



T.C.
KONYA TEKNİK ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ



Yeni Bir Açıklanabilir Yapay Zeka Modeli Geliştirilmesi

Ahmet Haydar ÖRNEK

DOKTORA TEZİ

Elektrik-Elektronik Mühendisliği Anabilim Dalı

Aralık-2023
KONYA
Her Hakkı Saklıdır

TEZ KABUL VE ONAYI

Ahmet Haydar ÖRNEK tarafından hazırlanan “Yeni Bir Açıklanabilir Yapay Zeka Modeli Geliştirilmesi” adlı tez çalışması 29/12/2023 tarihinde aşağıdaki jüri tarafından oy birliği / oy çokluğu ile Konya Teknik Üniversitesi Lisansüstü Eğitim Enstitüsü Elektrik-Elektronik Mühendisliği Anabilim Dalı’nda DOKTORA TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

İmza

Başkan

Prof. Dr. Seral ÖZŞEN

.....

Danışman

Prof. Dr. Murat CEYLAN

.....

Üye

Dr. Öğr. Üyesi Hüseyin Oktay ALTUN

.....

Üye

Dr. Öğr. Üyesi Hasan KOYUNCU

.....

Üye

Dr. Öğr. Üyesi Güzin ÖZMEN

.....

Yukarıdaki sonucu onaylarım.

Prof. Dr. Mevlüt UYAN
Enstitü Müdürü

TEZ BİLDİRİMİ

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

DECLARATION PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Ahmet Haydar ÖRNEK

Tarih:

ÖZET

DOKTORA TEZİ

Yeni Bir Açıklanabilir Yapay Zeka Modeli Geliştirilmesi

Ahmet Haydar ÖRNEK

**Konya Teknik Üniversitesi
Lisansüstü Eğitim Enstitüsü
Elektrik-Elektronik Mühendisliği Anabilim Dalı**

Danışman: Prof. Dr. Murat CEYLAN

2023, 63 Sayfa

Jüri

Prof. Dr. Seral ÖZŞEN

Prof. Dr. Murat CEYLAN

Dr. Öğr. Üyesi Hüseyin Oktay ALTUN

Dr. Öğr. Üyesi Hasan KOYUNCU

Dr. Öğr. Üyesi Güzin ÖZMEN

Bu tez çalışması, Açıklanabilir Yapay Zeka (AYZ) alanında görsel yaklaşımlardan biri olan Sınıf Aktivasyon Haritalarına (SAH) dayalı dört yeni yöntemin geliştirilmesi ve uygulanması konusunu ele almaktadır. Bu yöntemler, Evrişimli Sinir Ağları (ESA) temel alınarak tasarlanmış ve sırasıyla CodCAM, HayCAM, HayCAM+, ve HayCAMJ olarak adlandırılmıştır.

CodCAM, dört farklı görsel AYZ yöntemi (GradCAM, GradCAM++, LayerCAM, EigenCAM) kullanılarak SAH'ların oluşturulmasını sağlayan bir yöntemdir. CodCAM kullanılarak gerçekleştirilen birinci çalışmada, termal görüntülerle yeni doğan bebeklerin sınıflandırılma işlemi incelenmiştir. Bu çalışmanın sonuçları, ESA modelinin bebeklerin anatomik bölgelerine (örneğin kol, koltuk altı, baş, boyun, ayak ve vücut) odaklandığını göstermiştir. Bu sonuçlar, ESA modelinin yalnızca genel görüntüyü değil, aynı zamanda bebeklerin vücut yapılarını tanıdığını ve bu bilgiyi doğru sınıflandırmalar yapmak için kullandığını göstermektedir.

HayCAM, SAH'ları oluştururken Temel Bileşenler Analizi (TBA) kullanarak ESA'nın son katmanındaki filtreleri azaltmakta ve görüntü üzerinde yoğunlaşmış SAH'lar elde edilmesini sağlamaktadır. Maske takan ve takmayan insanları ayırt eden bir sınıflandırma modeliyle HayCAM kullanıldığında oluşturulan SAH'ların, ESA modelinin görüntülerde ağız bölgesine yoğunlaştığını göstermiştir. Oluşturulan SAH'lar kullanılarak nesne tespiti performansı ölçülmüş ve IoU değerleri (Gerçek bölge ve tahmin edilen bölgenin kesişim/birleşim oranı) hesaplanmıştır. Bu değerler GradCAM için 0.1922, GradCAM++ için 0.2472, EigenCAM için 0.3386 ve önerilen HayCAM için 0.3487 olarak bulunmuştur.

HayCAM+, HayCAM'in geliştirilmiş bir versiyonudur ve boyut azaltma sürecinde kullanılan temel filtreleri otomatik olarak hesaplayarak SAH'lar elde etmeyi amaçlar. HayCAM+ ile IoU değeri yaklaşık %2.5 artarak 0,3487'den 0,3740'a yükselmiştir. HayCAMJ ise sadece tek bir filtre kullanarak SAH'lar oluşturmayı amaçlar. Bu yöntem, basit ancak etkili bir yaklaşım sunarak tek bir filtre ile elde edilen SAH'ların nesne sınıflandırma sonuçlarını açıklar.

Geliştirilen CodCAM, HayCAM, HayCAM+, ve HayCAMJ yöntemleri, Resnet18, Resnet34, Resnet50 ve Mobilenet ESA modelleri kullanılarak Pascal VOC veri setleri üzerinde karşılaştırılmıştır. Pascal VOC 2007 için en yüksek IoU sonuçları sırasıyla HayCAM için 0.2422 (Resnet34), HayCAM+ için 0.3303 (Resnet18), HayCAMJ için 0.3277 (Resnet18) ve CodCAM için 0.4079 (Resnet34) olarak elde edilmiştir. Pascal VOC 2012 için de benzer şekilde HayCAM için 0.2484 (Resnet34), HayCAM+ için 0.3101 (Resnet18), HayCAMJ için 0.3118 (Resnet34) ve CodCAM için 0.3674 (Resnet34) olarak

bulunmuştur. Bu geliştirilen yöntemler, ESA tarafından verilen kararların görselleştirilerek anlaşılmasına katkı sağlamakta ve görselleştirilen çıktılar kullanılarak nesne tespiti uygulamasının performansının artırılmasını sağlamaktadır.

Anahtar Kelimeler: Açıklanabilir yapay zeka, codcam, görsel açıklama, haycam, haycam+, haycamj, sınıf aktivasyon haritaları



ABSTRACT

PhD THESIS

Developing A New Explainable Artificial Intelligence Model

Ahmet Haydar ÖRNEK

**Konya Technical University
Institute of Graduate Studies
Department of Electrical and Electronics Engineering**

Advisor: Prof. Dr. Murat CEYLAN

2023, 63 Pages

Jury

Prof. Dr. Seral ÖZŞEN

Prof. Dr. Murat CEYLAN

Assist. Prof. Dr. Hüseyin Oktay ALTUN

Assist. Prof. Dr. Hasan KOYUNCU

Assist. Prof. Dr. Güzin ÖZMEN

This thesis work addresses the development and implementation of four new methods based on Class Activation Maps (CAMs), a visual approach in the field of Explainable Artificial Intelligence (XAI). These methods, designed based on Convolutional Neural Networks (CNNs), are named CodCAM, HayCAM, HayCAM+, and HayCAMJ, respectively.

CodCAM is a method that generates CAMs using four different visual XAI techniques (GradCAM, GradCAM++, LayerCAM, EigenCAM). The first study using CodCAM examines the classification of thermal images belonging to newborn babies. The results of this study show that the CNN model focuses on anatomical regions of babies (such as arm, armpit, head, neck, foot, and body). These results demonstrate that the CNN model not only captures the overall image but also recognizes the anatomical structures of babies and uses this information for accurate classification.

HayCAM reduces the filters in the final layer of the CNN using Principal Component Analysis (PCA) during CAMs generation, resulting in focused CAMs on the image. When HayCAM is used with a classification model to distinguish people wearing masks from those who aren't, the generated CAMs show a concentration around the mouth region. These generated CAMs are used to measure object detection performance, and Intersection over Union (IoU) values are calculated. These values are found to be 0.1922 for GradCAM, 0.2472 for GradCAM++, 0.3386 for EigenCAM, and 0.3487 for the proposed HayCAM.

HayCAM+ is an enhanced version of HayCAM that aims to automatically calculate the essential filters used in the dimension reduction process to obtain CAMs. With HayCAM+, the IoU value increases by approximately 2.5% from 0.3487 to 0.3740. HayCAMJ, on the other hand, aims to create CAMs using only a single filter. This method offers a simple yet effective approach to explaining object classification results obtained with a single filter.

The developed CodCAM, HayCAM, HayCAM+, and HayCAMJ methods are compared using Resnet18, Resnet34, Resnet50, and Mobilenet CNN models on the Pascal VOC datasets. The highest IoU results for Pascal VOC 2007 are 0.2422 for HayCAM (Resnet34), 0.3303 for HayCAM+ (Resnet18), 0.3277 for HayCAMJ (Resnet18), and 0.4079 for CodCAM (Resnet34). Similarly, for Pascal VOC 2012, the results are 0.2484 for HayCAM (Resnet34), 0.3101 for HayCAM+ (Resnet18), 0.3118 for HayCAMJ (Resnet34), and 0.3674 for CodCAM (Resnet34). These developed methods contribute to visualizing the decisions made by the CNN models, enhancing object detection performance using the visualized outputs.

Keywords: Class activation maps, CodCAM, explainable artificial intelligence, HayCAM, HayCAM+, HayCAMJ, visual explanation.



ÖNSÖZ

Doktoraya başlama konusunda beni cesaretlendiren ve geniş vizyonuyla her zaman yol gösterici olan değerli danışmanım Konya Teknik Üniversitesi Mühendislik ve Doğa Bilimleri Fakültesi Elektrik-Elektronik Mühendisliği Öğretim Üyesi Prof. Dr. Murat CEYLAN'a içtenlikle teşekkür ederim.

Sadece akademik alanda değil, aynı zamanda sektöre yönelik deneyimlerimi ve birikimimi artırmama yardımcı olan Medianova, Aukilabs ve Huawei şirketlerine de teşekkürler.

Babamın, annemin ve kardeşlerimin desteği olmadan bu süreci tamamlayamazdım. Her kararında verdikleri destek için teşekkürler.

Ahmet Haydar ÖRNEK
KONYA-2023

İÇİNDEKİLER

ÖZET	iv
ABSTRACT.....	vi
ÖNSÖZ	viii
İÇİNDEKİLER	ix
SİMGELER VE KISALTMALAR	x
1. GİRİŞ	1
2. KAYNAK ARAŞTIRMASI	3
3. MATERYAL ve YÖNTEM	6
3.1. Yeni Doğan Bebeklere Ait Termal Görüntüler.....	6
3.2. Maske Görüntüleri	8
3.3. Pascal VOC 2007 ve 2012 Veri Setleri	9
3.4. Sınıflama	10
3.5. Yapay Sinir Ağı	11
3.6. Evrişimli Sinir Ağları.....	13
3.7. Transfer Öğrenme	17
3.8. Temel Bileşenler Analizi	19
3.9. Sınıf Aktivasyon Haritaları	19
3.10. CodCAM.....	23
3.11. HayCAM.....	24
3.12. HayCAM+	25
3.13. HayCAMJ	26
3.15. Zayıf Danışmanlı Nesne Tespiti	27
3.15. Değerlendirme Metrikleri	28
4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA.....	30
4.1. CodCAM Yönteminin Medikal Görüntüler Üzerindeki Sonuçları.....	30
4.2. HayCAM Yönteminin Maske Tanıma Problemi Üzerindeki Sonuçları	38
4.3. HayCAM+ Yönteminin Maske Tanıma Problemi Üzerindeki Sonuçları.....	42
4.4. HayCAMJ Yönteminin Pascal VOC 2007 ve 2012 Üzerine Sonuçları.....	46
4.5. Geliştirilen Yöntemlerin Pascal VOC 2007 ve 2012 Üzerinde Karşılaştırılması	50
5. SONUÇLAR VE ÖNERİLER	54
5.1 Sonuçlar	54
5.2 Öneriler	56
KAYNAKLAR	57
EKLER	63

SİMGELER VE KISALTMALAR

Simgeler

A	: Son evrişim katmanı
f [.]	: Filtre fonksiyonu
g [.]	: Giriş görüntüsü fonksiyonu
k	: Filtre boyutu
L	: Aktivasyon haritası
n_g	: Görüntünün giriş boyutu
p	: Görüntü etrafına eklenen sıfır sayısı
r	: Öğrenme oranı
s	: Filtre adım büyüklüğü
w_n	: Yapay sinir ağı ağırlıkları
w_s	: Aktivasyon ağırlıkları
x_n	: Yapay sinir ağı girişi
y_n	: Yapay sinir ağı çıkışı

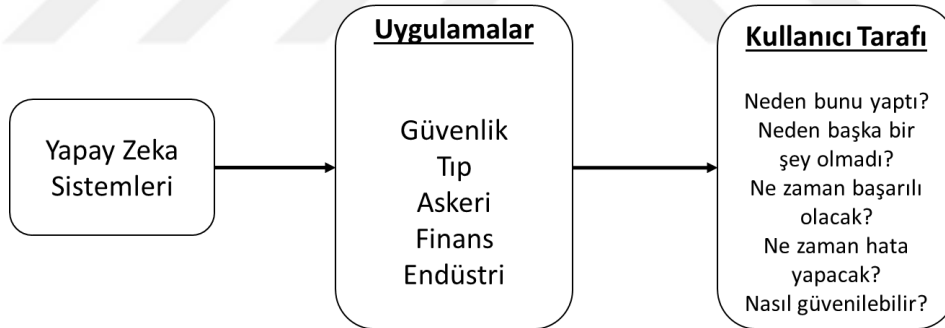
Kısaltmalar

AYZ	: Açıklanabilir yapay zeka
CodCAM	: Common decision class activation mapping
ESA	: Evrişimli sinir ağları
GIÜ	: Grafik işlem ünitesi
IoU	: Intersection over union
GradCAM	: Gradient weighted class activation mapping
SAH	: Sınıf aktivasyon haritaları
TBA	: Temel bileşenler analizi
TDA	: Tekil değer ayrışımı
YSA	: Yapay sinir ağları

1. GİRİŞ

Yapay zeka alanında yaşanan gelişmeler, çağımızın en umut verici teknolojik ilerlemelerinden birini işaret etmektedir. Yapay zeka, bilgisayarların insanlar gibi öğrenme ve karar verme yetenekleri kazanmasını sağlayan bir disiplindir (Winston, 1984). Yapay zeka askeri uygulamalar, tıbbi teşhis, finansal yaklaşımlar, oyunlar ve daha birçok alanda potansiyel uygulamalara sahiptir (Riedl ve Zook, 2013; Forrest ve Hoanca, 2015; Hamet ve Tremblay, 2017; Holzinger ve diğ., 2019; Hilpisch, 2020; Morgan ve diğ., 2020; Cao, 2022; Maathuis, 2022).

Yapay zekanın karar mekanizmasını anlamak ve açıklamak, Şekil 1.1’de görüldüğü gibi, bu teknolojinin kullanımında önemli bir problem olarak ortaya çıkmaktadır. Yapay zeka derin öğrenme algoritmaları kullanırken kararlarını açıklayabilme yeteneği oldukça sınırlıdır. Alınan kararların neden olduğunun gösterilmemesi sisteme duyulan güveni azaltmaktadır (Gunning ve Aha, 2019). Bu tez kapsamında yapay zeka modellerinin verdiği kararların anlaşılması için yeni yöntemler geliştirilmiştir.



Şekil 1.1 Yapay Zeka Sistemleri, Uygulama Alanları ve Kullanıcı Beklentileri

Açıklanabilir Yapay Zeka (AYZ), bir yapay zeka modelinin karar sürecinin insanlar tarafından anlaşılabilir bir şekilde ifade edilmesini içermektedir. Diğer bir deyişle yapay zeka kullanılarak alınan kararların/sonuçların açıklanması ve izah edilmesi anlamına gelir (Arrieta ve diğ., 2020).

Açıklanabilirlik, endüstriyel uygulamalardan akademik araştırmalara kadar geniş bir yelpazede büyük bir ilgi görmektedir. Araştırmacılar, yapay zeka sistemlerinin kararlarını anlamak için sayısal, kural tabanlı, görsel veya hibrit yöntemler gibi çeşitli yaklaşımlar geliştirmekte ve bu alandaki çalışmalarını sürdürmektedirler. (Das ve Rad, 2020).

Açıklanabilir modeller, yapay zeka sistemlerine duyulan güveni artırarak bu sistemlerin yaygın kabulünü desteklemektedir. Bir sistemin nasıl karar verdiğinin anlaşılması, insanların daha bilinçli hareket etmesini sağlamaktadır. Diğer bir taraftan bu sistemler etik ve hukuki sorumlulukların belirlenmesinde önemli bir adımdır (Adadi ve Berrada, 2018). Ancak, AYZ bazı sınırlamaları ve zorlukları da barındırmaktadır. Örneğin, açıklanabilirlik için kullanılan modellerin karmaşıklığı, performans ve anlaşılabilirlik arasında bir denge gerektirir. AYZ modelleri, eğitildikleri veriye bağımlıdır ve bu nedenle farklı veri kümeleri için farklı modeller eğitilmelidir.

Derin öğrenme bir yapay zeka yaklaşımı olarak birçok uygulama alanına sahiptir. Derin öğrenme modelleri içerdikleri fazla katmanlar sayesinde büyük ve çeşitliliği fazla veri setlerinin öğrenilmesini sağlamaktadır (LeCun ve diğ., 2015). Ancak bir model ne kadar fazla katman içeriyorsa açıklanabilirliği de o derece azalmaktadır (Singh ve diğ., 2020).

Yapay zeka modelleri yüksek doğruluklarda sonuç üretmesine rağmen bir kararı nasıl verdikleri tam olarak açıklanamamaktadır. Çünkü bir derin öğrenme modeli milyonlarca parametre ve doğrusal olmayan katmanlar içermektedir. Bu nedenle derin öğrenme modelleri “Kara Kutu” (Castelvecchi, 2016) modeller olarak bilinir.

