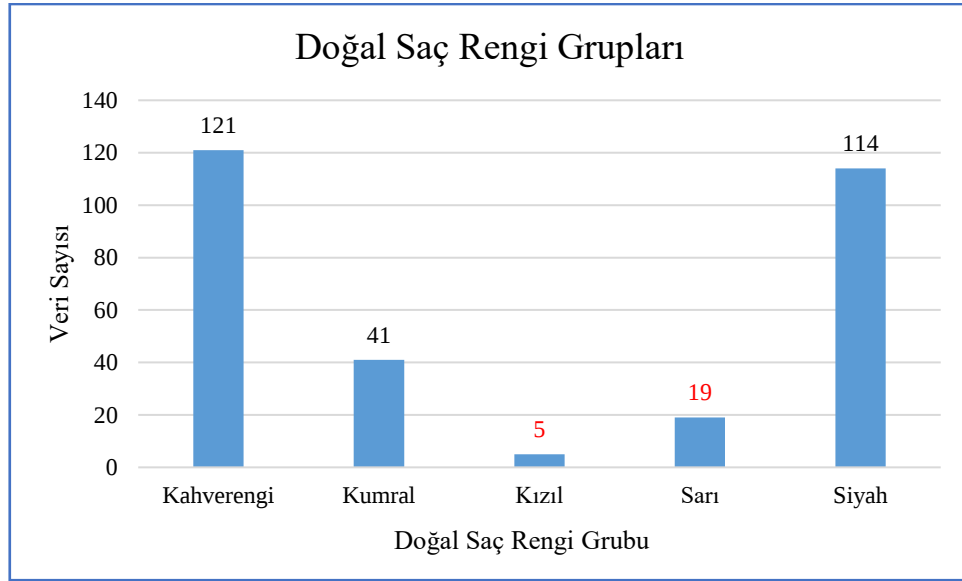
Göz rengi tahmini için her iki uygulamadaki algoritma testi sonuçlarına bakıldığında, Weka uygulamasındaki LibSVM ve DVM algoritmalarının %80,59 doğruluk ile göz rengi sınıflandırması için en başarılı algoritmalar olduğu belirlenmiştir. RapidMiner uygulamasında yapılan testlerde ise göz rengi tahmini için en başarılı algoritmanın %79,17 doğruluk ile LibSVM olduğu görülmüştür. Her iki uygulamada da test edilen algoritmaların iki uygulama bazında karşılaştırılması Tablo 7.33’de verilmiştir.

Tablo 7.33: VT2 için RapidMiner ve Weka uygulamalarının göz rengi tahmini için en başarılı algoritmalarının karşılaştırılması.

Uygulama	Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	S. Başarısı
RapidMiner	LibSVM	57	186	0	30	1	0,765	0,867	Kahverengi	%79,17
		0	16	31	200	0,340	1	0,508	Mavi	
		0	14	26	202	0,350	1	0,519	Ela	
Weka	LibSVM	48	183	3	27	0,984	0,792	0,878	Kahverengi	%80,59
		2	19	28	201	0,404	0,905	0,559	Mavi	
		3	18	22	202	0,450	0,857	0,590	Ela	
	DVM	53	186	0	32	1	0,778	0,875	Kahverengi	
		0	18	29	202	0,383	1	0,554	Mavi	
		0	16	24	204	0,400	1	0,571	Ela	

7.5.5. VT2 İçin Doğal Saç Rengi Tahmini

Araştırma kapsamında elde edilen verilerin doğal saç rengi sınıfı veri dağılımı incelendiğinde, “kızıl” ve “sarı” doğal saç rengine sahip verilerin %10’dan daha az olduğu görülmüştür. Bu durum göz önünde bulundurularak “sarı” ve “kızıl” saç rengine ait değerler algoritma testlerine dâhil edilmeyip, “kahverengi”, “siyah” ve “kumral” saç rengi grupları kullanılmıştır. Doğal saç rengi gruplarının veri sayısı dağılımları Şekil 7.16’da verilmiştir.



Şekil 7.20: VT2 için doğal saç rengi gruplarının veri dağılımları

Çalışmada tahmin edilmek istenen doğal saç rengi için RapidMiner uygulaması üzerindeki sınıflandırma algoritmaları test edildiğinde, en başarılı algoritmaların %57,62 ile Derin Öğrenme ve %56,16 ile DVM algoritmaları olduğu belirlenmiştir. Belirlenen algoritmalara ait detaylı sonuç verileri Tablo 7.34’de gösterilmiştir.

Tablo 7.34: VT2 için RapidMiner uygulamasında çalıştırılan Derin Öğrenme ve DVM algoritmalarının doğal saç rengi sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
Derin Öğrenme	54	77	44	82	0,636	0,588	0,611	Kahverengi	%57,62
	49	74	40	85	0,649	0,602	0,625	Siyah	
	14	8	33	151	0,195	0,364	0,254	Kumral	
DVM	108	111	10	44	0,917	0,507	0,487	Kahverengi	%56,16
	13	38	76	117	0,333	0,745	0,361	Siyah	
	0	6	35	149	0,146	1	0,431	Kumral	

RapidMiner uygulamasında yapılan algoritma testlerinde kullanılan aynı veri seti ile Weka uygulamasında yapılan algoritma testleri sonucunda, doğal saç rengi için en başarılı algoritmaların %57,25 ile DVM ve %55,44 ile LibSVM algoritmaları olduğu belirlenmiştir. DVM ve LibSVM algoritmalarına ait detaylı sonuç verileri Tablo 7.35’de verilmiştir.