Bir modelin nasıl karar verdiğinin açıklanması üzerine yapılan araştırmalar AYZ başlığı altında toplanmaktadır. Sayısal, kural tabanlı ve görsel AYZ yöntemleri vardır. Bu tez çalışması kapsamında derin öğrenme modellerinin karar mekanizmaları görsel yaklaşımlarla incelenerek yeni teknikler geliştirilmiştir. Geliştirilen yöntemler sayesinde bir model karar verdiği zaman modelin ilgili görüntünün neresine odaklandığı gösterilerek araştırmacıların verilen kararları değerlendirmesine olanak sağlanmıştır (Örnek ve Ceylan, 2022a). Ayrıca modellerin odaklandığı bölgeler kullanılarak nesne tanıma uygulamaları da gerçekleştirilmiş ve karşılaştırmalı olarak sonuçları gösterilmiştir. Elde edilen sonuçlar tez çalışması kapsamında geliştirilen yöntemlerin daha iyi çalıştığını ortaya koymaktadır.

2. KAYNAK ARAŞTIRMASI

AYZ çalışmalarına bakıldığında bu çalışmaların dört ana başlık altında toplandığı görülmektedir. Bunlar nümerik açıklamalar, kural tabanlı açıklamalar, görsel açıklamalar ve hibrit açıklamalardır (Minh ve diğ., 2022). Bir modelin bir sonucu oluştururken hangi özellikleri veya faktörleri dikkate aldığı AYZ yöntemleri ile açıklanmaktadır.

Nicel Giriş Etkisi (Datta ve diğ., 2016) ve Monoton Etki Ölçümleri (Sliwinski ve diğ., 2017) gibi yöntemler etki fonksiyonlarından (Koh ve Liang, 2017) yararlanarak her bir giriş özelliğinin katkısının değerlendirilmesini sağlar. Etki fonksiyonları bağımsız değişkenlerin dağılımındaki değişikliğe göre çıkış olasılığını ölçen istatistiksel bir yöntemdir. Nümerik açıklamalar giriş değişkenlerinin çıkışa olan katkısını nicel metriklerle açıklamaya odaklanır (Hase ve Bansal, 2020). Örneğin 10 farklı özelliğe sahip bir verinin sınıflandırılması işleminde, giriş özelliklerinin önemini anlamak için şu adımlar izlenmektedir: 10 farklı özellik ile eğitim yap, bir özellik eksiltip dokuz özellik ile eğitim yap, eksiltelen özellik dışındaki bir özelliği eksiltip bir eğitim yap. Böylelikle her bir özelliğin sınıflandırma sonucuna olan etkisi ortaya çıkarılmış olur.

Kural tabanlı açıklamalar girişten tahmine kadar olan süreci Eğer-İse kuralları oluşturarak açıklamaya çalışır (Van der Waa ve diğ., 2021). Karar ağaçları, en iyi bilinen kural tabanlı açıklamalar arasındadır (Navada ve diğ., 2011). Hibrit açıklamalar, birden fazla açıklama tekniğini birleştirerek daha kapsamlı açıklamalar sağlar (Joseph ve diğ., 2022). Örneğin bir model karar verdiğinde kural tabanlı açıklamalar kullanarak karara nasıl ulaştığı, görsel açıklamalar kullanarak ise verinin hangi kısımlarına odaklandığı ortaya çıkarılır.

Görsel açıklamalar bir modelin nasıl çalıştığını “Aktivasyon Haritaları” oluşturarak açıklamaya odaklanır. Örneğin Katman Bazında Alaka Yayılımı (Montavon ve diğ., 2019) sınıflayıcının çeşitli hesaplama katmanlarına ayrılarak her bir pikselin önemini ortaya çıkarır. Ters Evrişim Ağları (Zeiler ve diğ., 2010) danışmansız bir yöntem olarak evrişimsel ayrıştırmaya dayanmakta ve ayrıştırılan evrişimlerden bir hiyerarşi oluşturarak özelliklerin öğrenilmesini sağlamaktadır.

Molle ve diğ. kendi oluşturdukları dört katmanlı bir ESA modeli eğiterek %75 “eğri altında kalan alan” (Purves, 1992) değerine ulaşmış ve cilt lezyonu sınıflandırmasının aktivasyon haritalarını görselleştirmiştir (Van Molle ve diğ., 2018). Böylece ESA modelinin ve dermatologların kararlarının benzer olduğunu gösterilmiştir.

Pereira ve diğ. 2018 yılında tümör derecesini manyetik rezonans görüntülerinden korumak için ESA tabanlı bir yaklaşım önermektedir (Pereira ve diğ., 2018). Elde edilen yaklaşım %92.98 doğrulukla çalışmakta ve GradCAM tarafından oluşturulan aktivasyon haritalarının tümör bölgesi üzerinde görüldüğünü göstermektedir.

2019 yılında Sayres ve diğ., kararları aktivasyon haritaları ile görselleştirerek derin öğrenmenin diyabetik retinopati sınıflama üzerindeki etkilerine dikkat çekmişlerdir (Sayres ve diğ., 2019). Aktivasyon haritalarının, uzmanların sonuçları anlama hassasiyetini artırdığını ortaya çıkarmışlardır.

Yang ve diğ. yaşa bağlı makula dejenerasyonu kararlarını ortaya çıkarmak için ESA mimarilerinden VGG16 ve GoogleNet modellerini eğitmişlerdir (Yang ve diğ., 2019a). Aktivasyon haritaları ile görselleştirmeler yaparak, tomografi görüntülerindeki lezyonları %49.77'lik oranla lokalize etmişlerdir.

2019 yılında gerçekleştirilen bir çalışmada Eitel ve diğ., derin bir ESA modeli kullanarak Alzheimer hastalığının sınıflandırılması için aktivasyon haritası yöntemlerini karşılaştırmaktadır (Eitel ve diğ., 2019). Elde edilen sonuçlar doğruluğun gradyan geri yayılımı için %46.4, rehberli gradyan geri yayılımı için %68.6, katman bazında alaka yayılımı için %78 olduğunu göstermektedir.

Papanastasopoulos ve diğ. 2020 yılında östrojen reseptörü durumunu tahmin etmek için ESA modellerinden biri olan AlexNet mimarisini kullanmışlardır (Papanastasopoulos ve diğ., 2020). Oluşturulan aktivasyon haritalarının dinamik ve uzamsal alanlarda doğru noktaları gösterdiklerini kanıtlamışlardır.

Tez çalışmasında temel alınan Sınıf Aktivasyon Haritaları (SAH), ESA'nın son evrişim katmanındaki filtrelerden faydalanarak aktivasyon haritaları oluşturulmasını amaçlar. Zhou ve diğ. tarafından 2016 yılında sunulan SAH çalışması görüntü içerisinde istenilen sınıfla ilgili bölgelerin ortaya çıkarılmasını sağlasa da bazı dezavantajlara sahiptir (Zhou ve diğ., 2016). Bu dezavantajlar tam bağlı katmanın kaldırılması ve modelin yeniden eğitilmesi zorunluluğudur. Özellikle tam bağlı katmanın kaldırılması model performansının azalmasına neden olmaktadır. Temel SAH'ları kullanarak gerçekleştirdiğimiz yeni doğan bebeklere ait termogramların sınıflandırılması ve sınıflandırma sonuçlarının açıklanması çalışmasında sınıflama doğruluğu %88.77 olarak elde edilmiştir (Örnek ve Ceylan, 2021).

SAH'ın yukarıda ifade edilen dezavantajlarını ortadan kaldırmak için Selvaraju ve diğ. tarafından Gradyan tabanlı SAH (GradCAM) önerilmiştir (Selvaraju ve diğ., 2017). GradCAM'ler birinci dereceden gradyanlar kullanarak tam bağlı katman

kaldırılmadan veya model yeniden eğitilmeden aktivasyon haritalarının üretilmesini sağlamaktadır. Lei Chen ve diğ., GradCAM ile elde edilen eğitim verilerinin gradyanlarını toplayarak ve ağırlık transferi gerçekleştirerek bir resme ait aktivasyonları ortaya çıkarmaktadır. Resnet50 mimarisi kullanılarak gerçekleştirdikleri çalışmada doğruluk değerinin GradCAM'e göre %4.5 artarak %50.7'den %55.2'ye yükseldiğini göstermişlerdir (Chen ve diğ., 2020). Hars panuar ve diğ., Covid19 tespitinin X-RAY ve bilgisayarlı tomografi görüntülerinden yapılması için GradCAM yönteminden faydalanmışlardır. Böylece %94.03 hassasiyet ve %95.85 özgüllük değerine ulaşmışlardır (Panwar ve diğ., 2020).

GradCAM++ birinci dereceden gradyanlar yerine ikinci dereceden gradyanlar kullanarak nesnelerin konumunu daha kesin bir şekilde belirleme olanağı sağlar. (Chattopadhyay ve diğ., 2018). Bu, özellikle nesnelerin görüntü içindeki konumunun belirlenmesi bakımından önemlidir. Hesaplama maliyeti bakımından karşılaştırıldığında GradCAM++, GradCAM'e göre daha fazla işlem gerektirmektedir.

ScoreCAM ağırlıklar ve gradyanlar arasındaki bağımlılığı kaldırarak gradyanların doğrusal bir kombinasyon ile elde edilmesini sağlamaktadır (Wang ve diğ., 2020a). ScoreCAM, “enerji tabanlı gösterme oyunu” olarak adlandırılan görevde %63.7 doğruluğa ulaşırken GradCAM %48.1, GradCAM++ %49.3 doğruluğa ulaşmaktadır.

AblationCAM her özellik haritasının çıkış üzerine etkisini ölçer ancak çok fazla iterasyon gerektirdiği için zaman ve işlem maliyeti yüksektir (Ramaswamy, 2020). Gerçekleştirdikleri çalışmada GradCAM %30.96 doğruluğa ulaşırken AblationCAM %33.28 doğruluğa ulaşmıştır.

LayerCAM son katmanda bulunan filtre çözünürlüklerinin düşük olduğunu belirterek orta katmanlardaki filtrelerin de kullanılmasını sağlamaktadır (Jiang ve diğ., 2021). Elde ettikleri sonuçlara göre GradCAM %43.62, GradCAM++ %45.44, ScoreCAM %39.51 doğruluğa ulaşırken LayerCAM %46.67 doğruluğa ulaşmaktadır.

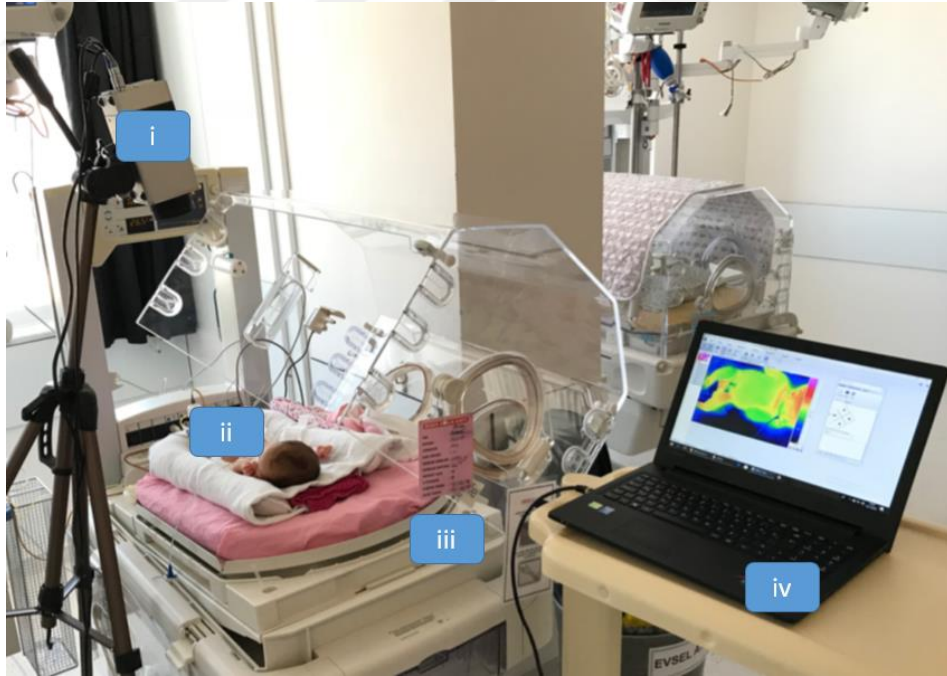
EigenCAM (Bany Muhammad ve Yeasin, 2021) Tekil Değer Ayırımı (TDA) (Henry ve Hofrichter, 1992) kullanarak son katmandaki filtrelerin azaltılmasını sağlar. Imagenet üzerinde yaptıkları çalışmada GradCAM %49.49 doğruluğa ulaşırken, EigenCAM % 52.33 doğruluğa ulaşmaktadır.

3. MATERYAL ve YÖNTEM

Derin öğrenme modellerinin eğitilmesi için çok sayıda veriye ihtiyaç duyulmaktadır. Tez çalışması kapsamında yeni doğan bebeklere ait termal görüntüler, maske takan ve takmayan insanlara ait görüntüler, Pascal VOC 2007 ve Pascal VOC 2012 veri setlerinden elde edilen görüntüler kullanılmıştır.

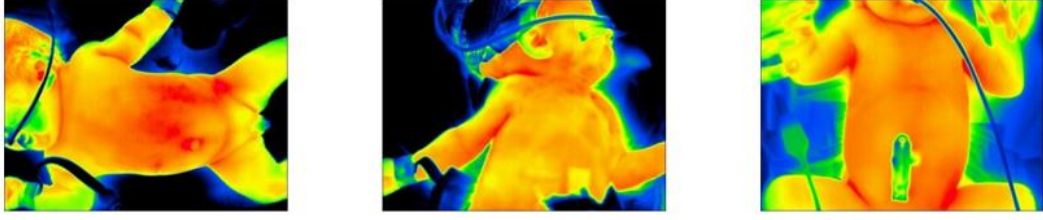
3.1. Yeni Doğan Bebeklere Ait Termal Görüntüler

Yeni doğan bebeklere ait termal görüntüler Selçuk Üniversitesi Yeni Doğan Yoğun Bakım Ünitesi'nden (Tıp Fakültesi) 215E019 TÜBİTAK projesi kapsamında ve 06.01.2015 tarih 2015/1 sayılı etik kurul izniyle alınmıştır. Etik kurul raporu Ek-1'de verilmiştir. Görüntüler alınırken Infratec© Variocam HD termal kamerası kullanılmıştır (Infratec, 2023). Termal kamera 640 x 480 piksel çözünürlüktedir ve 0.01 °C termal duyarlılığa sahiptir (Örnek ve Ceylan, 2020). Görüntü alma sistemi Şekil 3.1'de gösterilmiştir.



Şekil 3.1. Görüntü alma sistemi: (i) termal kamera (ii) yeni doğan bebek (iii) kuvöz (iv) bilgisayar

Şekildeki sistem kullanılarak elde edilen görüntülerden rastgele seçilen bazı örnekler Şekil 3.2'de verilmiştir.



Şekil 3.2. Rastgele seçilen termal görüntüler

Bu sistem ile elde edilen 19 hasta ve 19 sağlıklı bebeğe ait 3800 termal görüntü tez çalışmasında kullanılmıştır. Yeni doğan bebeklere ait hastalık bilgileri Çizelge 3.1’de gösterilmiştir.

Çizelge 3.1. Bebeklere Ait Hastalık Bilgileri (Ornek ve diğ., 2019)

İndeks	1. Hastalık	2. Hastalık
1	NEC	PDA
2	Hydrocephalus	Myelomen ingocele
3	Respiratory distress	Neonatal syphilis
4	AORT coarctation	Down syndrome
5	IKK	Hypothermia
6	Diaphragm hernia	Adeno
7	Respiratory distress	Moaning
8	NEC	Sepsis
9	Respiratory distress	DAB
10	NEC	—
11	Hypoplastic left heart	—
12	Respiratory distress	—
13	PAB	—
14	Intestinal obstruction	—
15	Ozofagus atrezi	—
16	Ozofagus atrezi	—
17	TTN	—
18	RDS	—
19	IUGR	—

3.2. Maske Görüntüleri

Maske takan ve takmayan insanlara ait görüntüler hem açık kaynak veri setlerinden (kaggle, 2020a; 2020b) hem de Şekil 3.3'ten görüldüğü gibi bir kamera sisteminden elde edilmiştir.



Şekil 3.3. Maske takma durumunu belirlemek amacıyla kurulan sistem

Sistemde kullanılan kamera Huawei M2150-10-EI (Huawei, 2023) ürünüdür. Kameranın görüntü sensörü 5MP, işlem kapasitesi 1 TOPS, çözünürlüğü 2560x1920 pikseldir ve çektiği görüntüleri FTP protokolü ile istenilen sunucuya gönderebilmektedir (Örnek ve diğ., 2021a).

Açık kaynak veri setlerinden elde edilen görüntülerin çözünürlüğü 25x25'ten 1800x1800'e kadar değişmektedir (kaggle, 2020a; 2020b). Yüksek çözünürlüklü görüntülerde birden fazla insan bulunmaktadır. Bu görüntüler içerisinde yüz görüntüleri manuel şekilde maskeli ve maskesiz olarak ayrılmıştır. Açık kaynak veri setlerinden elde edilen görüntülerden bazıları rastgele seçilerek Şekil 3.4'te gösterilmiştir.



Şekil 3.4. Açık kaynak veri setlerinden elde edilen örnek görüntüler

Hem kamera sistemi hem de açık kaynak veri setleri kullanılarak toplamda 23140 görüntülük bir veri seti oluşturulmuş ve “18400 eğitim, 200 geçerleme, 4540 test” verisi olacak şekilde çalışmalarda kullanılmıştır (Örnek ve diğ., 2021b).

3.3. Pascal VOC 2007 ve 2012 Veri Setleri

Pascal VOC veri setlerinin temel amacı gerçek dünyadan alınmış farklı nesnelere ait görüntülerin yapay öğrenme ile ayırt edilmesini sağlamaktır (Everingham ve diğ., 2007; Everingham ve diğ., 2012). Bu veri setlerinde, nesnelere ve görüntülere ait etiketler bulunduğundan sınıflama ve nesne tanıma uygulamaları için kullanılabilir. Pascal VOC veri setleri dört başlık altında toplanan 20 farklı kategoriye sahiptir. Bu başlıklar aşağıdaki gibidir;

1. İnsan
2. Hayvan: kuş, kedi, at, koyun, inek, köpek
3. Araç: otobüs, bisiklet, uçak, bot, araba, motosiklet, tren
4. İç mekan: masa, şişe, televizyon, sandalye, saksı, yemek masası

Veri seti her bir görüntüdeki nesnelerin hangi sınıfa ait olduklarını ve çevreleyici kutuların koordinatlarını içermektedir. Şekil 3.5’te Pascal VOC veri setlerinden rastgele seçilmiş görüntüler bulunmaktadır.