Tablo 7.35: VT2 için Weka uygulamasında çalıştırılan DVM ve LibSVM algoritmalarının doğal saç rengi sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
DVM	101	108	13	50	0,893	0,517	0,655	Kahverengi	%57,25
	17	42	72	116	0,368	0,712	0,486	Siyah	
	0	8	33	150	0,195	1	0,486	Kumral	
LibSVM	61	79	42	74	0,653	0,564	0,605	Kahverengi	%55,44
	46	63	51	90	0,553	0,578	0,565	Siyah	
	16	11	30	142	0,268	0,407	0,324	Kumral	

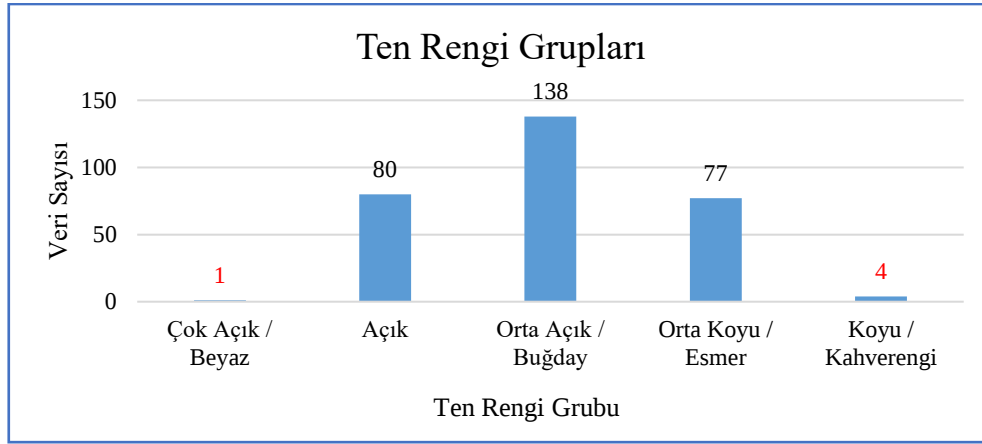
Doğal saç rengi tahmini için her iki uygulamadaki sınıflandırma algoritmalarının doğruluklarına bakıldığında, en yüksek doğruluğa sahip algoritmanın %57,62 ile RapidMiner uygulamasındaki Derin Öğrenme algoritması olduğu sonucuna ulaşılmıştır. Ayrıca Weka programındaki en başarılı algoritmanın %57,25 ile DVM algoritması olduğu görülmüştür. Her iki uygulamanın en başarılı algoritmalarının karşılaştırması Tablo 7.36’da görülmektedir.

Tablo 7.36: VT2 için RapidMiner ve Weka uygulamalarının doğal saç rengi tahmini üzerindeki en başarılı algoritmalarının karşılaştırılması.

Uygulama	Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	S. Başarısı
RapidMiner	Derin Öğrenme	54	77	44	82	0,636	0,588	0,611	Kahverengi	%57,62
		49	74	40	85	0,649	0,602	0,625	Siyah	
		14	8	33	151	0,195	0,364	0,254	Kumral	
Weka	DVM	101	108	13	50	0,893	0,517	0,655	Kahverengi	%57,25
		17	42	72	116	0,368	0,712	0,486	Siyah	
		0	8	33	150	0,195	1	0,486	Kumral	

7.5.6. VT2 İçin Ten Rengi Tahmini

Araştırma kapsamında elde edilen verilerdeki ten rengi sınıfına bakıldığında, “Çok Açık / Beyaz” ve “Koyu / Kahverengi” grubuna ait %10’dan az sayıda veri olduğu görülmüştür. Bu nedenle sadece "Açık" "Orta Açık / Buğday" ve "Orta Koyu / Esmer" ten rengi grupları ile algoritma testleri gerçekleştirilmiştir. Ten rengi gruplarının veri sayısı dağılımları Şekil 7.17’de verilmiştir.



Şekil 7.21: VT2 için ten rengi gruplarının veri dağılımları

RapidMiner uygulaması içerisinde bulunan sınıflandırma algoritmaları ile ten rengi tahmini testi yapılması sonucunda, en başarılı algoritmaların %60,67 ile LibSVM ve %60,37 ile Rastgele Orman algoritmaları olduğu görülmüştür. Bu algoritmaların detaylı sonuç bilgileri Tablo 7.37’de verilmiştir.

Tablo 7.37: VT2 için RapidMiner uygulamasında çalıştırılan LibSVM ve Rastgele Orman algoritmalarının ten rengi sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	S. Başarısı
LibSVM	0	20	60	159	0,250	1	0,400	Açık	%60,67
	116	138	0	41	1	0,543	0,704	Orta Açık / Buğday	
	0	21	56	158	0,273	1	0,429	Orta Koyu / Esmer	
Rastgele Orman	9	23	57	155	0,288	0,712	0,410	Açık	%60,37
	101	129	9	49	0,935	0,561	0,701	Orta Açık / Buğday	
	7	26	51	152	0,338	0,788	0,473	Orta Koyu / Esmer	

RapidMiner uygulamasında gerçekleştirilen ten rengi tahmini için kullanılan aynı veri seti ile Weka uygulamasında yapılan ten rengi tahmini için algoritmalar test edildiğinde, en başarılı algoritmaların %60,68 ile LibSVM ve %60 ile DVM algoritmaları olduğu sonucuna ulaşılmıştır. Bu sonuçların detaylı verileri Tablo 7.38’de verilmiştir.

Tablo 7.38: VT2 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının ten rengi sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	S. Başarısı
LibSVM	20	31	49	148	0,388	0,608	0,473	Açık	%60,68
	73	114	24	65	0,826	0,610	0,702	Orta Açık / Buğday	
	23	34	43	145	0,442	0,596	0,507	Orta Koyu / Esmer	
DVM	0	20	60	157	0,250	1	0,400	Açık	%60,00
	118	138	0	39	1	0,539	0,701	Orta Açık / Buğday	
	0	19	58	158	0,247	1	0,396	Orta Koyu / Esmer	

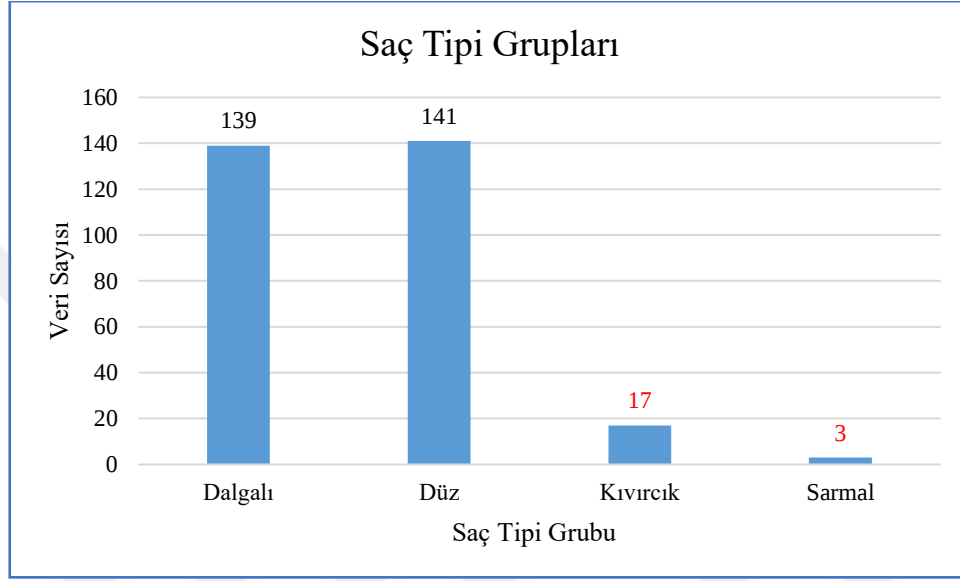
Algoritma testlerinin gerçekleştirildiği iki farklı uygulamanın algoritma testi doğruluklarına bakıldığında, en başarılı sınıflandırma algoritmasının %60,68 sınıflandırma başarısı ile Weka uygulamasında gerçekleştirilen LibSVM algoritması olduğu görülmüştür. RapidMiner uygulamasında ise en başarılı algoritma %60,67 başarı ile LibSVM algoritması olmuştur. İki uygulamanın en başarılı algoritmalarının karşılaştırılması Tablo 7.39’da verilmiştir.

Tablo 7.39: VT2 için RapidMiner ve Weka uygulamalarının ten rengi tahmini üzerindeki en başarılı algoritmalarının karşılaştırılması.

Uygulama	Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	S. Başarısı
RapidMiner	LibSVM	0	20	60	159	0,250	1	0,400	Açık	%60,67
		116	138	0	41	1	0,543	0,704	Orta A. / B.	
		0	21	56	158	0,273	1	0,429	Orta K. / E.	
Weka	LibSVM	20	31	49	148	0,388	0,608	0,473	Açık	%60,68
		73	114	24	65	0,826	0,610	0,702	Orta A. / B.	
		23	34	43	145	0,442	0,596	0,507	Orta K. / E.	

7.5.7. VT2 İçin Saç Tipi Tahmini

Çalışmada kullanılan veri tabanındaki saç tipi sınıfındaki veri dağılımları incelendiğinde, “kıvrırcık” ve “sarmal” saç tipi grubuna ait verilerin %10’dan daha az sayıda olduğu görülmüştür. Bu nedenle algoritma testlerine “kıvrırcık” ve “sarmal” grularına ait veriler dâhil edilmemiştir. Saç tipi gruplarının veri sayısı dağılımları Şekil 7.18’de verilmiştir.



Şekil 7.22: VT2 için saç tipi rengi gruplarının veri dağılımları

RapidMiner uygulaması üzerindeki sınıflandırma algoritmalarının saç tipi tahmini üzerinden testleri gerçekleştirildiğinde, en başarılı algoritmaların %69,06 ile LibSVM ve %65,50 ile Rastgele Orman algoritmaları oldukları sonucuna ulaşılmıştır. Algoritmaların detaylı sonuçları Tablo 7.40’da verilmiştir.

Tablo 7.40: VT2 için RapidMiner uygulamasında çalıştırılan LibSVM ve Rastgele Orman algoritmalarının saç tipi sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
LibSVM	0	52	87	142	0,374	1	0,544	Dalgalı	%69,06
	87	142	0	52	1	0,620	0,765	Düz	
Rastgele Orman	42	84	55	100	0,604	0,667	0,634	Dalgalı	%65,50
	55	100	42	84	0,704	0,645	0,673	Düz	

Weka uygulaması üzerindeki algoritmalar ile yapılacak saç tipi sınıflandırması için yapılan algoritma testlerinde, en yüksek sınıflandırma başarısına sahip algoritmaların %66,55 ile LibSVM ve %65,48 ile DVM algoritmaları olduğu görülmüştür. Saç tipi için Weka uygulamasında test edilen algoritmaların detaylı sonuçları Tablo 7.41’de verilmiştir.

Tablo 7.41: VT2 için Weka uygulamasında çalıştırılan LibSVM ve DVM algoritmalarının saç tipi sınıflandırma sonucu.

Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	Sınıflandırma Başarısı
LibSVM	45	90	45	97	0,647	0,667	0,657	Dalgalı	%66,55
	49	97	49	90	0,383	0,664	0,674	Düz	
DVM	48	90	49	94	0,647	0,652	0,650	Dalgalı	%65,48
	49	94	48	90	0,662	0,657	0,660	Düz	

Her iki uygulamada yapılan algoritma testleri sonucunda, saç tipi tahmini için en başarılı algoritmanın %69,06 başarı oranı ile RapidMiner uygulamasındaki LibSVM algoritması olduğu sonucuna ulaşılmıştır. Weka uygulamasında ise aynı algoritmanın %66,55 başarı ile sınıflandırma yaptığı görülmüştür. Her iki uygulamada da en başarılı sonuca sahip LibSVM algoritmasının karşılaştırması Tablo 7.42’de verilmiştir.