Şekil 3.5. Pascal VOC veri setlerine ait rastgele seçilmiş görüntüler

Tez çalışması kapsamında sınıflayıcı modeller kullanılarak nesne tanıma yapıldığından Pascal VOC veri setleri seçilmiştir. Bu veri setleri görsel AYZ çalışmalarında örnek veri seti olarak kullanılmaktadır. Pascal VOC 2007 veri seti toplamda 20 sınıfa ait 9963 görüntü ve 24640 etiketli nesne içermektedir. Pascal VOC 2012 veri seti 20 sınıfa ait 11530 görüntü ve 27450 etiketli nesne içermektedir.

3.4. Sınıflama

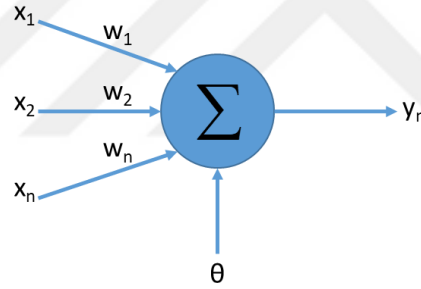
Yapay öğrenme içerisinde sınıflama, temel bir uygulama olarak bir verinin hangi kategoriye ait olduğunun bulunması işlemidir (Robert, 2014). Sınıflanacak veri bir görüntü verisi, bir ses verisi veya bir yazı verisi olabilir. Sınıflama işleminin ilk adımı veriden problemin amacına göre değişiklik gösterebilecek anlamlı özelliklerin çıkarılmasıdır (Khalid ve diğ., 2014). Burada bahsedilen anlamlı özellik tanımlanan problemin amacına göre değişiklik gösterir. Bir görüntü sınıflama işleminde kırmızı meyveleri birbirinden ayıran bir yapay öğrenme modeli eğitildiğini varsayarsak, meyvelere ait anlamlı özellikler şekil, doku veya alan gibi özellikler olacaktır. Ancak tüm veriler baştan kırmızı olarak seçildiği için renk verisi bu problem için anlamlı bir özellik olmayacaktır.

Görüntü sınıflama problemlerinde veriden özellik çıkarma için çeşitli yöntemler kullanılmaktadır. Bunlardan en temel olanları Kenar Bulma ve Yerel İkili Desen gibi yöntemlerdir (Ziou ve Tabbone, 1998; Owotogbe ve diğ., 2019; Zhao ve diğ., 2019; Kaplan ve diğ., 2020). Ancak özelliklerin çıkarılması bir yapay öğrenme modeli

eğitmek için yeterli değildir. Birden fazla sayıda özellik çıkarma yöntemi kullanıldığında çıkarılan özellikler birbirleri ile benzerlik gösterebilirler. Bu durumda özellikler arasında seçim yapılması gerekebilir ve bu yöntemler özellik seçme yöntemleri olarak adlandırılır (Remeseiro ve Bolon-Canedo, 2019; Khaire ve Dhanalakshmi, 2022). Özellik çıkarma ve özellik seçme işlemleri tamamlandıktan sonra elde edilen nihai özellikler Yapay Sinir Ağı (YSA) gibi yapay öğrenme modellerine giriş olarak verilirler. Böylece elde edilen özellikler kullanılarak sınıflama işlemi gerçekleştirilir.

3.5. Yapay Sinir Ağı

YSA'lar insan doğasından esinlenilerek oluşturulmuş ve içeriğinde nöron, aktivasyon fonksiyonları ve giriş-çıkış gibi çeşitli katmanlar bulunduran yapay öğrenme modelleridir (Zupan, 1994). Basit bir nöron yapısı Şekil 3.6'da gösterilmiştir.



Şekil 3.6. Basit bir nöron yapısı

Bir nöron, kendisine gelen giriş (x_n) değerlerini, ağırlık değerleri (w_n) ile ağırlıklandırarak ve bias (θ) değeri ekleyerek çıkış değeri (y_n) elde edilmesini sağlar (Eşitlik 3.1).

$$y_n = \sum_n x_n w_n + \Theta \quad 3.1$$

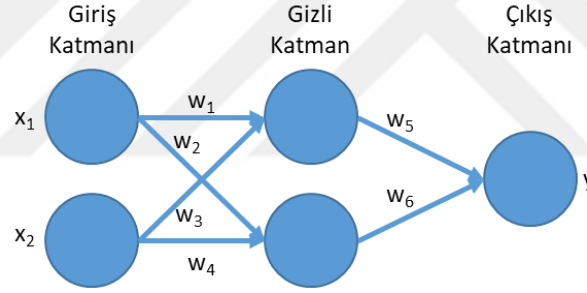
Burada θ bias değeri olarak adlandırılmaktadır ve nöronun öğrenme sürecine dahil edilir. Elde edilen çıkış değeri (y_n) bir aktivasyon fonksiyonundan geçirilerek bu değer belirlili bir aralıkta oluşması sağlanır. Diğer bir ifade ile belirsiz bir aralıkta meydana gelecek çok yüksek veya çok düşük değerlerin önüne geçilmiş olur. Aktivasyon fonksiyonlarına örnek olarak Sigmoid (Eşitlik 3.2) ve ReLU (Eşitlik 3.3)

verilebilir (Marreiros ve diğ., 2008; Sharma ve diğ., 2017; Lau ve Lim, 2018; Schmidt-Hieber, 2020).

$$\text{sig}(y) = \frac{1}{1 + e^{-y_n}} \quad 3.2$$

$$\text{ReLU} = \begin{cases} 0 & \Rightarrow y_n \leq 0 \\ y_n & \Rightarrow y_n > 0 \end{cases} \quad 3.3$$

Sigmoid fonksiyonu, çıkış değerini herhangi bir giriş için 0 ile 1 arasına sıkıştırırken, ReLU fonksiyonu negatif değerlerin sıfırlanmasını gerçekleştirir. Bu aktivasyon fonksiyonları, modeli doğrusal olmayan bir hale getirerek öğrenme sürecinin etkili bir şekilde gerçekleştirilmesine olanak tanır. Bir YSA mimarisi giriş katmanı, gizli katmanlar ve çıkış katmanı içerir. Her bir katman nöronlardan meydana gelir. Örnek bir tek gizli katmanlı YSA mimarisi Şekil 3.7’de gösterilmiştir.



Şekil 3.7. Örnek Bir Tek Gizli Katmanlı YSA Mimarisi

Giriş katmanı özellik seçimiyle elde edilen çıktıları ağırlıklandırarak aktivasyon fonksiyonlarından geçirir ve gizli katmana aktarır. Gizli katman kendine gelen girişleri ağırlıklandırarak aktivasyon fonksiyonlarından geçirir ve çıktı katmanına aktarır. YSA’da bulunan gizli katmanların sayısının artırılması ile derin modeller meydana getirilir ve bu şekilde oluşturulan öğrenme modeli “derin öğrenme” olarak adlandırılır. YSA aynı zamanda tam bağlı katman veya çok katmanlı algılayıcı olarak da isimlendirilmektedir (Gardner ve Dorling, 1998; Basha ve diğ., 2020).

YSA’da öğrenme işlemi belirli adımlarla açıklanabilir. Bu adımlar ileri yayılım, hata hesaplama, geri yayılım ve ağırlık güncellemedir. İleri yayılımda giriş değerleri (x_n) nöronlar ve aktivasyon fonksiyonlarından geçerek çıkış değeri (y_n) elde edilir.

İkinci aşama olarak elde edilen çıkış değeri (y_n) ile gerçek çıkış değeri (y_h) arasındaki fark Ortalama Karesel Hata (Eşitlik 3.4) gibi bir hata fonksiyonu ile hesaplanır (Das ve diğ., 2004; Moore ve DeNero, 2011; Lecué ve Lerasle, 2020; Wang ve diğ., 2020b).

$$hata = \frac{1}{n} \sum (y_h - y_n)^2 \quad 3.4$$

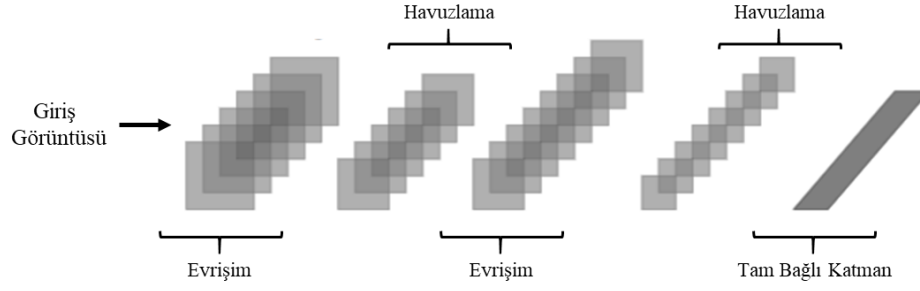
Hata değeri elde edildikten sonra geriye yayılım ile ağırlık değerlerinin (w_n) ne kadar değişeceği hesaplanır ve ağırlık güncelleme ile yeni ağırlık değerleri bulunur (Eşitlik 3.5).

$$w_n = w_{n-1} - r \frac{\partial Hata}{\partial w_{n-1}} \quad 3.5$$

Buradaki “r” öğrenme oranı olarak bilinir ve hesaplanan ağırlık farkının yeni değeri ne kadar etkileyeceğini belirler, buradaki ağırlık farkı bir önceki ağırlık değerinin hataya göre kısmi türevi alınarak ($\partial Hata / \partial w_{n-1}$) hesaplanır. Ağırlık güncelleme işlemi ardından yeniden ileri yayılım gerçekleştirilerek öğrenme işlemine devam edilir. Öğrenme işlemi belirli bir iterasyona ulaşana kadar veya bir hata değerinin altına inene kadar devam ettirilir (Amari, 1993).

3.6. Evrişimli Sinir Ağları

ESA bir derin öğrenme modelidir ve içeriğinde evrişim, havuzlama, aktivasyon fonksiyonları ve tam bağlı katman gibi katmanlar barındırır (YSA burada tam bağlı katman olarak adlandırılmaktadır). Şekil 3.8’de örnek bir ESA mimarisi gösterilmektedir.



Şekil 3.8. Örnek bir ESA Mimarisi

ESA mimarisinde, öncelikle giriş görüntüsünden özellik çıkarma işlemi gerçekleştirilmektedir. Bu işlem evrişim katmanları ile sağlanmaktadır. Bir görüntüye farklı sayılarda evrişim uygulanarak çıkarılan özelliklerin sayısı artırılmaktadır. Şekil 3.9’da örnek bir görüntü (a) ve evrişim filtresi (b) yer almaktadır.

Görüntü Matrisi						
G_{11}	G_{12}	G_{13}	G_{14}	G_{15}	G_{16}	G_{17}
G_{21}	G_{22}	G_{23}	G_{24}	G_{25}	G_{26}	G_{27}
G_{31}	G_{32}	G_{33}	G_{34}	G_{35}	G_{36}	G_{37}
G_{41}	G_{42}	G_{43}	G_{44}	G_{45}	G_{46}	G_{47}
G_{51}	G_{52}	G_{53}	G_{54}	G_{55}	G_{56}	G_{57}
G_{61}	G_{62}	G_{63}	G_{64}	G_{65}	G_{66}	G_{67}
G_{71}	G_{72}	G_{73}	G_{74}	G_{75}	G_{76}	G_{77}

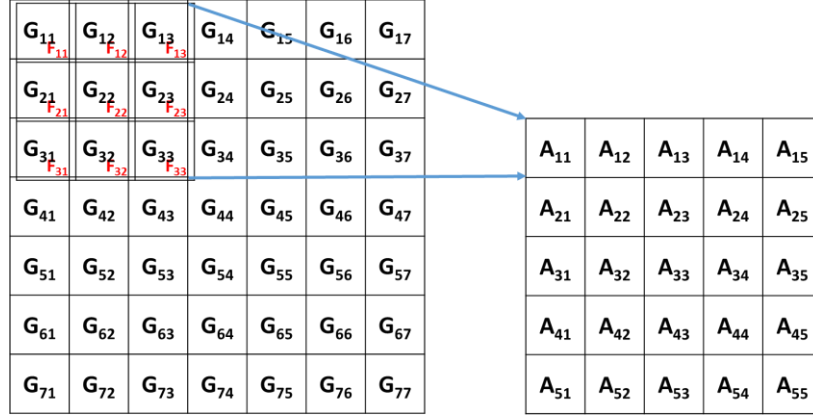
(a)

Filtre Matrisi		
F_{11}	F_{12}	F_{13}
F_{21}	F_{22}	F_{23}
F_{31}	F_{32}	F_{33}

(b)

Şekil 3.9. Örnek bir görüntü ve evrişim filtresi. G değerleri görüntünün piksel değerlerini, F evrişim filtresinin katsayılarını göstermektedir

Evrişim uygulama işleminde ilgili filtre görüntü üzerine yerleştirilerek (Şekil 3.10) eşleşen ağırlıklar birbirleri ile çarpılıp toplanırlar ve filtre kaydırılarak aynı işlem tüm görüntü için tekrarlanır.



Şekil 3.10. Evrişim işlemi

Evrişim işlemi matematiksel olarak Eşitlik 3.6'daki gibi ifade edilir.

$$g[m,n] * f[m,n] = \sum_{j=-\infty}^{+\infty} \sum_{i=-\infty}^{+\infty} g[i,j] \bullet f[m-i,n-j] \quad 3.6$$

Burada g[.] fonksiyonu giriş görüntüsünü, f[.] fonksiyonu uygulanacak filtreyi, i satır ve j ise sütun sayısını temsil etmektedir. Şekil 3.10'daki A₁₁ değerinin hesaplanması için Eşitlik 3.6 kullanıldığında aşağıdaki denklem elde edilmektedir.

$$A_{11} = G_{11}F_{11} + G_{12}F_{12} + G_{13}F_{13} + G_{21}F_{21} + G_{22}F_{22} + G_{23}F_{23} + G_{31}F_{31} + G_{32}F_{32} + G_{33}F_{33}$$

Şekil 3.10'da görüldüğü gibi evrişim işlemi tamamlandıktan sonra görüntü boyutu 7x7'den 5x5'e inmiştir. Bu durum istenmediği durumda görüntü etrafına sıfır ekleme işlemi gerçekleştirilebilir. Şekil 3.11'de örnek bir sıfır ekleme işlemi gösterilmiştir.

0	0	0	0	0	0	0	0	0
0	G ₁₁	G ₁₂	G ₁₃	G ₁₄	G ₁₅	G ₁₆	G ₁₇	0
0	G ₂₁	G ₂₂	G ₂₃	G ₂₄	G ₂₅	G ₂₆	G ₂₇	0
0	G ₃₁	G ₃₂	G ₃₃	G ₃₄	G ₃₅	G ₃₆	G ₃₇	0
0	G ₄₁	G ₄₂	G ₄₃	G ₄₄	G ₄₅	G ₄₆	G ₄₇	0
0	G ₅₁	G ₅₂	G ₅₃	G ₅₄	G ₅₅	G ₅₆	G ₅₇	0
0	G ₆₁	G ₆₂	G ₆₃	G ₆₄	G ₆₅	G ₆₆	G ₆₇	0
0	G ₇₁	G ₇₂	G ₇₃	G ₇₄	G ₇₅	G ₇₆	G ₇₇	0
0	0	0	0	0	0	0	0	0

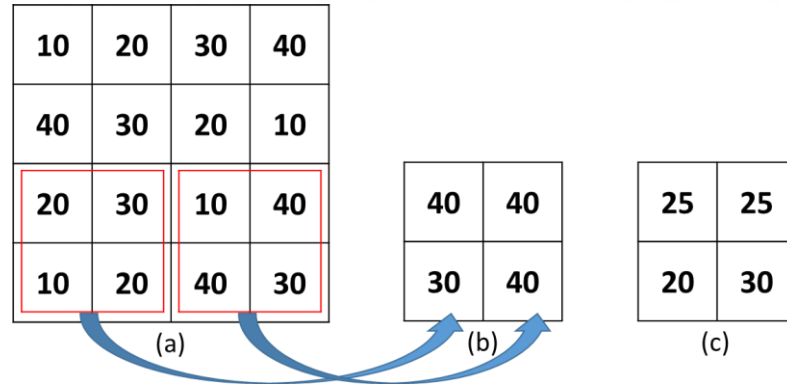
Şekil 3.11. Sıfır eklenmiş görüntü

Evrişim işlemi uygulandığında meydana gelecek görüntünün boyutu Eşitlik 3.7 ile hesaplanır.

$$\text{çıkış boyutu} = \frac{n_g + 2p - k}{s} + 1 \quad 3.7$$

Burada n_g giriş boyutu, p eklenen sıfır katman sayısı, k evrişim boyutu, s ise kaydırma sayısıdır. Şekil 3.11'deki örnek için; sıfır eklenmediğinde $n_g:7$, $p:0$, $k:3$, $s:1$ olacaktır ve çıkış boyutu 5 olarak elde edilecektir. Ancak $p:1$ olduğu durumda $n_g:7$, $k:3$, $s:1$ olacağından çıkış boyutu 7 olarak elde edilir. Yani çıkış boyutu giriş boyutuna eşit olacaktır.

Özellik çıkarma işlemi tamamlandıktan sonra özellik seçme işlemi havuzlama katmanı ile gerçekleştirilir. Havuzlama katmanı belirlenen parametrelere göre evrişim boyutunu düşürerek özelliklerin azaltılmasını sağlar. Maksimum havuzlama veya ortalama havuzlama gibi havuzlama çeşitleri bulunmaktadır. Şekil 3.12'de örnek bir havuzlama işlemi gösterilmiştir.