Tablo 7.42: VT2 için RapidMiner ve Weka uygulamalarının saç tipi tahmini için en başarılı algoritmalarının karşılaştırılması.

Uygulama	Algoritma	Yanlış Pozitif	Doğru Pozitif	Yanlış Negatif	Doğru Negatif	Recall	Precision	F1 Skoru	Sınıf	S. Başarısı
RapidMiner	LibSVM	0	52	87	142	0,374	1	0,544	Dalgalı	%69,06
		87	142	0	52	1	0,620	0,765	Düz	
Weka		45	90	45	97	0,647	0,667	0,657	Dalgalı	%66,55
		49	97	49	90	0,383	0,664	0,674	Düz	

8- SONUÇ VE ÖNERİLER

Çalışma kapsamında görüntü işleme teknikleriyle oluşturulan kulak veri seti kullanılarak, kulak görüntüleri alınan kişilere ait anketlerde yer alan soft biyometrik özelliklerin yapay zekâ teknikleri ile tahmin edilmesi amacıyla RapidMiner ve Weka veri işleme uygulamaları üzerinde çeşitli sınıflandırma algoritmaları kullanılarak testler yapılmıştır. Her bir soft biyometrik için yapılan çeşitli sınıflandırma testleri sonucunda, soft biyometriklerle ait en yüksek sınıflandırma başarısına sahip algoritmaların hangi uygulamadan elde edildiği ve sınıflandırma başarısı Tablo 8.1’de verilmiştir.

Tablo 8.1: Soft biyometriklerin sınıflandırma sonuçları.

Tahmin Grubu	VERİ TABANI 1			VERİ TABANI 2		
	Sınıflandırma Algoritması	Uygulama	Sınıflandırma Başarısı	Sınıflandırma Algoritması	Uygulama	Sınıflandırma Başarısı
Cinsiyet	LibSVM	Weka	<u>%84,76</u>	LibSVM ve DVM	Weka	<u>%86,33</u>
Yaş	LibSVM	RapidMiner	<u>%40,00</u>	LibSVM	RapidMiner	<u>%68,40</u>
Boy	LibSVM	RapidMiner	<u>%44,82</u>	LibSVM	Weka	<u>%60,74</u>
Göz Rengi	Derin Öğrenme	RapidMiner	<u>%68,47</u>	LibSVM	Weka	<u>%80,59</u>
Doğal Saç Rengi	Derin Öğrenme	RapidMiner	<u>%45,91</u>	Derin Öğrenme	RapidMiner	<u>%57,62</u>
Ten Rengi	LibSVM	Weka	<u>%50,00</u>	LibSVM	Weka	<u>%60,68</u>
Saç Tipi	LibSVM	RapidMiner	<u>%68,89</u>	LibSVM	RapidMiner	<u>%69,06</u>

Çalışmada kullanılan sınıflandırma algoritmalarının testlerine bakıldığında, en başarılı algoritmaların LibSVM, Derin Öğrenme ve DVM algoritmaları oldukları görülmektedir. Ayrıca aynı algoritmanın farklı uygulamalarda kullanılması ile de farklı sonuçlar ortaya çıkabileceği gözlemlenmiştir. Yapılan çalışma sonucunda elde edilen sınıflandırma başarısı sonuçları ile yakın çalışmaların sonuçlarının karşılaştırılmasına Tablo 8.2’de yer verilmiştir.

Tablo 8.2: Literatürdeki yakın çalışmaların yapılan çalışma ile karşılaştırması.

Referans	Veri Seti Boyutu	Cinsiyet Tahmin Başarısı	Yaş Tahmin Başarısı	Boy Tahmin Başarısı	Göz Rengi Tahmin Başarısı	Saç Rengi Tahmin Başarısı	Ten Rengi Tahmin Başarısı	Saç Tipi Tahmin Başarısı
Yapılan Çalışma VT1	105	%84,76	%40,00	%44,82	%68,47	%45,91	%50,00	%68,89
Yapılan Çalışma VT2	300	%86,33	%68,40	%60,74	%80,59	%57,62	%60,68	%69,06
Yaman vd. (2018)	270	%94,00	%52,00	İlgili özelliklerin tahmini yapılmamıştır.				
Nguyen-Quoc vd. (2020)	28412	%80,36	İlgili özelliklerin tahmini yapılmamıştır.					
Karasulu vd. (2022)	28412	%85,16	İlgili özelliklerin tahmini yapılmamıştır.					
Karasulu vd. (2022)	2071	%87,61	İlgili özelliklerin tahmini yapılmamıştır.					
Zhang ve Whang (2011)	28412	%97.75	İlgili özelliklerin tahmini yapılmamıştır.					
Khorsandi ve Abdel-Mottaleb (2013)	2340	%89,49	İlgili özelliklerin tahmini yapılmamıştır.					
Gnanasivam ve Muttan (2013)	342	%90,42	İlgili özelliklerin tahmini yapılmamıştır.					

Cinsiyet tahmini için yapılan testlere bakıldığında, VT1 için LibSVM algoritması %84,76 doğruluk oranı ile doğru sınıflandırma yapabilmektedir. Buna ek olarak VT2 ile yapılan testler sonucunda da LibSVM algoritmasının %86,33 ile cinsiyet sınıflandırması yapabildiği görülmüştür. Literatür göz önünde bulundurulduğunda, cinsiyetin doğru tahmin edilme oranının literatürdeki sonuçlara benzerliğinden ötürü, sistemin cinsiyet tahmini için başarılı sonuçlar üretebileceği görülmüştür. Literatürde kulak görüntüsünden cinsiyet sınıflandırma başarısı daha yüksek olan beş çalışma vardır. Fakat bu çalışmaların az da olsa daha iyi olmasının nedenleri arasında özellikle kulak veri setlerinin çok daha fazla veriye sahip olmasından dolayı gerçekleştirdikleri modelin daha iyi eğitilmesi gösterilebilir.