Şekil 3.12. örnek bir çıktı (b) maksimum havuzlama (c) ortalama havuzlama

Şekil 3.12 (a)'daki örnekte $n_g:4$, $p:0$, $k:2$, $s:2$ olduğundan çıkış boyutu 2 olarak elde edilmiştir. Maksimum havuzlamada 2x2 boyutundaki en yüksek değer seçilmekte (Şekil 3.12 b), ortalama havuzlamada ise 2x2 boyutundaki değerlerin ortalaması alınmaktadır (Şekil 3.12 c). Evrişim ve havuzlama işlemleri iki boyutlu özellikler meydana getirdiğinden tam bağlı katmandan önce düzleştirme katmanı kullanılmaktadır. Düzleştirme katmanı iki boyutlu özelliği tek boyutlu bir vektör haline

getirir. Nihai olarak elde edilen özellikler sınıflayıcı katmana verilerek sınıflama işleminin gerçekleşmesi sağlanır.

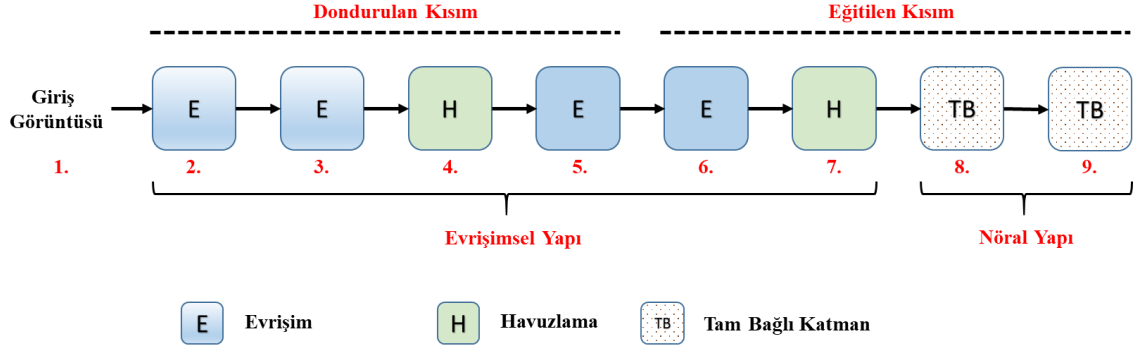
3.7. Transfer Öğrenme

ESA modellerinin ilk evrişim katmanları düşük seviyeli özellikleri (kenar, köşe gibi) öğrenirken son katmanları ise yüksek seviyeli özellikleri (doku gibi) öğrenmektedir. Bu durum, ESA modelleri arasında ilk katmanlarda bulunan evrişim parametrelerinin paylaşılabileceği fikrini ortaya çıkarmıştır (Weiss ve diğ., 2016; Niu ve diğ., 2020; Kim ve diğ., 2022; Yu ve diğ., 2022). Örneğin, bir sınıflandırma işlemi gerçekleştirilirken, ESA modelini sıfırdan eğitmek yerine ilk katmanlarının parametreleri, Resnet (Praveen ve diğ., 2022), Alexnet (Liu ve Deng, 2015), VGG (Mateen ve diğ., 2018) veya Inception (Szegedy ve diğ., 2017) gibi ön eğitilmiş bir modelden aktarılabilmektedir.

Bilgisayarlar üzerinde, model karmaşıklığı veya veri setinin büyüklüğü gibi nedenlerle bazı modellerin eğitilmesi oldukça zor olabilir. Bu nedenle, ESA gibi modellerin eğitim işlemleri için grafik işleme ünitelerinden (GIU) faydalanılmaktadır. Önceden eğitilmiş bir ESA modelini GIU'lerle gerçekleştirilecek işleme adapte etmek daha etkin sonuçlar elde etmeyi sağlar (Ornek ve Ceylan, 2022). Önceden eğitilmiş ESA mimarisinin evrişim katmanı kısmı sabit tutularak, yalnızca çıkış katmanındaki sınıfları değiştirmek veya evrişim kısımda bulunan belirli ağırlıkları ve katmanları değiştirmek gibi transfer öğrenme yöntemleri kullanılmaktadır. Bu şekilde, eğitilmiş modellerin bilgi ve özellikleri, yeni bir görev veya veri setine adapte edilerek kullanılabilir.

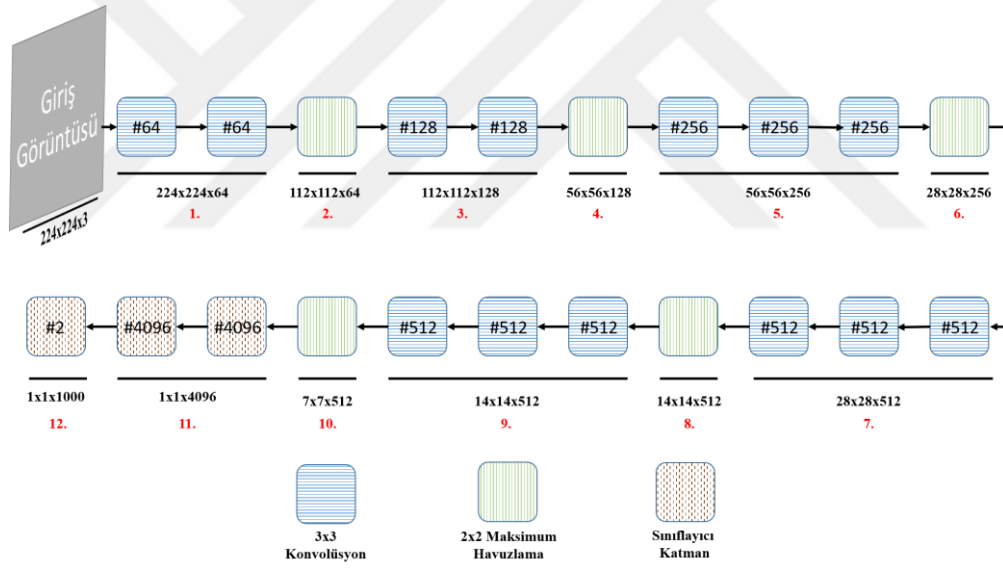
Transfer öğrenme, ESA modellerinin verimli bir şekilde kullanılmasını sağlayan önemli bir yaklaşımdır. Bu yöntem, hem eğitim süresini kısaltmak hem de az veriye sahip verisetleri üzerinde bile iyi performans elde etmek için kullanılmaktadır. Önceden eğitilmiş modellerin başarısı ve ağ genelleme kabiliyeti, farklı alanlardaki problemleri çözmede büyük bir avantaj sağlamaktadır.

Şekil 3.13'in önceden eğitilmiş bir ESA modeline ait olduğu düşünülürse, transfer öğrenme gerçekleştirmek için bu mimarinin 2. ve 5. bölümleri arası korunur ve eğitilebilirliği dondurulur. Kalan kısımlar da probleme özgü olarak düzenlenir ve eğitim işlemi sırasında yalnızca bu kısımların eğitilmesi sağlanır.



Şekil 3.13. Örnek bir transfer öğrenme yapısı

Ön eğitilmiş modellerden biri olan VGG16 1000 sınıf içeren bir görüntü sınıflama probleminin çözülmesi için eğitilmiş bir derin öğrenme modeldir. Bu mimari kullanılarak iki sınıflı bir görüntü sınıflama problemi çözülmek istendiğinde Şekil 3.14'teki gibi bir mimari oluşturulmalıdır.



Şekil 3.14. Transfer Öğrenme Yönteminin 2 Sınıflı Bir Probleme Adapte Edilmesi

VGG16'da 12. kısımda 1000 adet çıktı varken bu sayı 2'ye indirilmiştir. Diğer bir deyişle 2 sınıflı bir görüntü sınıflandırma problemi çözüleceği için son katmanda bulunan nöron sayısı 1000'den 2'ye indirilmiştir. Eğitim işlemi sırasında 1. kısımdan 8. kısıma kadar olan ağırlıkların öğrenilebilirliği durdurularak geriye kalan ağırlıklar eğitilmektedir.

3.8. Temel Bileşenler Analizi

Sınıflama bölümünde açıklandığı gibi bir veri doğası gereği birbirinden farklı özellikler içermektedir. Ancak tüm bu özellikler çözülmek istenen problem ile ilgili olmayabilir. Bu sebeple özellik çıkarma işlemi gerçekleştirildikten sonra özellik seçme veya boyut azaltma yöntemleri kullanılmaktadır (Espadoto ve diğ., 2019).

Temel Bileşenler Analizi (TBA) boyut azaltma yöntemlerinden birisi olarak kendisine gelen özellikleri özdeğerlere göre sıralar ve önemli özelliklerin ortaya çıkarılmasını sağlar (Rocchi ve diğ., 2004; Ringnér, 2008; Vidal ve diğ., 2016). TBA algoritması aşağıdaki gibidir:

1. Giriş değerlerinin ortalaması hesaplanır
2. Ortalama değer tüm giriş değerlerinden çıkarılır
3. Elde edilen değerlerden kovaryans matrisi oluşturulur
4. Kovaryans matrisinden özdeğerler ve özvektörler hesaplanır
5. Özdeğerler büyükten küçüğe doğru sıralanır
6. Sıralı özdeğerler arasından istenilen sayıda özellik seçilir
7. Özdeğerlere karşılık gelen özvektörler seçilir
8. Özvektörler ile giriş değerleri çarpılarak çıkış elde edilir

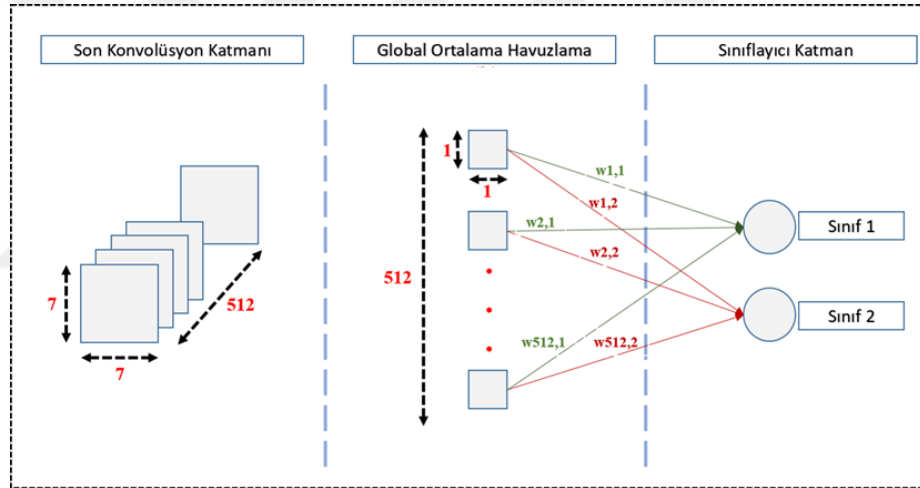
Belirli bir özellik sayısını manuel olarak seçmek yerine bu işlem otomatikleştirmek isteniyorsa, bu işlem varyans değeri kullanılarak gerçekleştirilebilir. Özdeğerler büyükten küçüğe doğru sıralandıktan sonra ilgili özvektörlerin varyans değerleri hesaplanır (Kramer ve Kramer, 2016). Bu özvektörler arasında bir seçim yapılmadığında tüm vektörlere karşılık gelen varyans değerlerinin toplamı bire eşittir. Bu bağlamda, toplam değeri belirli bir eşik değerinin üzerine çıkana kadar ilk vektörden itibaren bir toplama işlemi gerçekleştirilir. Söz konusu eşik değeri (örneğin 0.25, 0.35, 0.85, 0.95), araştırmacı tarafından deneysel olarak belirlenir.

3.9. Sınıf Aktivasyon Haritaları

SAH görsel bir açıklanabilir yapay zeka yöntemidir ve ESA mimarisinin son evrişim katmanında bulunan filtrelerin görselleştirilmesine dayanır (Jung ve Oh, 2021). Bir sınıflama işlemi gerçekleştirilmeden önce bir görüntü seti hazırlanarak bu

görüntüler sınıf isimleri ile etiketlenmektedir. Eğitim işlemi sırasında ise görüntüden evrişim işlemi ile özellikler çıkarılmakta ve her bir evrişim işleminde farklı sayıda filtreler elde edilmektedir.

İlk evrişim işlemi sonucunda ortaya çıkan filtreler kenar ve köşe gibi düşük nitelikli özelliklere odaklanırken, daha sonraki evrişim işlemleri sonucunda elde edilen filtreler, doku gibi yüksek nitelikli özelliklere odaklanmaktadır. Buradaki problem ise son evrişim katmanındaki filtrelerin hangisinde bu doku özelliğinin bulunduğu bilinmemesidir. Örneğin tez çalışması kapsamında kullanılan modellerden biri olan Resnet18, ön eğitilmiş transfer öğrenme modelidir ve derin ESA ağı eğitilmesine olanak tanımaktadır. Resnet18'in son katmanında 512 adet 7×7 boyunda filtre bulunmaktadır. Şekil 3.15'te Resnet18'in son 3 katmanı olan son evrişim katmanı, global havuzlama katmanı ve sınıflayıcı katmanı gösterilmiştir.

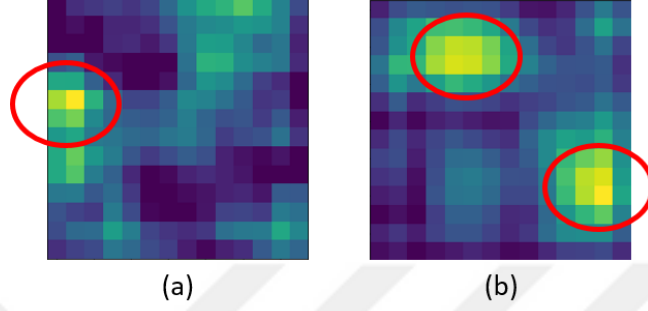


Şekil 3.15. Resnet18 Mimarisi Son 3 Katman

Bu mimariye göre son konvolüsyon katmanından sonra bir global havuzlama katmanı kullanılmaktadır. Global ortalama havuzlama ile beraber 7×7 olan filtre boyutları 1×1 boyutuna indirilmekte, yani katmanda bulunan 49 farklı değer tek bir sayı ile ifade edilmektedir. Bu işlem bir özellik azaltma işlemi olarak da düşünülebilir. Global ortalama havuzlama işleminden sonra ise bir sınıflayıcı katman bulunmaktadır. Eğitim işlemi tamamlandığında filtreler ağırlıklandırılarak (Eşitlik 3.8) Şekil 3.16'deki gibi sınıf aktivasyon haritaları elde edilir. Şekil 3.16 (a)'da bir yoğunlaşmış bölge, (b)'de iki yoğunlaşmış bölge bulunmaktadır. Aktivasyonların yoğunlaştığı yerler ilgili nesnenin hangi bölümlerinin sınıflama için önemli olduğunu gösterir.

$$L^s = w^s A \quad 3.8$$

Burada s sınıf bilgisini, L^s s sınıfına ait aktivasyon haritasını, w^s s sınıfa karşılık gelen ağırlıkları ve A evrişimli katmanı ifade etmektedir. Bu şekilde yapılan görselleştirme işlemi klasik SAH olarak adlandırılmaktadır.



Şekil 3.16. Çeşitli Sınıf Aktivasyon Haritası Çıktıları (a) tek yoğunlaşmış bölge (b) yoğunlaşmış iki bölge

Klasik SAH ile görselleştirme sağlanması için model mimarisinin Şekil 3.15'teki gibi düzenlenmesi ve modelin yeniden eğitilmesi gerekmektedir (Zhou ve diğ., 2016). Ancak bu mimaride tam bağlı katman kullanılmadığından bu şekilde eğitilen modeller çok yüksek başarımlar gösterememektedir (Selvaraju ve diğ., 2017). Bu dezavantajları ortadan kaldırmak için Selvaraju ve diğ. tarafından, Eşitlik 3.8'de bulunan w^s ağırlıklarının elde edilmesi için gradyanlardan faydalanılması gerektiği ortaya konulmuştur (Eşitlik 3.9). Böylece model mimarisinde bir değişiklik yapmadan aktivasyon haritaları elde edilmekte ve görselleştirme sağlanmaktadır. Bu yöntem GradCAM yöntemi olarak adlandırılmaktadır.

$$w_{(GSAH)}^s = \frac{1}{N} \sum \frac{\partial Y^s}{\partial A} \quad 3.9$$

Gradyan değerleri Eşitlik 3.9'a göre hesaplandıktan sonra Eşitlik 3.8'te yerine koyularak aktivasyon haritası oluşturulmaktadır. GradCAM++ (Chattopadhyay ve diğ., 2018) yöntemi ise ağırlık değerleri hesaplanırken pozitif değerlerden ve ikinci derece gradyanlardan faydalanılması gerektiğini ortaya koymuştur. Buna göre ağırlık değerleri Eşitlik 3.10'a göre hesaplanmakta ve Eşitlik 3.8'de yerine koyularak aktivasyon haritaları meydana getirilmektedir.

$$w_{(GSAH++)}^s = \sum \alpha^s \text{ReLU} \left(\frac{\partial Y^s}{\partial A} \right) \quad 3.10$$

Eşitlik 3.10’da α^s piksel düzeyinde gradyanları ifade etmektedir. CAM, GradCAM ve GradCAM++ yöntemleri son evrişim katmanında bulunan filtreleri kullanarak aktivasyon haritaları oluştururken, LayerCAM (Jiang ve diğ., 2021) yöntemi sadece son katmandan faydalanmak yerine ara katmanlardan (A') da faydalanarak aktivasyon haritaları üretme fikrini ortaya atmıştır. Eşitlik 3.11’de görüldüğü gibi öncelikle farklı katmanlardaki aktivasyonlar ilgili ağırlıklar ile ağırlıklandırılmakta ardından Eşitlik 3.12’deki gibi doğrusal birleşimleri sağlanmaktadır.

$$A' = w^s A \quad 3.11$$

$$L_{(LayerCAM)} = \text{ReLU} \left(\sum A' \right) \quad 3.12$$

EigenCAM (Bany Muhammad ve Yeasin, 2021) çalışmasında, daha hızlı bir şekilde CAM sonuçları elde etmek için son evrişim katmanında bulunan tüm filtrelerin kullanılmaması gerektiği fikri ortaya atılmıştır. Son evrişim katmanındaki bütün filtreleri kullanmak yerine, bir boyut azaltma yöntemi olan TDA’lardan (Eşitlik 3.13) faydalanarak filtre sayısının bire indirilmesi sağlanmıştır.

$$A = UEV^t \quad 3.13$$

Eşitlik 3.13’te V ile U ortogonal matris, E diagonal matris olarak bilinir. V sağ tekil vektörleri, U sol tekil vektörleri ve E tekil değerleri ifade eder. V ’nin ilk elementi V_1 , EigenCAM tarafından temel bileşen olarak kullanılır. Bu değer Eşitlik 3.14’te yerine koyularak EigenCAM çıktıları elde edilir.

$$L_{(EigenCAM)} = AV_1 \quad 3.14$$

Özetlenecek olursa klasik SAH yöntemi son katmandaki ağırlıklar ile son katmanda bulunan filtreleri ağırlıklandırmaktadır. GradCAM’de son katmandaki ağırlıklar birinci dereceden gradyan hesabı ile bulunurken, GradCAM++’da ikinci dereceden gradyan hesabı ile bulunur. LayerCAM son katman dışında orta katmanları

da kullanır. EigenCAM ise son katmanda bulunan filtre sayısının TDA ile bire indirmesini sağlar.