Kulak görüntüsünden cinsiyet ve yaş tahmin eden çalışmalar az da olsa olmakla birlikte, saç tipi, doğal saç rengi, boy ve göz rengi tahmini ile ilgili herhangi bir çalışmaya rastlanmamıştır. Bu bakımdan tez çalışmamız literatürde ilk kez kulak görüntüsünden saç tipi, doğal saç rengi, boy ve göz rengi tahmin sonuçlarını içermektedir.

Çalışmada VT1 ile yapılan yaş, boy ve doğal saç rengi tahminlerinde, veri kümesinin az sayıda olması ve sınıflardaki değişkenlere sahip kişi sayısının az olması sebebiyle diğer gruplara nazaran daha düşük sonuçlar çıkmıştır. Bunu destekler bir sonuç olarak VT2 ile yapılan testlerde bu sonuçların yükseldiği görülebilmektedir.

VT1 ile yapılan algoritma testleri sonucunda saç tipi ve göz rengi gruplarının tahminlerinin sırasıyla %68,89 ve %68,47 oldukları görülmektedir. Bu grupların tahmini yaş, boy ve doğal saç rengi gruplarından daha başarılı olmasının sebebinin daha sınıflarda yer alan değişken sayısının az olması gösterilebilir.

Algoritma testlerinin yapıldığı uygulamalar ile ilgili sonuçlar göz önünde bulundurulduğunda, iki farklı uygulamada aynı algoritma kullanılsa da farklı sonuçlar alınabildiği görülmüştür. Bunun sebebinin uygulamaların algoritmaları uyarlama yöntemlerinin ufak farklılıklar gösterse bile sonuçlarda değişiklikler yaratması olduğu düşünülmektedir. Bu sonucu destekler bir çalışma olarak Foozy vd. (2017) çalışmalarında, şüpheli kısa mesajları tespit için Weka ve Rapidminer programlarını kullanarak sınıflandırma yapmışlardır. Yaptıkları çalışma sonucunda Weka ve Rapidminer programlarının aynı algoritmalarda farklı sonuçlar verebildiklerini görmüşlerdir.

Çalışma kapsamında oluşturulan veri tabanları ile ilgili sonuçlar incelendiğinde, daha fazla veriye sahip ikinci veri tabanı ile yapılan testlerin daha yüksek sınıflandırma başarılarına sahip oldukları görülmüştür. Bu sonuçtan yola çıkarak; gelecek çalışmalarda, veri setinin her bir veri grubu soft biyometrik için artırılıp, veri grupları içerisindeki sınıf dağılımlarının eşit sağlanması önerilmektedir. Bu sayede bütün gruplarda daha yüksek sınıflandırma başarısına sahip olunabilir. Ayrıca, ön işlemlerin sayısının artırılması önerilmektedir. Bu sayede bütün kulak görüntüleri arasındaki ortamdan ve diğer etmenlerden oluşabilecek problemlerin önüne geçilmesi sağlanabilir.

KAYNAKLAR

- Aksoy, B. (2019). Yüz Tanıma Sistemlerinde Doğruluk Performanslarının Değerlendirilmesi. *Mühendislik Bilimleri ve Tasarım Dergisi*, 7(4), 835-842.
- Alshazly, H., Linse, C., Barth, E., & Martinetz, T. (2020). Deep convolutional neural networks for unconstrained ear recognition. *IEEE Access*, 8, 170295-170310.
- Anwar, A. S., Ghany, K. K. A., & Elmahdy, H. (2015). Human ear recognition using geometrical features extraction. *Procedia Computer Science*, 65, 529-537.
- Arı, A., & Berberler, M. E. (2017). Interface design for the solution of prediction and classification problems with ANN. *Acta Infologica*, 1(2), 55-73.
- Arpacı, S. A., & Kalıpsız, O. (2018). Yazılım Hata Sınıflandırmasında Farklı Naive Bayes Tekniklerin Kıyaslanması. *Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi*, 7(1), 1-13.
- Arunadevi, D. J., Ramya, S., & Raja, M. R. (2018). A study of classification algorithms using Rapidminer. *International Journal of Pure and Applied Mathematics*, 119(12), 15977-15988.
- Ayhan, S., & Erdoğan, Ş. (2014). Destek vektör makineleriyle sınıflandırma problemlerinin çözümü için çekirdek fonksiyonu seçimi. *Eskişehir Osmangazi Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 9(1), 175-201.
- Bektaş, B., Babur, S., Turhal, U., & Köse, E. (2016). Makine öğrenmesi yardımıyla optik karakter tanıma sistemi. 5. *ULUSLARARASI MATBAA TEKNOLOJİLERİ SEMPOZYUMU*. s. 487-494
- Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test*, 25(2), 197-227.
- Boodoo-Jahangeer, N. B., & Baichoo, S. (2013). LBP-based ear recognition. In *13th IEEE international conference on bioinformatics and bioengineering* (s. 1-4). IEEE.
- Boyraz, Ö. F., Pala, M. A., Çimen, M. E., Boz, A. F., & Yıldız, M. Z. (2019). Mikrobilgisayar Tabanlı El-Bilek Damar Örüntüleri Kullanılarak Biyometrik Kimlik Doğrulama İşleminin Yapılması. *Academic Perspective Procedia*, 2(3), 593-600.
- Ceyhan, E. B. (2012). 10 parmakten herhangi biri temel alınarak yeni bir yaklaşımla parmak izinden cinsiyet tespit eden zeki sistem. Yüksek Lisans Tezi, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, Ankara, s. 105.
- Ceyhan, E. B., Sağiroğlu, Ş., Tatoğlu, S., & Atagün, E. (2014). Age estimation from fingerprints: Examination of the population in Turkey. In *2014 13th International Conference on Machine Learning and Applications* (s. 478-481). IEEE.
- Chang, C. C., & Lin, C. J. (2011). LIBSVM: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, 2(3), 1-27.