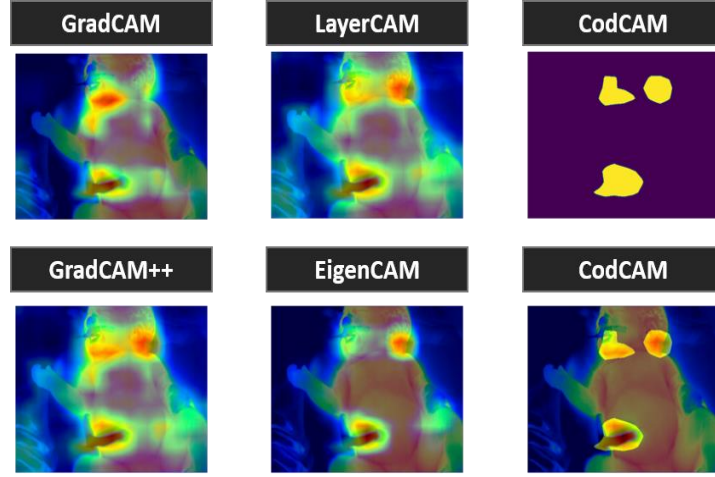
3.10. CodCAM

Bu tez çalışması kapsamında geliştirilmiş yöntemlerden birisi olan CodCAM (Ornek ve Ceylan, 2023), literatürde mevcut olan GradCAM, GradCAM++, LayerCAM ve EigenCAM yöntemlerinin aktivasyon haritalarını alarak (Eşitlik 3.15) ortak bir karar verilmesi (Eşitlik 3.16) fikrine dayanmaktadır. Yeni doğan bebeklere ait termal görüntülerde olduğu gibi gerçek çıktının bilinmediği uygulamalarda veya hassasiyetin ön planda olduğu nesne tespiti gibi uygulamalarda, CodCAM farklı yöntemlerin ortak kararlarını nihai bir karar verilmesini sağlamaktadır.

$$L_{(ortak)} = L_{(GSAH)} + L_{(GSAH+)} + L_{(LayerCAM)} + L_{(EigenCAM)} \quad 3.15$$

$$L_{(CodCAM)} = L_{ortak} > 2 \quad 3.16$$

Her bir AYZ yöntemi çıkış olarak 0 ve 1 arasında değerler üretmektedir. CodCAM'e göre her bir AYZ yönteminin ürettiği değerlerin toplamının en az 2 olması gerekmektedir. Yani bir pikselin önemli olarak kabul edilebilmesi için ortalama olarak en az 0.5 değeri olmalıdır. Bir örnek ile açıklanırsa, giriş görüntüsünün bir pikseli GradCAM'den 0.55, GradCAM++'dan 0.72, LayerCAM'den 0.64 ve EigenCAM'den 0.58 çıktı aldıysa bu piksel önemli bir piksel olarak ele alınmakta ve aktivasyon haritasında gösterilmektedir. Ancak eğer bir piksel GradCAM'den 0.33, GradCAM++'dan 0.12, LayerCAM'den 0.60 ve EigenCAM'den 0.49 aldıysa bu piksel göz ardı edilmekte ve aktivasyon haritasında yer almamaktadır. Her bir yöntemin çıktısı ayrıca görselleştirildiğinden (Şekil 3.17) göz ardı edilen piksellerin bir probleme neden olabileceği düşünülmemelidir. Buradaki amaç hem tüm yöntemlerin çıktılarını hem de ortak bir kararı değerlendiriciye sunmak ve karar vermesine destek olmayı sağlamaktır.



Şekil 3.17. CodCAM ve diğer yöntemlerin çıktılarının görselleştirilmesi

3.11. HayCAM

Tez kapsamında gerçekleştirilen bir diğer çalışma olan HayCAM (Örnek ve Ceylan, 2022b), EigenCAM yaklaşımında olduğu gibi son evrişim katmanındaki tüm filtrelerin kullanılması yerine seçilen filtrelerin kullanılmasını önermektedir. Çünkü son katmanda hedef sınıfla ilgili anlamlı özellikleri öğrenemeyen filtreler de bulunur. Örneğin tez çalışmasında kullanılan ESA modellerinden biri olan Resnet18'in son evrişim katmanında 512 filtre bulunmaktadır. Bütün filtrelerin ve seçilmiş filtrelerin kullanıldığı durumlarda meydana gelen görsel ve sayısal farklılıklar Bölüm 4.2'de verilmiştir.

HayCAM boyut azaltmayı TDA ile yapmak yerine istenilen sayıda filtre seçimini sağlamak için TBA ile gerçekleştirmektedir. Ağırlıklar ve filtreler Eşitlik 3.17'ye göre seçildikten sonra görselleştirme işlemi Eşitlik 3.18'e göre yapılır.

$$w_{s(HayCAM)}, A_{(HayCAM)}^s = TBA \left(\frac{1}{N} \sum \frac{\partial Y^s}{\partial A}, A^s \right) \quad 3.17$$

$$L_{(HayCAM)} = w_{s(HayCAM)} A_{(HayCAM)}^s \quad 3.18$$

HayCAM algoritması aşağıdaki gibidir:

1. Giriş görüntüsü ESA'dan geçirilerek evrişim filtreleri elde edilir
2. Filtreler $n \times n$ boyutundan $n^2 \times 1$ boyutuna çevrilir
3. TBA ile ilk k adet filtre ve bu filtrelere karşılık gelen ağırlıklar seçilir
4. Filtreler ağırlık değerleri ile ağırlıklandırılarak aktivasyon haritası elde edilir
5. Aktivasyon haritası $n^2 \times 1$ boyutundan $n \times n$ boyutuna çevrilir

3.12. HayCAM+

HayCAM yönteminin dezavantajlarından birisi TBA'da kullanılan filtre sayısının (örneğin 19) deneysel olarak hesaplanmasıdır. Bu durumu ortadan kaldırmak için tez kapsamında HayCAM+ yöntemi geliştirilmiştir. HayCAM+, TBA'da kullanılacak filtre sayısını özvektörlerin varyans değerlerini kullanarak hesaplar (Eşitlik 3.19) ve kullanılacak filtreler ağırlıklandırma yapmak yerine direkt olarak toplanır (Eşitlik 3.20). Böylece aktivasyon haritaları oluşturulur.

$$A_{(HayCAM+)}^s = TBA(A^s) \quad 3.19$$

$$L_{(HayCAM+)} = \sum A_{(HayCAM+)}^s \quad 3.20$$

HayCAM+ algoritması aşağıdaki gibidir:

1. Tüm filtreler kullanılarak kovaryans matrisi oluşturulur
2. Öz değerler ve öz vektörlerin hesaplanır
3. Öz değerler sıralanır
4. Her öz vektöre karşılık gelen varyans değeri hesaplanır
5. Her varyans değeri toplam varyans değerine bölünür
6. En yüksek öz değerli vektörden başlayarak varyans değerleri toplanır
7. k . öz vektörde varyans toplamı eşik değerini geçtiği anda işlem sonlandırılır
8. Seçilen filtreler toplanarak aktivasyon haritası oluşturulur
9. Aktivasyon haritası $n^2 \times 1$ boyutundan $n \times n$ boyutuna çevrilir

Varyans toplamı için geçerli eşik değeri araştırmacı tarafından birkaç deneme sonucunda belirlenir. Çeşitli varyans eşik değerlerine göre HayCAM+'ın performans değerleri Bölüm 4.5'te gösterilmiştir. HayCAM yöntemi ile karşılaştırıldığında, HayCAM yönteminde kullanılacak filtre sayısı deneysel olarak belirlenirken, HayCAM+ yönteminde varyans değeri deneysel olarak belirlenmektedir.

3.13. HayCAMJ

Tez çalışması kapsamında geliştirilen son yöntem olarak HayCAMJ, TBA ile son evrişim katmanındaki filtrelerin bire indirilmesini sağlamaktadır. Diğer bir deyişle TBA'daki temel bileşen ana filtre olarak seçilmektedir. EigenCAM'den farklı olarak filtreler TDA yerine TBA ile boyut azaltma işleminden geçirilmektedir. HayCAMJ algoritması aşağıdaki gibidir:

1. Tüm filtreler kullanılarak kovaryans matrisi oluşturulur
2. Öz değerler ve öz vektörlerin hesaplanır
3. Öz değerler sıralanır
4. En yüksek öz değerli vektör alınır
5. Aktivasyon haritası elde edilir
6. Aktivasyon haritası $n^2 \times 1$ boyutundan $n \times n$ boyutuna çevrilir

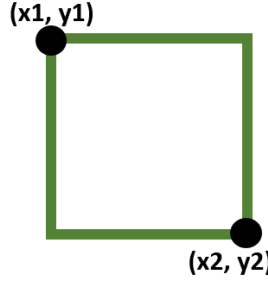
Tez çalışması kapsamında geliştirilen CodCAM, HayCAM, HayCAM+ ve HayCAMJ yöntemlerinin farklılıkları Çizelge 3.2'de verilmiştir.

Çizelge 3.2. Geliştirilen Yöntemlerin Farklılıkları

Yöntem	Hiperparametre	Açıklama
CodCAM	-	AYZ yöntemlerinden elde edilen tüm filtreler nihai çıktının elde edilmesi için kullanılmaktadır.
HayCAM	filtre sayısı	Filtre sayısı araştırmacı tarafından belirlenmektedir. TBA'da filtreler azaltıldıktan sonra ilgili ağırlıklar ile ağırlıklandırılır.
HayCAM+	varyans değeri	Varyans değeri araştırmacı tarafından belirlenmektedir. TBA'da filtreler azaltıldıktan sonra birim ağırlıklar ile ağırlıklandırma yapılır.
HayCAMJ	-	TBA'da elde edilen temel filtre kullanılır.

3.15. Zayıf Danışmanlı Nesne Tespiti

Bir nesne tanıma işlemi gerçekleştirilmek istendiğinde Şekil 3.18’de görüldüğü gibi nesneye ait dört farklı koordinat değerinin hesaplanması gerekmektedir (Amit ve diğ., 2020).



Şekil 3.18. Örnek bir çevreleyici kutu

Bir görüntüde onlarca nesne olduğu düşünüldüğünde, her bir nesnenin ayrı ayrı etiketlenmesi zaman alıcı bir işlemdir. Zayıf danışmanlı nesne tespiti, bu süreci hızlandırmak için kullanılan yaklaşımlardan birisidir (Yang ve diğ., 2019b; Cheng ve diğ., 2020; Zhang ve diğ., 2021). Bu yaklaşımda, bir görüntüdeki nesnelerin hangi sınıfa ait olduğu belirtilirken, nesnelerin koordinatları bir sınıflayıcı model tarafından tahmin edilmektedir. Bu şekilde, tüm nesnelerin ayrıntılı etiketlemesi yerine sadece sınıf bilgisi ile birlikte nesnelerin konumlarının belirlenmesi sağlanmaktadır.

Zayıf danışmanlı nesne tespitinde en yaygın olarak kullanılan yöntemlerden biri SAH’lardır (Bae ve diğ., 2020). SAH yöntemleri, bir görüntüdeki nesne konumlarını tahmin etmek için aktivasyon haritalarının yoğunlaştığı piksellerin ilgili nesne olduğunu ortaya çıkarmaktadır. Böylece SAH yöntemleri, nesne tespiti ve sınıflandırma arasında bir bağlantı kurarak nesne konumlarının tahmin edilmesini sağlar.

Bu tez çalışmasında, literatürde mevcut olan SAH yöntemleri ile beraber yeni geliştirilen CodCAM, HayCAM, HayCAM+ ve HayCAMJ yöntemleri de kullanılarak zayıf danışmanlı nesne tespiti uygulamaları yapılmıştır.

3.15. Değerlendirme Metrikleri

Tez çalışması kapsamında sınıflayıcı modellerin performansının hesaplanması için karmaşıklık matrisleri oluşturulmuş (Çizelge 3.3) ve doğruluk (Eşitlik 3.21), hassasiyet (Eşitlik 3.22) ve özgüllük (Eşitlik 3.23) değerleri hesaplanmıştır (Marino ve diğ., 2013).

Çizelge 3.3. İkili sınıfa ait bir karmaşıklık matrisi

DP	YP
YN	DN

Çizelge 3.3’deki, DP (doğru pozitif) ilk sınıfa ait doğru olarak sınıflandırılan verilerin sayısını, DN (doğru negatif) ikinci sınıfa ait doğru olarak sınıflandırılan verilerin sayısını, YP (yanlış pozitif) birinci sınıfa ait yanlış olarak sınıflandırılan verilerin sayısını ve YN (yanlış negatif) ikinci sınıfa ait yanlış olarak sınıflandırılan verilerin sayısını belirtmektedir.

$$\text{Doğruluk} = \frac{DP + DN}{DP + YP + YN + DN} \quad 3.21$$

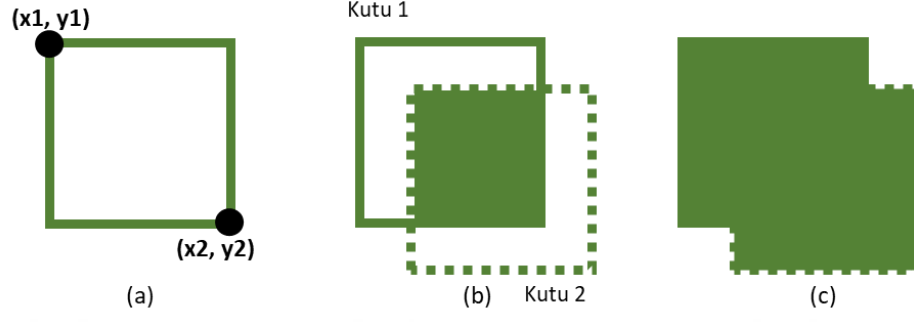
$$\text{Hassasiyet} = \frac{DP}{DP + YP} \quad 3.22$$

$$\text{Özgüllük} = \frac{DN}{YN + DN} \quad 3.23$$

Şekilde 3.19’te görüldüğü gibi nesne tanıma probleminin amacı bir nesneye ait dört parametrenin belirlenmesidir. Nesne tanıma performansının ölçülmesi için ise IoU (Eşitlik 3.24) ve Dice (Eşitlik 3.25) metrikleri kullanılmıştır (Zhou ve diğ., 2019; Kang ve diğ., 2020). Bu değerler [0,1] aralığında oluşmakta ve 1’e yakın sonuçlar performansın daha iyi olduğunu göstermektedir. Eşitliklerde bulunan kesişim ve birleşim bölgeleri Şekil 3.19 (a) ve (b)’de sırasıyla temsil edilmiştir. Düz çizgiyle görselleştirilen Kutu 1’i, kesikli görselleştirilen belirtilen Kutu 2’yi ifade etmektedir.

$$\text{IoU} = \frac{|Kutu1| \cap |Kutu2|}{|Kutu1| \cup |Kutu2|} \quad 3.24$$

$$\text{Dice} = \frac{2 * (|Kutu1| \cap |Kutu2|)}{|Kutu1| + |Kutu2|} \quad 3.25$$



Şekil 3.19. (a) çevreleyici kutu (b) kesişim (c) birleşim

4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA

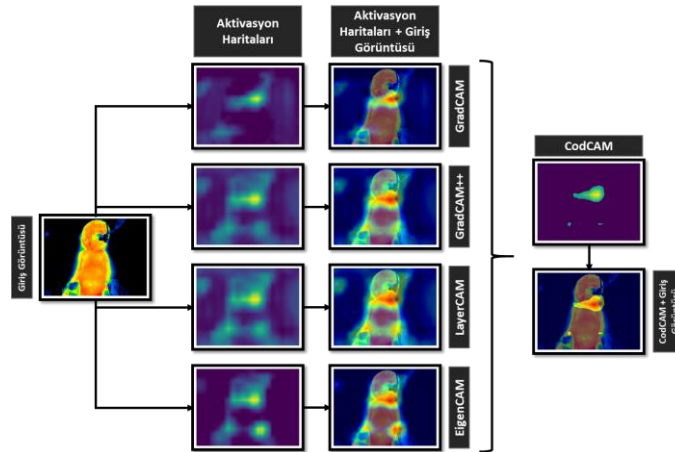
Tez çalışması kapsamında beş farklı uygulama gerçekleştirilmiştir:

1. CodCAM yönteminin medikal termal görüntüler üzerindeki etkisi,
2. HayCAM yönteminin maske tanıma problemi üzerindeki etkisi,
3. HayCAM+ yönteminin maske tanıma üzerindeki etkisi,
4. HayCAMJ yönteminin Pascal VOC 2007 ve 2012 veri setleri üzerindeki etkisi
5. Geliştirilen dört yöntemin Pascal VOC 2007 ve 2012 veri setleri üzerindeki etkisi

4.1. CodCAM Yönteminin Medikal Görüntüler Üzerindeki Sonuçları

Medikal uygulamalar için yapay zeka modellerinin açıklanması hayati önem taşımaktadır. Yapay zekanın verdiği kararlar doktorlar tarafından anlaşılmadığında verilen karara duyulan güven azalmaktadır (Antoniadi ve diğ., 2021). Bu nedenle yapay zeka modelinin görüntünün hangi kısmına bakarak karar verdiğinin ortaya çıkarılması verilen karara duyulan güvenin artmasını sağlamaktadır.

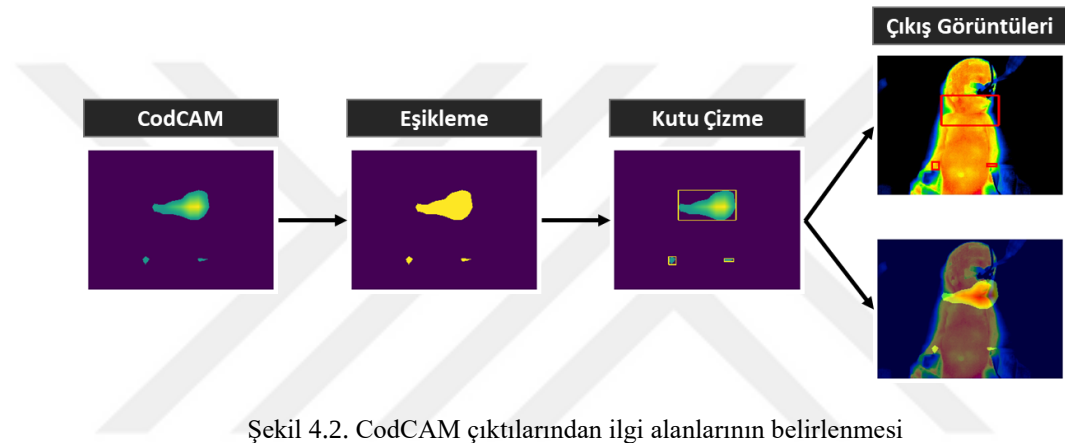
CodCAM çalışmasında, yeni doğan bebeklerin hasta olup olmadıklarını ortaya çıkaran yapay zeka modellerinin “Kara Kutu” doğasını açmak için dört farklı görsel AYZ yöntemi (GradCAM, GradCAM++, LayerCAM, EigenCAM) kullanılmış ve önerilen CodCAM yönteminin sonuçları da gösterilmiştir (Şekil 4.1).



Şekil 4.1. Önerilen CodCAM yönteminin görselleştirilmesi

Önerilen yaklaşımda, bir termal görüntü yapay zeka modeline verildiği durumda öncelikle dört farklı görsel AYZ yöntemi ile görselleştirme gerçekleştirilmektedir. Buradaki amaç bir topluluk kararı oluşturulmasını sağlamaktır. CodCAM diğer dört yöntemin ortak çıktısını alarak nihai bir görselleştirme sunar.

CodCAM çalışmasının sonuçları şu şekildedir: (i) Medikal termal görüntülerin sağlıklı ve hasta olarak sınıflandırılması, (ii) Sınıflama için kritik olan alanların dört farklı AYZ yöntemi tarafından çıkarılması ve görselleştirilmesi, (iii) Yeni bir yöntem geliştirilmesi ve (iv) Şekil 4.2'de gösterildiği gibi uzman değerlendirmesi için ilgi alanlarının tanımlanması.



Şekil 4.2. CodCAM çıktılarından ilgi alanlarının belirlenmesi

Şekil 4.2’de eşikleme yapmak için OpenCV’nin “inRange” özelliği, eşikleme yaptıktan sonra kenarların bulunması için “findContours” özelliği ve kenarları içeren en iyi kutunun tespit edilmesi için “boundingRect” özelliği kullanılmıştır. Burada kullanılan parametreler şu şekildedir: `inRange(0.3)`, `findContours(retr_tree, chain_approx_simple)`, `boundingRect(edges)`.

Aktivasyon haritalarının oluşturulması için birinci aşama sınıflama modelinin eğitilmesidir. Medikal görüntülerin sınıflanması için sınıflama performansı yüksek olan Resnet50 ön-eğitilmiş modeli kullanılmıştır. Termal görüntüler hasta ve sağlıklı olmak üzere 2 sınıfa ait oldukları için Resnet50’nin son katmanında bulunan 1000 nöron kaldırılarak yerine 2 nöron yerleştirilmiş ve Çizelge 4.1’deki özellikler kullanılarak model eğitimi gerçekleştirilmiştir. Eğitim için 2280, geçerleme için 380 ve test için 1140 görüntü kullanılmıştır.

Çizelge 4.1. CodCAM için kullanılan donanım ve yazılım özellikleri

Derin Öğrenme Kütüphanesi	PyTorch
Ortam	Linux
Programlama Dili	Python
GIÜ	Tesla K80
Tüm Görüntüler	3800 (38x100)
Eğitim Görüntüleri	2280 (38x60)
Geçerleme Görüntüleri	380 (38x10)
Test Görüntüleri	1140 (38x30)

Resnet50 modeli, Çizelge 4.2’de görülen farklı parametreler ile eğitilip test edildiğinde (2280 Eğitim, 380 Geçerleme, 1140 Test verisi) sonuçlar Şekil 4.3’teki gibi elde edilmektedir.

Çizelge 4.2. Resnet50 modelinin farklı parametreler ile eğitilmesi

Deney No	Optimizasyon Algoritması	Öğrenme Oranı	Momentum	Eğitim (%)	Geçerleme (%)	Test (%)
1	SGD	0.01	0.8	99	77	91
2	SGD	0.01	0.3	99	98	87
3	SGD	0.001	0.5	94	36	82
4	SGD	0.001	0.8	99	99	99
5	ADAM	0.01	-	92	97	70
6	ADAM	0.001	-	99	98	88
7	ADAM	0.0001	-	99	72	91

513	57	570	0	514	56	568	2
42	528	147	423	146	424	0	570
1		2		3		4	
327	243	533	37	527	43		
90	480	96	474	53	517		
5		6		7			

Şekil 4.3. Resnet50 modelinin farklı parametreler ile eğitildiğinde elde edilen test sonuçları

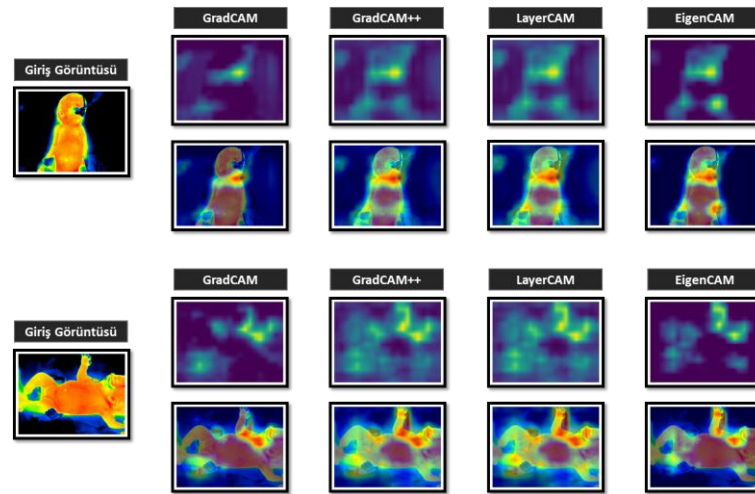
Çizelge 4.2’den görüldüğü gibi, en iyi sonuç Deney 4’te elde edilmiştir. Resnet50 modelinin en iyi şekilde öğrendiği durumda elde edilen eğitim, geçerleme ve test sonuçları Şekil 4.4’te verilmiştir.

1134	6	189	1	568	2
0	1140	1	189	0	570
Eğitim		Geçerleme		Test	

Şekil 4.4. Eğitilen Resnet50 modelinin eğitim, geçerleme ve test sonuçları

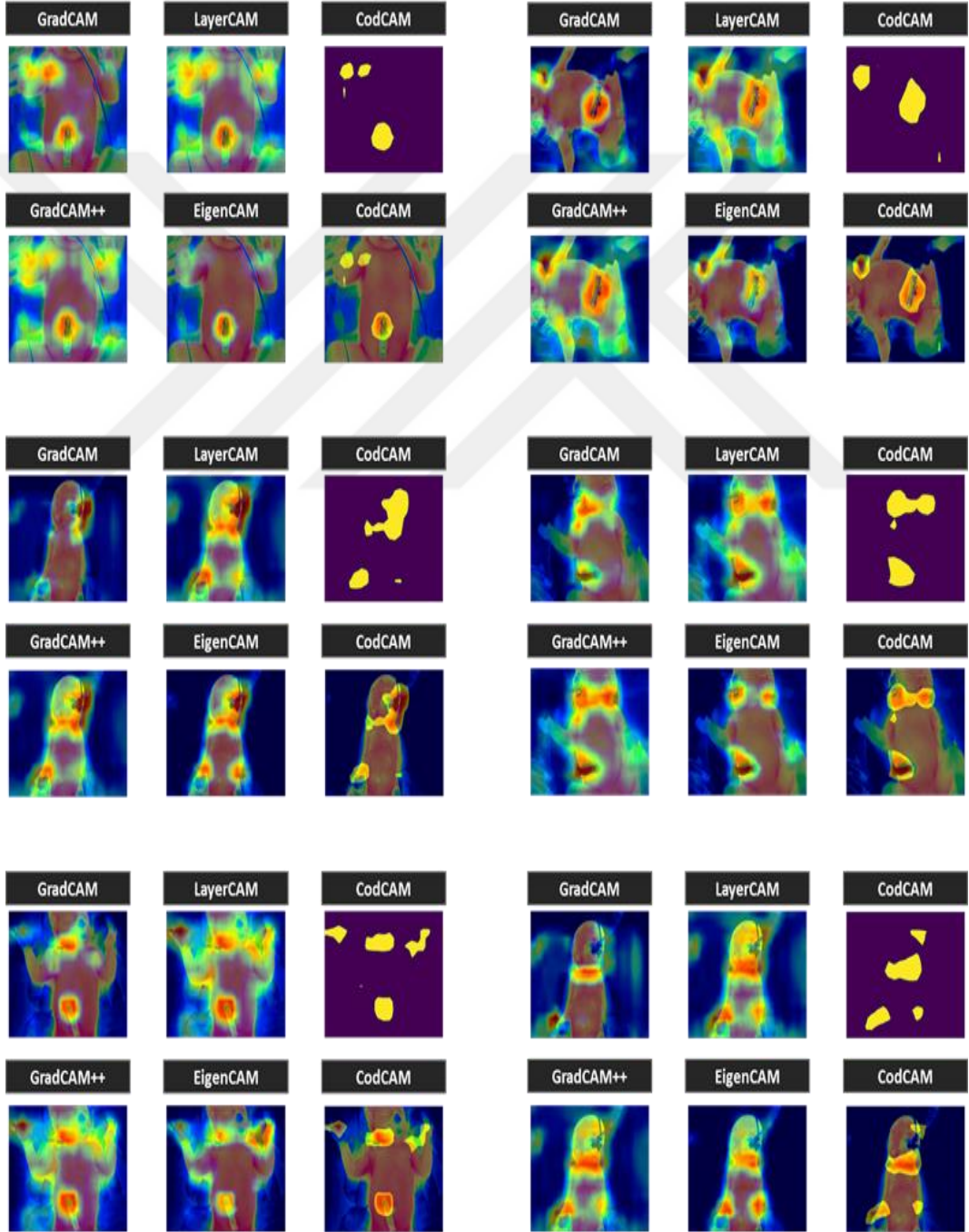
Şekil 4.4’ten görüldüğü gibi yeni doğan bebeklere ait görüntüler %99.47 duyarlılık, %100 hassasiyet ve %99.91 doğruluk ile öğrenilmiştir. Diğer bir deyişle 2280 görüntüden 6 tanesi yanlış sınıflanmıştır. Geçerleme setine bakıldığında ise %99.47 doğruluk görülmektedir yani eğitim işlemi sırasında bir aşırı öğrenme meydana gelmemiştir. Test işleminde eğitim ve geçerlemede kullanılmayan görüntüler modele verilmektedir. 1140 görüntüden 1138 tanesini doğru sınıflanarak %99.82 doğruluk oranına ulaşılmıştır.

Eğitilen model %99 doğruluğun üzerinde çalışırken yalnızca hasta ve sağlıklı olarak çıktı vermekte fakat bunu nasıl yaptığı AYZ yöntemleri kullanılmadan bilinmemektedir. Doktorların verilen karara olan güvenlerinin artırılması için Şekil 4.5’te görüldüğü gibi görsel AYZ yöntemleri kullanılarak modelin kararını etkileyen kısımlar ortaya çıkarılmıştır.

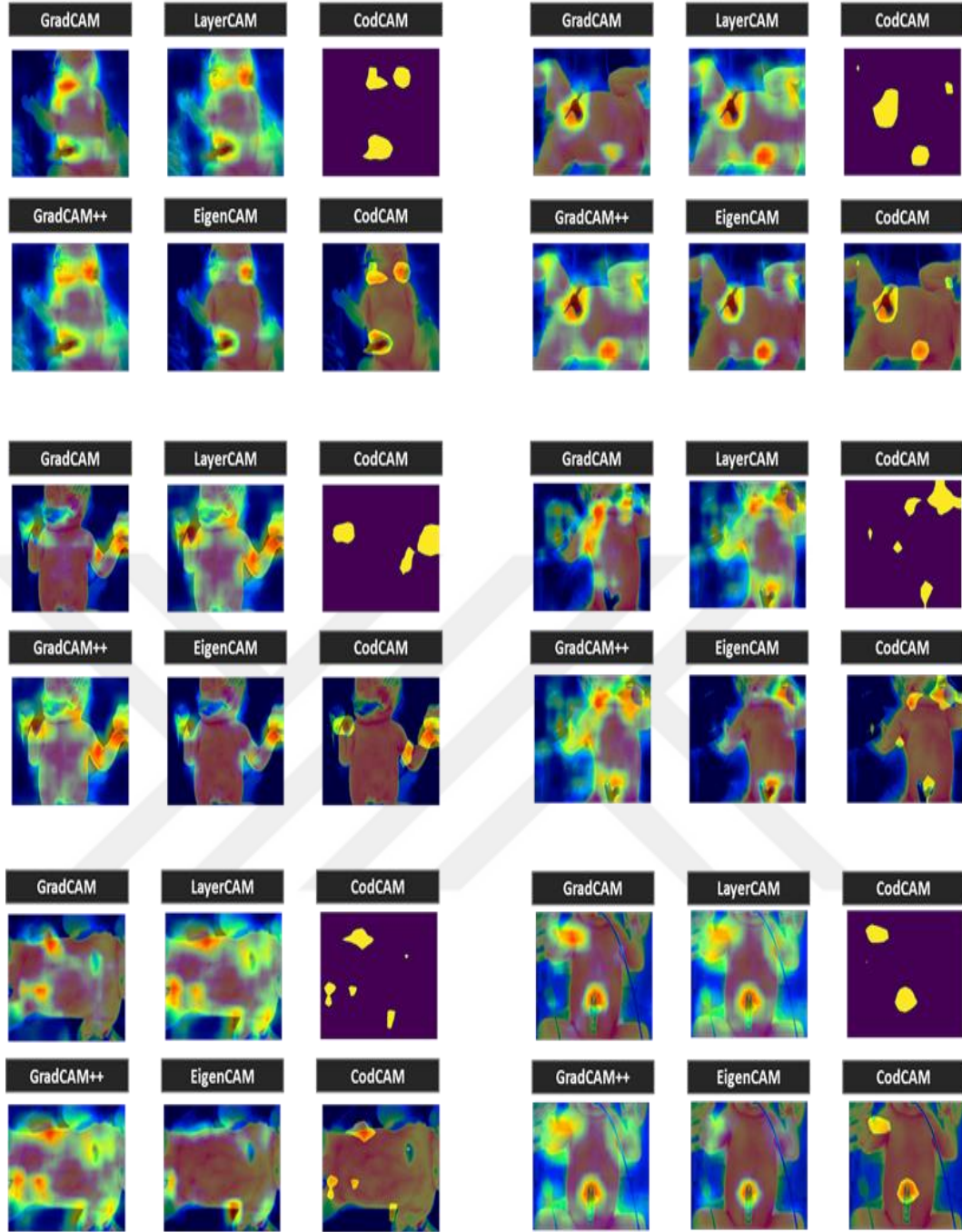


Şekil 4.5. Yapay zeka kararlarının görsel AYZ yöntemleri ile görselleştirilmesi

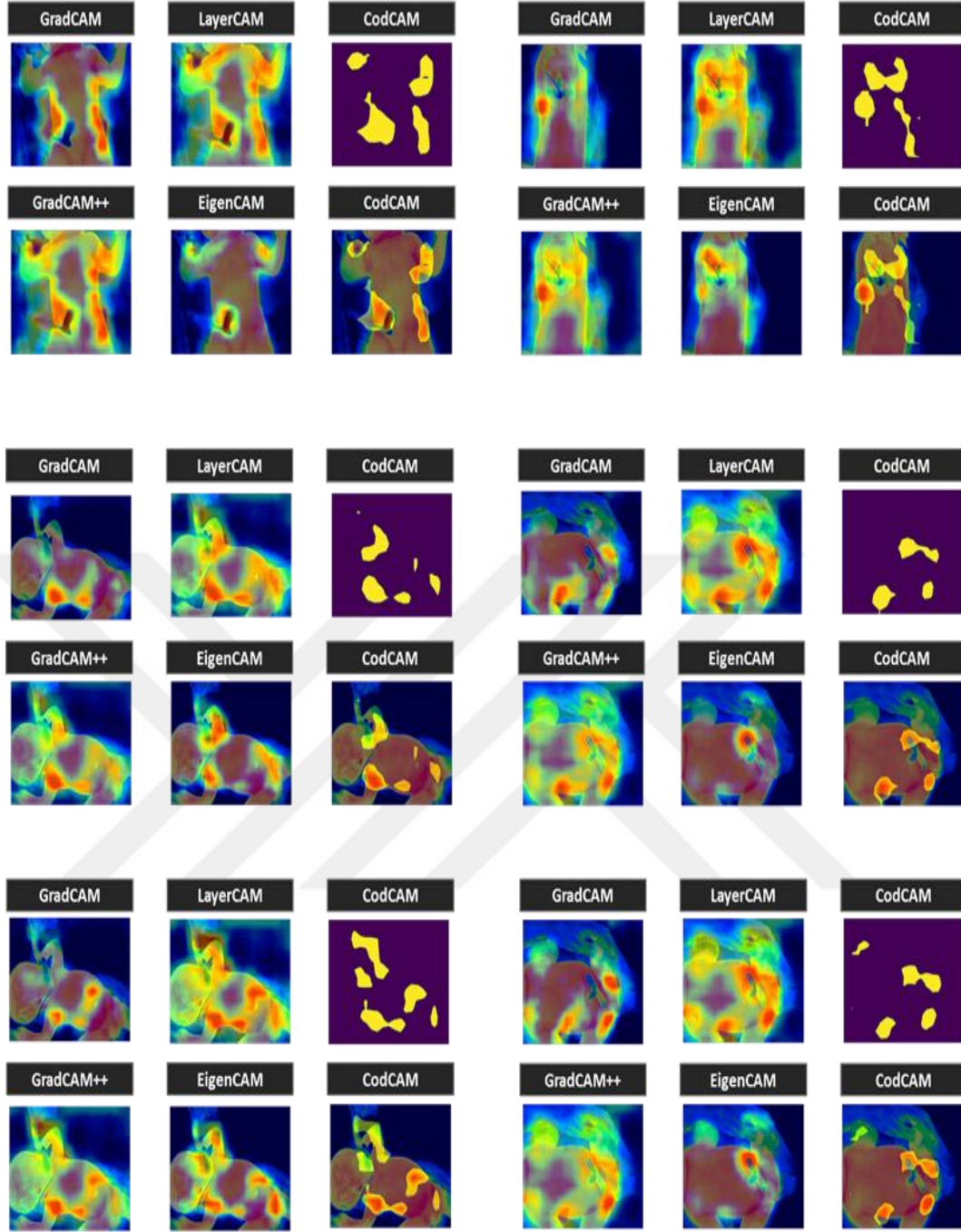
Şekil 4.5'ten görüldüğü gibi her bir AYZ yöntemi bebek üzerinde bir alana odaklanmakta ve bu bölgeler göz önünde bulundurularak kararın verildiği söylenmektedir. Ancak her yöntem hem ortak hem de farklı kısımları ortaya çıkarmaktadır. Her yöntemin ortaya çıkardığı ortak alanları bulmak amacıyla CodCAM yöntemi geliştirilmiş ve böylece bir topluluk kararı verilmesi sağlanmıştır. Elde edilen sonuçlar Şekil 4.6 (Birinci hasta grup), Şekil 4.7 (ikinci hasta grup), Şekil 4.8 (birinci sağlıklı grup) ve Şekil 4.8'de (ikinci sağlıklı grup) görselleştirilmiştir.