- Chen, T., Ju, S., Ren, F., Fan, M., & Gu, Y. (2020). EEG emotion recognition model based on the LIBSVM classifier. *Measurement*, 164, 108047.
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273-297.
- Çakır, M. Y. (2017). *Gerçek zamanlı yüksek kalitede ses tanıma*. Yayınlanmamış Yüksek Lisans Tezi. İstanbul Sabahattin Zaim Üniversitesi, İstanbul.
- Çetli, E., Tatar, D., & Özkoçak, V. (2019). Adli Bilimlerde DNA Parmak İzine Adli Genetik ve Adli Antropolojik Bakış. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, 8(4), 1545-1556.
- Deng, L., Hinton, G., & Kingsbury, B. (2013). New types of deep neural network learning for speech recognition and related applications: An overview. In *2013 IEEE international conference on acoustics, speech and signal processing* (s. 8599-8603). IEEE.
- Deshpande, U., & Malemath, V. (2022). A Study on Automatic Latent Fingerprint Identification System. *Journal of Computer Science Research*, 4(1).
- Dodge, S., Mounsef, J., & Karam, L. (2018). Unconstrained ear recognition using deep neural networks. *IET Biometrics*, 7(3), 207-214.
- Dreiseitl, S., & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: a methodology review. *Journal of biomedical informatics*, 35(5-6), 352-359.
- Fırtına, N., & Silahtaroglu, G. (2014). Kulaktan Kişi ve Kimlik Tespiti için Örnek Bir Uygulama. *Beykent Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, 7(2), 21-46.
- Foozy, C. F. M., Ahmad, R., Abdollah, M. F., & Wen, C. C. (2017). A comparative study with RapidMiner and WEKA tools over some classification techniques for SMS spam. In *IOP Conference Series: Materials Science and Engineering* (226,1), (s. 12-100). IOP Publishing.
- Gajjar, R., & Zaveri, T. (2017). Defocus blur radius classification using random forest classifier. In *2017 International Conference on Innovations in Electronics, Signal Processing and Communication (IESC)* (s. 219-223). IEEE.
- Galdámez, P. L., Raveane, W., & Arrieta, A. G. (2017). A brief review of the ear recognition process using deep neural networks. *Journal of Applied Logic*, 24, 62-70.
- Ganapathi, I. I., Ali, S. S., & Prakash, S. (2020). Geometric statistics-based descriptor for 3D ear recognition. *The Visual Computer*, 36(1), 161-173.
- Gillani, S. I., Azam, M. A., & Ehatisham-Ul-Haq, M. (2020). Age estimation and gender classification based on human gait analysis. In *2020 International conference on emerging trends in smart technologies (ICETST)* (s. 1-6). IEEE.

- Gnanasivam, P., & Muttan, S. (2013). Gender classification using ear biometrics. In *Proceedings of the Fourth International Conference on Signal and Image Processing 2012 (ICSIP 2012)* (s. 137-148). Springer, India.
- Grandini, M., Bagli, E., & Visani, G. (2020). Metrics for multi-class classification: an overview. *arXiv preprint arXiv:2008.05756*.
- Guenouni, S., Mansouri, A., & Ahaitouf, A. (2019). Biometric systems and their applications. In *Visual Impairment and Blindness-What We Know and What We Have to Know. IntechOpen*. s. 348.
- Haltaş, A., & Alkan, A. (2013). İmmunohistokimyasal Boyalar İle Tiroid Tümörü Teşhisinde Naive Bayes Algoritması Kullanılması. *Ankara Bilişim Konferansı*.
- Hansley, E. E., Segundo, M. P., & Sarkar, S. (2018). Employing fusion of learned and handcrafted features for unconstrained ear recognition. *IET Biometrics*, 7(3), 215-223.
- Hassaballah, M., Alshazly, H. A., & Ali, A. A. (2020). Robust local oriented patterns for ear recognition. *Multimedia Tools and Applications*, 79(41), 31183-31204.
- Hassan, B., Izquierdo, E., & Piatrik, T. (2021). Soft biometrics: a survey. *Multimedia Tools and Applications*, 1-44.
- İnel Özkiper, Z. (2019). *Biyometrik kimlik tanımlama sistem tasarımı* (Master's thesis, Fen Bilimleri Enstitüsü).
- Jacob, L., & Raju, G. (2014). Ear recognition using texture features-a novel approach. In *Advances in Signal Processing and Intelligent Recognition Systems* (s. 1-12). Springer, Cham.
- Karasulu, B., Yücalar, F., & Borandağ, E. (2022) İnsan kulağı görüntüleri kullanarak cinsiyet tanıma için derin öğrenme tabanlı melez bir yaklaşım. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 37(3), 1579-1594.
- Kavipriya, A., & Muthukumar, A. (2019). Human Age Estimation based on ear Biometrics using KNN. In *2019 IEEE International Conference on Clean Energy and Energy Efficient Electronics Circuit for Sustainable Development (INCCES)* (s. 1-5). IEEE.
- Keskenler, M. F., & Keskenler, E. F. (2017). Geçmişten günümüze yapay sinir ağları ve tarihçesi. *Takvim-i Vekayi*, 5(2), 8-18.
- Khalidi, Y., Benzaoui, A., Ouahabi, A., Jacques, S., & Taleb-Ahmed, A. (2021). Ear recognition based on deep unsupervised active learning. *IEEE Sensors Journal*, 21(18), 20704-20713.
- Khorsandi, R., & Abdel-Mottaleb, M. (2013). Gender classification using 2-D ear images and sparse representation. In *2013 IEEE Workshop on applications of computer vision (WACV)* (s. 461-466). IEEE.