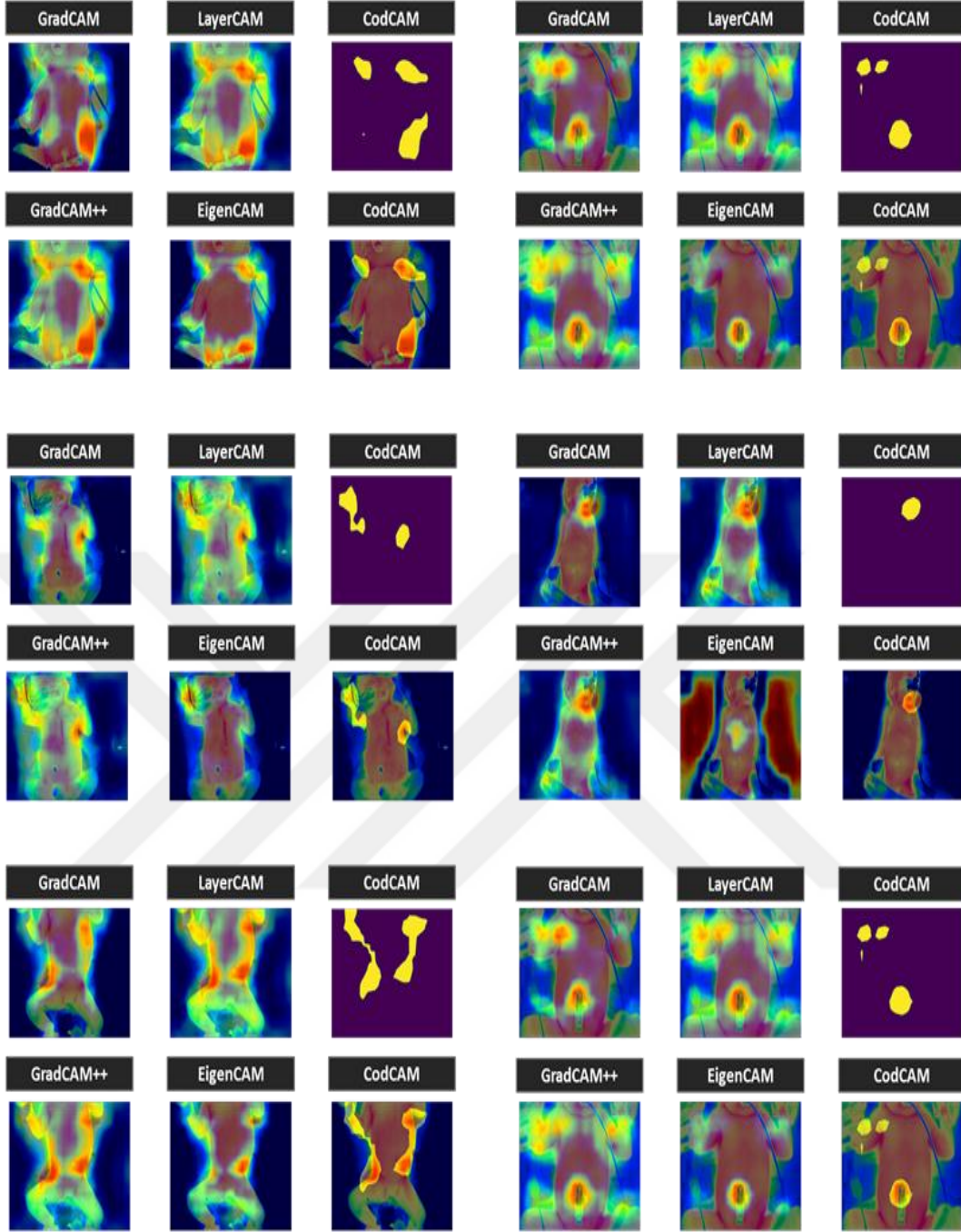
Şekil 4.6. Birinci hasta gruba ait görselleştirmeler



Şekil 4.7. İkinci hasta gruba ait görselleştirmeler



Şekil 4.8. Birinci sağlıklı gruba ait görselleştirmeler



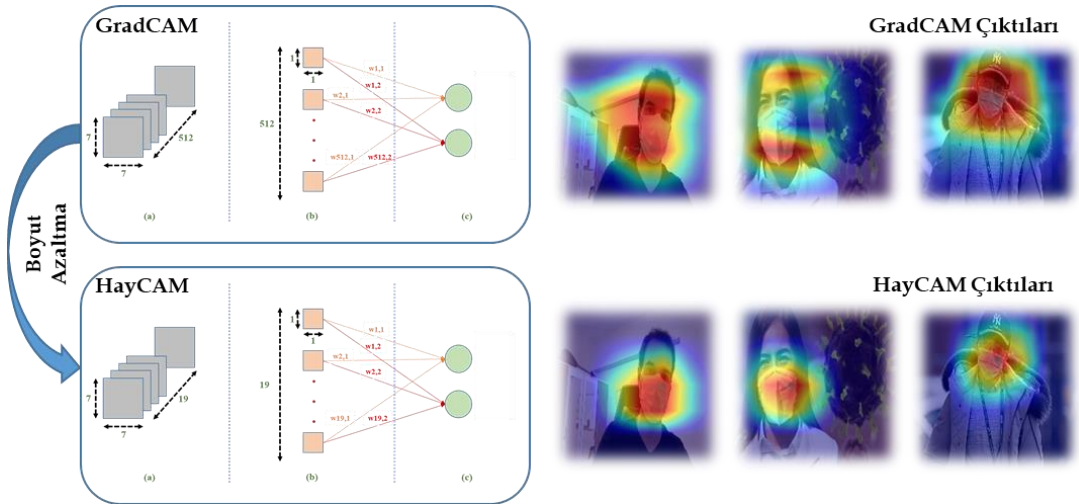
Şekil 4.9. İkinci sağlıklı gruba ait görselleştirmeler

Görselleştirme yapılmadığında model kararları sadece “hasta” ve “sağlıklı” sonuçlarından ibaret olduğundan uzmanlar elde edilen kararların nedenlerini de öğrenmek ve ona göre teşhisi doğrulamak ihtiyacı duymaktadır (Örnek ve Ceylan, 2021). Aynı zamanda yapay zeka tarafında model eğiten ve sonuçları test eden araştırmacılar da modellerin doğru öğrenip öğrenmediğini anlamak için görsel açıklamalara ihtiyaç duymaktadır. CodCAM hem dört farklı yöntemin (GradCAM, GradCAM++, LayerCAM ve EigenCAM) ayrı ayrı çıktılarını hem de ortak kararlarını

görselleştirerek hem uzmanlara da hem de araştırmacılara kararların değerlendirilmesinde önemli bir veri sağlamaktadır. Şekil 4.6, 4.7, 4.8 ve 4.9 incelendiğinde, eğitilen Resnet50 modelinin, görüntülerin arka planını öğrenmek yerine bebeklerin vücuduna ait özellikleri ayırt etmeyi başardığı görülmüştür. Örnekler arasında, eğitilen modelin karar verirken bebeklerin kol, koltuk altı, baş, boyun, ayak ve vücut gibi belirgin anatomik bölgelerine odaklandığı açıkça görülmektedir. Bu sonuçlar, modelin sadece genel görüntüyü değil, aynı zamanda bebeklerin vücut yapılarını tanıdığını ve bu özellikleri kullanarak doğru sınıflandırmalar yaptığını göstermektedir.

4.2. HayCAM Yönteminin Maske Tanıma Problemi Üzerindeki Sonuçları

Görsel olarak açıklanabilir bir yöntem olan SAH, sınıflandırma gradyanlarını kullanarak girdi görüntülerinin önemli bölgelerini vurgulamak için kullanılır. Mevcut yöntemlerin eksikliği, dağınık ve odaklanmamış aktivasyon haritalarına neden olan son evrişim katmanındaki tüm filtrelerin kullanılmasıdır. Yeni bir görselleştirme yöntemi olarak HayCAM, boyut azaltma kullanarak Şekil 4.10'da görüldüğü gibi daha odaklanmış SAH'lar oluşturulmasını sağlamakta ve bu SAH'lar kullanılarak nesne tanıma işlemi de gerçekleştirilmektedir. Önerilen HayCAM yöntemi Şekilde 4.9'da gösterilmiştir.



Şekil 4.10. Önerilen HayCAM Yönteminin SAH'lar Üzerindeki Etkisi

Maske tespiti çalışmasında 18400 eğitim, 200 doğrulama ve 4540 test olmak üzere toplamda 23140 görüntü kullanılmıştır. ESA modeli olarak sınıflama performansı yüksek ön eğitilmiş bir model olan Resnet18 seçilmiş ve 18400 eğitim görüntüsü ile

model eğitilmiştir. Maske tespiti problemi için kullanılan ortam ve veri bilgileri Çizelge 4.3’de gösterilmiştir.

Çizelge 4.3. Çalışmada Kullanılan Ortam ve Veri Bilgileri

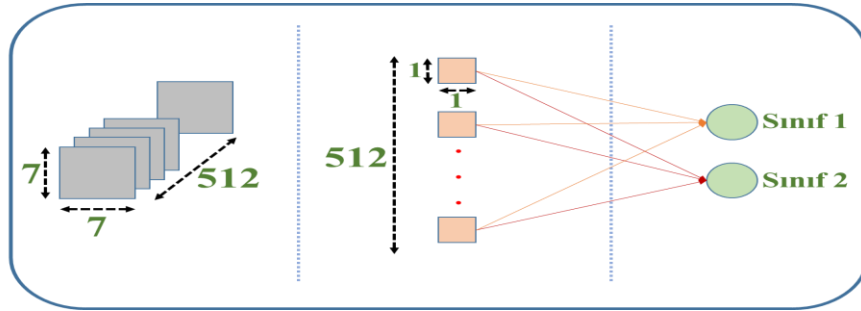
Programlama Dili	Python
Eğitim Ortamı	Linux
Derin Öğrenme Kütüphanesi	PyTorch
GIÜ	Tesla K80
Maskeli Görüntüler	11570
Maskesiz Görüntüler	11570

Eğitim sırasında aşırı öğrenme olup olmadığının kontrol edilmesi için 200 adet geçerleme görüntüsü, modelin test edilmesi için ise 4540 test görüntüsü kullanılmıştır. Model sınıflama sonuçları Çizelge 4.4’te gösterilmiştir.

Çizelge 4.4. Maske Sınıflama Sonuçları

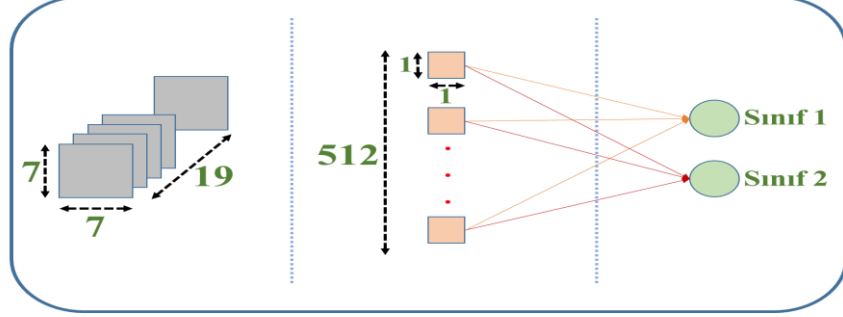
Hassasiyet	%95.16
Özgüllük	%96.69
Doğruluk	%96.58

Eğitilen model 4290 maskeli görüntünün 4090’ını, 310 maskesiz görüntünün 295’ini doğru olarak sınıflamıştır. Sınıflama işlemi tamamlandıktan sonra görselleştirme gerçekleştirilmektedir. Şekil 4.11’de görüldüğü gibi Resnet18’in son katmanında 512 evrişim bulunmaktadır.



Şekil 4.11. Resnet18 Modelinin Son 3 Katmanı

HayCAM, TBA kullanarak son katmanda bulunan 512 filtrenin 19 filtreye indirilmesini sağlamaktadır. 19 filtreli model Şekil 4.12’de gösterilmiştir.

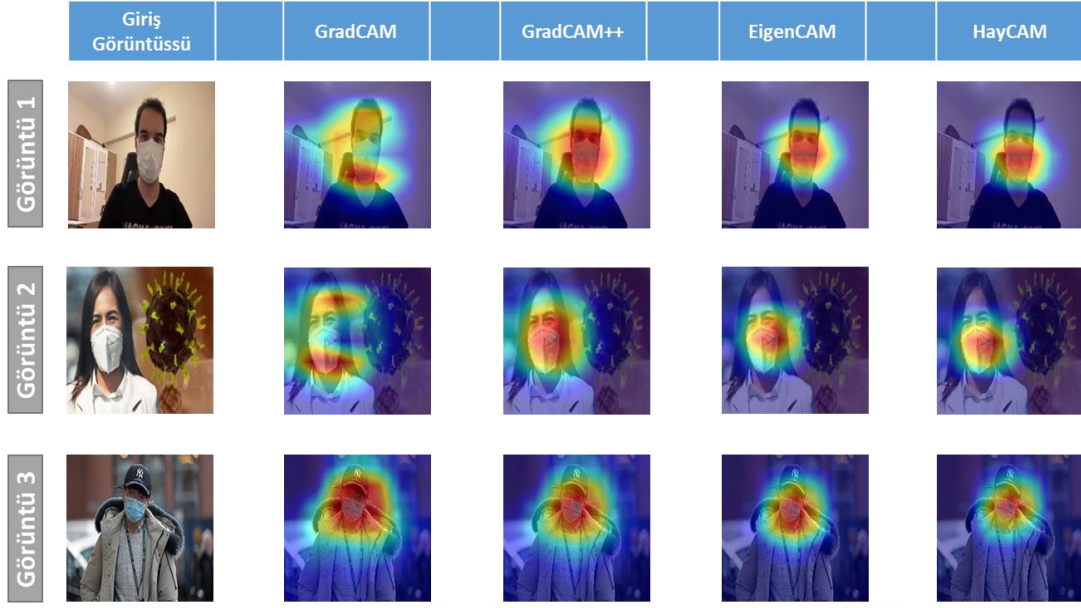


Şekil 4.12. 19 Filtreye Sahip Resnet18 Modeli

Şekil 4.12’de sunulan yapı kullanıldığında HayCAM, Şekil 4.13’de görüldüğü gibi diğer yöntemlere göre daha odaklanmış SAH’lar oluşturmaktadır. Oluşturulan SAH’ların giriş görüntülerine uygulanmasıyla elde edilen sonuçlar Şekil 4.14’te gösterilmiştir.

	Giriş Görüntüsü	GradCAM	GradCAM++	EigenCAM	HayCAM
Görüntü 1					
Görüntü 2					
Görüntü 3					

Şekil 4.13. HayCAM ile Oluşturulan SAH'lar



Şekil 4.14. SAH'ların Giriş Görüntülerine Uygulanması

Şekil 4.14'ten görüldüğü üzere oluşturulan SAH'lar arasında maske üzerine en çok odaklanan yöntem önerilen HayCAM yöntemidir. SAH'ların çevresine çevreleyici kutular çizilerek elde edilen nesne tanıma sonuçları ise Şekil 4.15'te gösterilmiştir. Nesne tanıma performansını ölçmek için IoU metriği kullanılmıştır. Elde edilen IoU değerleri Çizelge 4.5'te gösterilmiştir.