- Kılıç, S. (2013). Doğrusal regresyon analizi. *Journal of Mood Disorders*, 3(2), 90-92.
- Kortli, Y., Jridi, M., Al Falou, A., & Atri, M. (2020). Face recognition systems: A survey. *Sensors*, 20(2), 342.
- Küçüksille, E. U., & Ateş, N. (2013). Destek vektör makineleri ile yaramaz elektronik postaların filtrelenmesi. *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 6(1).
- Lei, J., Zhou, J., & Abdel-Mottaleb, M. (2013). Gender classification using automatically detected and aligned 3D ear range data. In 2013 International Conference on Biometrics (ICB) (s. 1-7). IEEE.
- Ma, X., Spinner, S., Venditti, A., Li, Z. & Tang, S. (2019). Initial Margin Simulation with Deep Learning. SSRN Electronic Journal. 10.2139/ssrn.3357626.
- Manav, C. (2010). Görüntü İşleme Yardımı İle Kimlik Tespiti. Yüksek Lisans Tezi, Gazi Üniversitesi Fen Bilimleri Enstitüsü, , Ankara, s. 105.
- Manzoor, S. I., & Selwal, A. (2018). An analysis of biometric based security systems. In 2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC) (s. 306-311). IEEE.
- Meng, D., Nixon, M. S., & Mahmoodi, S. (2021). On distinctiveness and symmetry in ear biometrics. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 3(2), 155-165.
- Mutlag, W. K., Ali, S. K., Aydam, Z. M., & Taher, B. H. (2020). Feature extraction methods: a review. In Journal of Physics: Conference Series (Vol. 1591, No. 1, p. 012028). IOP Publishing.
- Namasudra, S., Devi, D., Kadry, S., Sundarasekar, R., & Shanthini, A. (2020). Towards DNA based data security in the cloud computing environment. *Computer Communications*, 151, 539-547.
- Nieles, M., Dempsey, K., & Pillitteri, V. Y. (2017). An introduction to information security. *NIST special publication*, 800(12), 101.
- Nguyen-Quoc, H., & Hoang, V. T. (2020). Gender recognition based on ear images: a comparative experimental study. In 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI) (s. 451-456). IEEE.
- Oravec, M., Pavlovičová, J., Sopiak, D., Jirka, V., Loderer, M., Lehota, Ľ., ... & Gerát, J. (2016). Mobile ear recognition application. In 2016 International Conference on Systems, Signals and Image Processing (IWSSIP) (s. 1-4). IEEE.
- Özlem, A., & Güngör, O. (2012). Rastgele orman algoritması kullanılarak çok bantlı görüntülerin sınıflandırılması. *Jeodezi ve Jeoinformasyon Dergisi*, (106), (s. 139-146).

- Öztürk, K., & Şahin, M. E. (2018). Yapay sinir ağları ve yapay zekâ'ya genel bir bakış. *Takvim-i Vekayi*, 6(2), 25-36.
- Panesar, S. S., D'Souza, R. N., Yeh, F. C., & Fernandez-Miranda, J. C. (2019). Machine learning versus logistic regression methods for 2-year mortality prognostication in a small, heterogeneous glioma database. *World neurosurgery*: X, 2, 100012.
- Potok, T. (2021). Adiabatic quantum linear regression. *Scientific reports*, 11(1), 1-10.
- Pramanik, S., Chatterjee, S., Kumar, G. S., & Devi, P. S. (2018). Egg-shell derived carbon dots for base pair selective DNA binding and recognition. *Physical Chemistry Chemical Physics*, 20(31), 20476-20488.
- Rahman, M., Sadi, M. S., & Islam, M. R. (2014). Human ear recognition using geometric features. In *2013 international conference on Electrical information and communication technology (EICT)* (s. 1-4).
- Riaz, N., Riaz, A., & Khan, S. A. (2018). Biometric template security: an overview. *Sensor Review*.
- Sarhan, S., Alhassan, S., & Elmougy, S. (2017). Multimodal biometric systems: a comparative study. *Arabian Journal for Science and Engineering*, 42(2), 443-457.
- Singla, S. K., Singh, M., & Kanwal, N. (2021). Biometric System-Challenges and Future Trends. In *2021 8th International Conference on Computing for Sustainable Global Development (INDIACom)* (s. 647-651). IEEE.
- Smith, M., Mann, M., & Urbas, G. (2018). *Biometrics, crime and security*. Routledge.
- Solmaz, R., Günay, M., & Alkan, A. (2014). Fonksiyonel Tiroit Hastalığı Tanısında Naive Bayes Sınıflandırıcısının Kullanılması. *Akademik Bilişim'14-XVI. Akademik Bilişim Konferansı Bildirileri 5-7 Şubat 2014 Mersin Üniversitesi*, 891-897.
- Somepalli, S. H., Tangella, S. K. R., & Yalamanchili, S. (2020). Information Security Management. HOLISTICA. *Journal of Business and Public Administration*, 11(2), 1-16.
- Şenel, S., & Alatli, B. (2014). Lojistik regresyon analizinin kullanıldığı makaleler üzerine bir inceleme. *Journal of Measurement and Evaluation in Education and Psychology*, 5(1), 35-52.
- Travers, A., & Muskhelishvili, G. (2015). DNA structure and function. *The FEBS journal*, 282(12), 2279-2295.
- Trikoš, M., Tot, I., Bajčetić, J., Lalović, K., Jovanović, B., & Bogićević, D. (2019"). Biometric security standardization. In *2019 Zooming Innovation in Consumer Technologies Conference (ZINC)* (s. 17-20). IEEE.
- Wang, J., & Wang, G. (2017). Quality-specific hand vein recognition system. *IEEE Transactions on Information Forensics and Security*, 12(11), 2599-2610.

- Wang, Z., Yang, J., & Zhu, Y. (2021). Review of ear biometrics. *Archives of Computational Methods in Engineering*, 28(1), 149-180.
- Xia, S., Liu, Y., Yuan, G., Zhu, M., & Wang, Z. (2017). Indoor fingerprint positioning based on Wi-Fi: An overview. *ISPRS International Journal of Geo-Information*, 6(5), 135.
- Yaman, D., Eyiokur, F. I., Sezgin, N., & Ekenel, H. K. (2018). Age and gender classification from ear images. In *2018 International Workshop on Biometrics and Forensics (IWBF)* (s. 1-7). IEEE.
- Yang, W. Y., Zhang, H., Song, D. Z., & Lai, F. N. (2018). Review of face recognition methods. *DEStech Transactions on Computer Science and Engineering*, (msota).
- Ying, T., Shining, W., & Wanxiang, L. (2018). Human ear recognition based on deep convolutional neural network. In *2018 Chinese Control And Decision Conference (CCDC)* (s. 1830-1835). IEEE.
- Zhang, G., & Wang, Y. (2011). Hierarchical and discriminative bag of features for face profile and ear based gender classification. In *2011 International joint conference on biometrics (IJCB)* (s. 1-8). IEEE.

EKLER

EK-1: Etik Kurul Onay Belgesi.



T.C.
BARTIN ÜNİVERSİTESİ REKTÖRLÜĞÜ
Sosyal ve Beşeri Bilimler Etik Kurulu



Sayı : E-23688910-050.01.04-2100041742
Konu : Sosyal ve Beşeri Bilimler Etik Kurulu
Onay Formu

27.05.2021

Protokol No:	2021-SBB-0234
Araştırmanın Başlığı:	Kulak Görüntülerinden Soft Biyometrik Özellikleri Derin Öğrenme İle Tahmin Eden Zeki Sistem
Proje Yürütücüsü:	Mertkan Sinoplu
Başvuru Formunun Geliş Tarih:	29.04.2021
Karar Tarihi:	26.05.2021
Toplantı No:	8

Başvuru dosyasında etik sorun oluşturabilecek sorular/maddeler, süreçler ya da unsurlar bulunmadığından 26.05.2021 tarihli ve 8 numaralı toplantıda 2021-SBB-0234 numaralı başvuruya araştırma için ETİK KURUL ONAY belgesinin verilmesine oy çokluğu ile karar verilmiştir.

Prof. Dr. Ash YAZICI
Kurul Başkanı

Doç. Dr. Ayşe Derya IŞIK
Başkan yardımcısı

Dr. Öğr. Üyesi Hasan Basri
KANSIZOĞLU
Üye

Dr. Öğr. Üyesi Emine GENÇ
Üye

Dr. Öğr. Üyesi Emel GENÇ
Üye

Dr. Öğr. Üyesi Veysel GENÇİL
Üye

Dr. Öğr. Üyesi İlknur DOLU
Üye

Belge Doğrulama Kodu: F9CTC4F

Bu belge, güvenli elektronik imza ile imzalanmıştır.

Belge Takip Adresi: <http://ubya.bartin.edu.tr/ERMS/Record/ConfirmationPage/Index>

Adres: Ağdacı Mahallesi Fakülte Caddesi No:54 Bartın

Telefon No: (0 378) 2235500

e-Posta:

Kap Adresi: barinuniv@barin.edu.tr

Faks No: (0 378) 2235042

İnternet Adresi: <http://www.bartin.edu.tr/>

Bilgi için :

Ash Yazıcı
Kurul Başkanı

Telefon No:



EK-2: Çalışma Kapsamında Kullanılan Anket Formu.

Katılımcı No:

Kulak Görüntülerinden Soft Biyometrik Özellikleri Derin Öğrenme İle Tahmin Eden Zeki Sistem Anket Formu

Lütfen aşağıdaki soruları eksiksiz bir şekilde doldurunuz. Her soru için tek bir kutucuğu işaretleyiniz. Bu ankette verdiğiniz cevaplar, araştırma kapsamında çekilecek olan kulak fotoğraflarınız ile ilişkilendirilmek için kullanılacaktır. Verdiğiniz bilgiler güvenli bir veri tabanında saklanacaktır ve akademik çalışmalar dışında herhangi bir çalışmada kullanılmayacaktır.

1- Cinsiyetiniz:

- ☐ Kadın
☐ Erkek

2- Yaş aralığınız:

- ☐ 18 – 28
☐ 29 – 38
☐ 39 – 48
☐ 49 – 58
☐ 59 – 64
☐ 65 ve üzeri

3- Kilo Aralığınız:

- ☐ 50 – 59 kg
☐ 60 – 69 kg
☐ 70 – 79 kg
☐ 80 – 89 kg
☐ 90 – 99 kg
☐ 100 - 109 kg
☐ 110 – 119 kg
☐ 120 – 129 kg
☐ 130 – 139 kg
☐ 140 kg ve üzeri

4- Boy Aralığınız

- ☐ 150 – 159 cm
☐ 160 – 169 cm
☐ 170 – 179 cm
☐ 180 – 189 cm
☐ 190 – 199 cm
☐ 200 – 209 cm
☐ 210 cm ve üzeri

5- Göz Renginiz

- ☐ Kahverengi
☐ Mavi
☐ Ela
☐ Kehribar
☐ Gri
☐ Yeşil

6- Doğal Saç Renginiz

- ☐ Kahverengi
☐ Sarı
☐ Siyah
☐ Kumral
☐ Kızıl
☐ Diğer

7- Ten Renginiz

- ☐ Çok açık / Beyaz Cilt
☐ Açık Cilt
☐ Orta Açık / Buğday Cilt
☐ Orta koyu / Esmer Cilt
☐ Koyu / Kahverengi Cilt
☐ Çok koyu / Siyah Cilt

8- Saç Tipiniz

- ☐ Düz Saç
☐ Dalgalı Saç
☐ Kıvrıkcık Saç
☐ Sarmal Saç

ÖZGEÇMİŞ

İş Deneyimi

Stajlar : MSU Yazılım (2017 Temmuz-Ağustos)
MSU Yazılım (2018 Temmuz-Ağustos)
Çalıştığı Kurumlar : Teklas

İletişim

E-Posta Adresi : mertkansinoplu@gmail.com

Tarih : 03/05/2023 (Tez Savunma Tarihi)