	Giriş Görüntüsü	GradCAM	GradCAM++	EigenCAM	HayCAM	Beklenen Çıktı
Görüntü 1						
	Image 1, IOU:	0.1721	0.2438	0.3246	0.3627	
Görüntü 2						
	Image 2, IOU:	0.2128	0.3070	0.3666	0.4822	
Görüntü 3						
	Image 3, IOU:	0.1116	0.1463	0.1959	0.2436	

Şekil 4.15. Nesne Tanıma Uygulaması Görsel Sonuçlar

Çizelge 4.5. HayCAM ile Maske Tespiti Uygulaması için IoU Değerleri

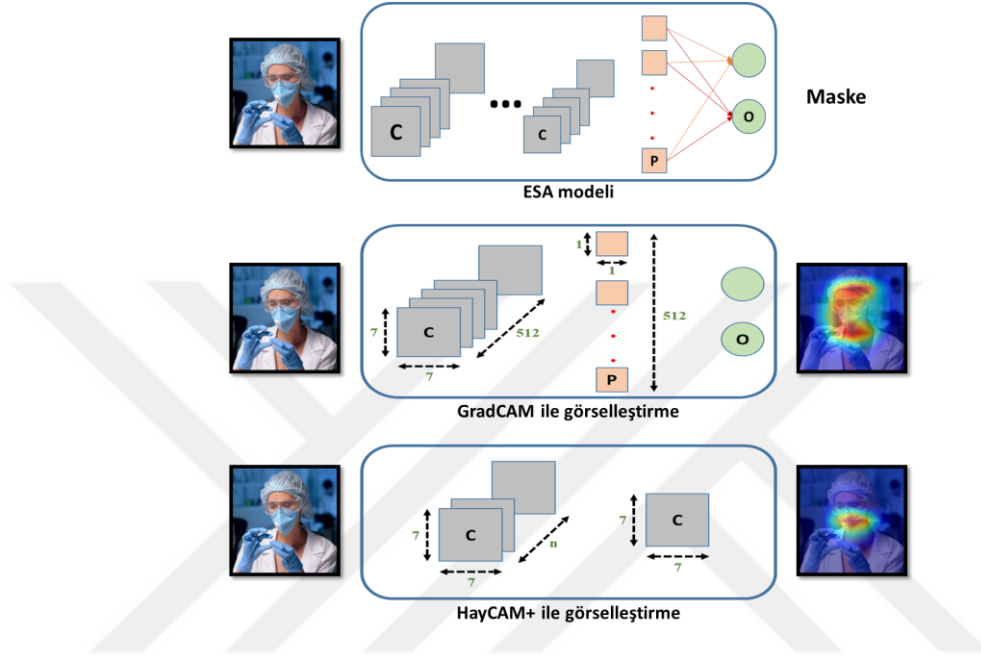
Görüntü No	GradCAM	GradCAM++	EigenCAM	HayCAM
1	0.1721	0.2438	0.3246	0.3627
2	0.2128	0.3070	0.3666	0.4822
3	0.1116	0.1463	0.1959	0.2436
4	0.2612	0.2669	0.4425	0.5269
5	0.2291	0.3415	0.4298	0.4420
6	0.0523	0.1214	0.1784	0.1963
7	0.0520	0.1148	0.1701	0.1766
8	0.0499	0.1188	0.1647	0.1829
9	0.1077	0.1651	0.2316	0.2529
10	0.0953	0.1664	0.2404	0.2553
11	0.2501	0.3101	0.4144	0.4483
...	...	...	...	...
151	0.2259	0.2797	0.3407	0.4381
152	0.2217	0.2832	0.3392	0.4219
153	0.1529	0.2438	0.3397	0.4037
154	0.1486	0.1898	0.3104	0.3676
155	0.1359	0.1856	0.2771	0.3466
156	0.1232	0.1695	0.2972	0.3187
157	0.1205	0.1484	0.2018	0.2115
158	0.1613	0.2226	0.3061	0.3223
159	0.1651	0.2341	0.3021	0.3454
160	0.1486	0.1898	0.3104	0.4277
Ortalama	0.1922	0.2472	0.3386	0.3487

Çizelge 4.5'ten görüldüğü gibi IoU değerleri GradCAM için 0,1922; GradCAM++ için 0,2472; EigenCAM için 0,3386 ve önerilen HayCAM için 0,3487 olarak elde edilmiştir. Şekil 4.15'te de görüldüğü gibi, elde edilen sonuçlar HayCAM'in boyut azaltma ile en iyi SAH'ları elde ettiğini göstermektedir.

4.3. HayCAM+ Yönteminin Maske Tanıma Problemi Üzerindeki Sonuçları

HayCAM yöntemi son katmanda bulunan evrişimlere boyut azaltma uygulayarak sadece önemli evrişimlerin ve ağırlıkların kullanılmasını sağlamaktadır. Böylece ortaya çıkarılan aktivasyon haritaları Şekil 4.15'te görüldüğü gibi daha odaklanmış bir şekilde oluşmaktadır. Ancak bu yöntemin en önemli dezavantajlarından birisi boyut azaltmada kullanılan temel filtrelerin sayısının deneysel olarak belirlenmesidir.

HayCAM+ yönteminde ise TBA'da kullanılan filtre sayıları otomatik olarak hesaplanmaktadır ve tam bağlı katman ile ağırlıklandırma işlemi ortadan kaldırılmıştır. Şekil 4.16'da HayCAM+ yönteminin özet bir temsili verilmiştir. Çalışmada kullanılan tüm veriler, çalışma ortamı ve model HayCAM ile aynıdır ancak bu çalışmada ek olarak LayerCAM yöntemi de karşılaştırma çizelgesine eklenmiştir.

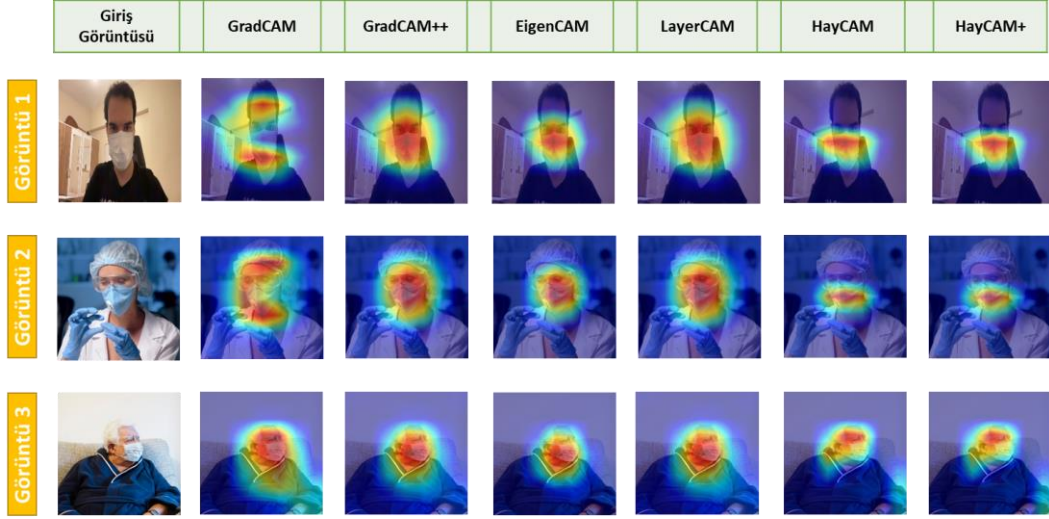


Şekil 4.16. HayCAM+ Yönteminin SAH'lar Üzerindeki Etkisi.

HayCAM+ ile üretilen SAH'lar Şekil 4.17'de sunulmuştur. Şekil 4.18, bu SAH'ların giriş görüntüleriyle birleştirildiği hali göstermektedir.

	Giriş Görüntüsü	GradCAM	GradCAM++	EigenCAM	LayerCAM	HayCAM	HayCAM+
Görüntü 1							
Görüntü 2							
Görüntü 3							

Şekil 4.17. HayCAM+ ile Oluşturulan SAH'lar.



Şekil 4.18. Giriş Görüntüleri ile SAH'ların Birleştirilmesi

Şekil 4.17 ve Şekil 4.18'den görüldüğü gibi HayCAM+ en odaklanmış SAH'ların oluşmasını sağlamaktadır. SAH'ların etrafına çevreleyen kutular çizildiğinde elde edilen nesne tanıma sonuçları ise Şekil 4.19'da görselleştirilmiştir. Nesne tanıma performansının ölçülmesi için IoU ve Dice skorları kullanılmıştır.

	Beklenen Çıktı	GradCAM	GradCAM++	EigenCAM	LayerCAM	HayCAM	Haycam+
Görüntü 1							
	IoU	0.3752	0.2950	0.3558	0.2971	0.4048	0.4443
	Dice	0.5456	0.4556	0.5248	0.4588	0.5764	0.6153
Görüntü 2							
	IoU	0.2679	0.3728	0.5238	0.3694	0.5095	0.5898
	Dice	0.4225	0.5431	0.6875	0.5395	0.6750	0.7420
Görüntü 3							
	IoU	0.0994	0.1620	0.2087	0.1603	0.2173	0.2201
	Dice	0.1809	0.2789	0.3454	0.2764	0.3571	0.3607

Şekil 4.19. HayCAM+ ile Maske Tespiti Uygulaması için Görsel Sonuçlar.

Şekil 4.19'dan görüldüğü gibi GradCAM [3], GradCAM++ [4], EigenCAM [5], LayerCAM [6], HayCAM ve HayCAM+ karşılaştırıldığında en iyi nesne tanıma

performansı HayCAM+ ile elde edilmiştir. Çizelge 4.6 ve Çizelge 4.7’da sırasıyla nesne tanımayaya ait IoU ve Dice skorları verilmiştir.

Çizelge 4.6. Maske Tespiti Uygulaması IoU Değerleri

Görüntü No	GradCAM	GradCAM++	EigenCAM	LayerCAM	HayCAM	HayCAM+
1	0.0523	0.1214	0.1784	0.1214	0.2828	0.3156
2	0.0499	0.1188	0.1647	0.1188	0.2816	0.2895
3	0.0835	0.1488	0.2087	0.1488	0.4269	0.4782
4	0.1162	0.1752	0.2457	0.1752	0.3024	0.3105
5	0.1014	0.1900	0.2685	0.1881	0.2699	0.2706
6	0.1146	0.1345	0.1879	0.1358	0.3337	0.3568
...						
154	0.1285	0.1534	0.2131	0.1519	0.4179	0.4575
155	0.2039	0.2555	0.3397	0.2555	0.3460	0.3948
156	0.2275	0.2803	0.3643	0.2830	0.3735	0.4205
157	0.2329	0.2866	0.3593	0.2893	0.4017	0.4231
158	0.1205	0.1484	0.2018	0.1469	0.2845	0.2905
159	0.1293	0.1550	0.2129	0.1534	0.3199	0.4132
160	0.1486	0.1898	0.3104	0.1916	0.3468	0.4039
Ortalama	0.1922	0.2472	0.3386	0.2476	0.3487	0.3740

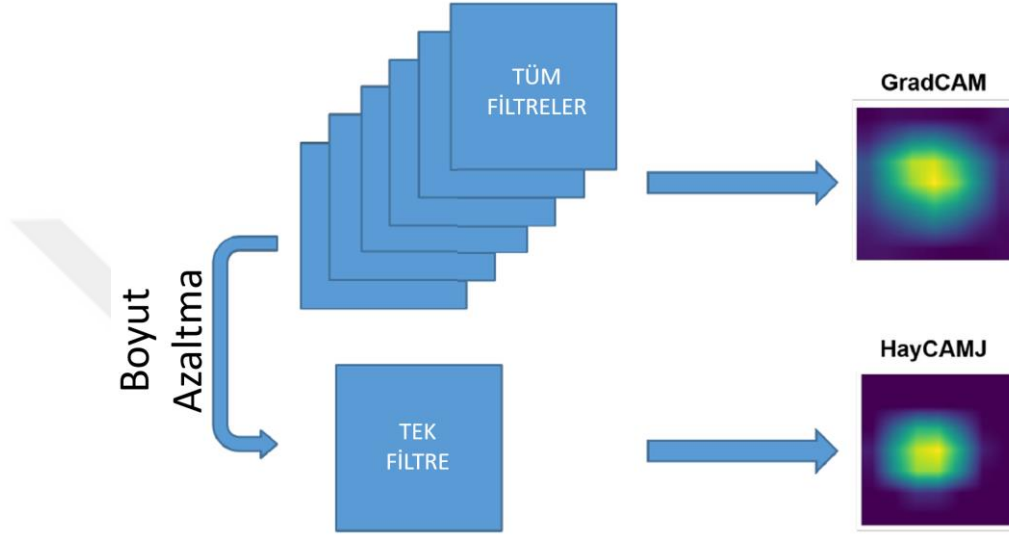
Çizelge 4.7. Nesne Tespiti Uygulaması Dice Değerleri

Görüntü No	GradCAM	GradCAM++	EigenCAM	LayerCAM	HayCAM	HayCAM+
1	0.0994	0.2165	0.3029	0.2165	0.4409	0.4798
2	0.0951	0.2123	0.2828	0.2123	0.4395	0.4490
3	0.1542	0.2148	0.2941	0.2129	0.2924	0.2903
4	0.2082	0.2983	0.3945	0.2983	0.4644	0.4738
5	0.1842	0.3193	0.4233	0.3166	0.4251	0.4260
6	0.2056	0.2371	0.3164	0.2391	0.5004	0.5260
...						
154	0.2278	0.2660	0.3513	0.2637	0.5895	0.6278
155	0.3388	0.4070	0.5072	0.4070	0.5141	0.5661
156	0.3707	0.4379	0.5341	0.4411	0.5438	0.5921
157	0.3778	0.4455	0.5286	0.4487	0.5731	0.5946
158	0.2151	0.2584	0.3358	0.2562	0.4430	0.4502
159	0.2290	0.2683	0.3510	0.2660	0.4847	0.5847
160	0.2588	0.3190	0.4738	0.3215	0.5150	0.5754
Ortalama	0.3153	0.3923	0.5003	0.3928	0.5098	0.5376

Çizelge 4.6 ve Çizelge 4.7’den görüldüğü üzere sırasıyla en yüksek IoU ve Dice değerleri %37.40 ve %53.76 olarak HayCAM+ ile elde edilmiştir. Böylece HayCAM yöntemine göre yaklaşık %2.5’lik bir iyileştirme gerçekleştirilmiştir.

4.4. HayCAMJ Yönteminin Pascal VOC 2007 ve 2012 Üzerine Sonuçları

HayCAMJ, ESA'nın son katmanında bulunan evrişimlere TBA ile boyut azaltma uygulayıp öz değeri en yüksek olan bileşenin alınmasını sağlamakta ve nesne tanıma işlemi gerçekleştirmektedir. Şekil 4.20'de örnek bir, tüm filtreleri kullanan GradCAM çıktısı ve tek filtre kullanan HayCAMJ çıktısı gösterilmiştir.



Şekil 4.20. HayCAMJ ile TBA kullanılarak filtrelerin tek bir filtreye indirilmesi

HayCAMJ yönteminin etkinliğinin gösterilmesi için literatürde sıklıkla geçen Pascal VOC 2007 ve Pascal VOC 2012 veri setleri kullanılmıştır. Pascal VOC 2007 veri seti toplamda 9963 görüntü ve 24640 etiketli nesne içermektedir. Pascal VOC 2012 veri seti ise 11530 görüntü ve 27450 etiketli nesne içermektedir. Veri setleri içerisinde rastgele seçilen görüntüler Şekil 4.21'de gösterilmiştir.



Şekil 4.21. Pascal VOC 2007 ve 2012 Veri Setleri Arasından Rastgele Seçilen Görüntüler

Pascal VOC veri setlerinde bazı nesneler görüntünün tümünü kaplayacak boyuta ulaşmakta, yani nesne/görüntü oranı yüksek olabilmektedir. Ancak HayCAM, HayCAM+ ve HayCAMJ yöntemlerinin temel amacı daha odaklanmış aktivasyon haritaları oluşturmak olduğu için nesne/görüntü oranı büyük olan görüntüler üzerinde nesne tanıma performansı azalmaktadır. Nesne/görüntü oranlarına göre IoU performansı Çizelge 4.8’de Resnet18, Çizelge 4.9’da Resnet34, Çizelge 4.10’da Resnet50 ve Çizelge 4.11’de Mobilenet için verilmiştir.

Çizelge 4.8. Resnet18 için Çeşitli nesne/görüntü oranlarına göre IoU değerleri

nesne/görüntü oranı <	GradCAM	GradCAM++	EigenCAM	HayCAMJ
0.95	0.4764	0.4838	0.4545	0.4452
0.90	0.4825	0.4778	0.4511	0.4461
0.85	0.4753	0.4783	0.4436	0.4468
0.80	0.4713	0.4738	0.4353	0.4775
0.75	0.4676	0.4685	0.4269	0.4479
0.70	0.4635	0.4631	0.4192	0.4479
0.65	0.4597	0.4585	0.4129	0.4485
0.60	0.4551	0.4525	0.4044	0.4484
0.55	0.4488	0.4454	0.3949	0.4480
0.50	0.4421	0.4380	0.3854	0.4468
0.45	0.4352	0.4303	0.3763	0.4448
0.40	0.4285	0.4230	0.3678	0.4427
0.35	0.4216	0.4157	0.3597	0.4397
0.30	0.4149	0.4088	0.3523	0.4358
0.25	0.4085	0.4023	0.3457	0.4318

Çizelge 4.9. Resnet34 için Çeşitli nesne/görüntü oranlarına göre IoU değerleri

nesne/görüntü oranı <	GradCAM	GradCAM++	EigenCAM	HayCAMJ
0.95	0.4856	0.4966	0.4388	0.4539
0.90	0.4869	0.4970	0.4363	0.4549
0.85	0.4834	0.4931	0.4276	0.4554
0.80	0.4791	0.4882	0.4190	0.4555
0.75	0.4759	0.4822	0.4102	0.4549
0.70	0.4718	0.4763	0.4023	0.4534
0.65	0.4680	0.4712	0.3958	0.4522
0.60	0.4628	0.4647	0.3875	0.4510
0.55	0.4568	0.4569	0.3785	0.4491
0.50	0.4506	0.4493	0.3696	0.4471
0.45	0.4440	0.4415	0.3610	0.4440
0.40	0.4378	0.4342	0.3530	0.4408
0.35	0.4315	0.4270	0.3454	0.4370
0.30	0.4251	0.4201	0.3384	0.4327
0.25	0.4195	0.4136	0.3320	0.4284

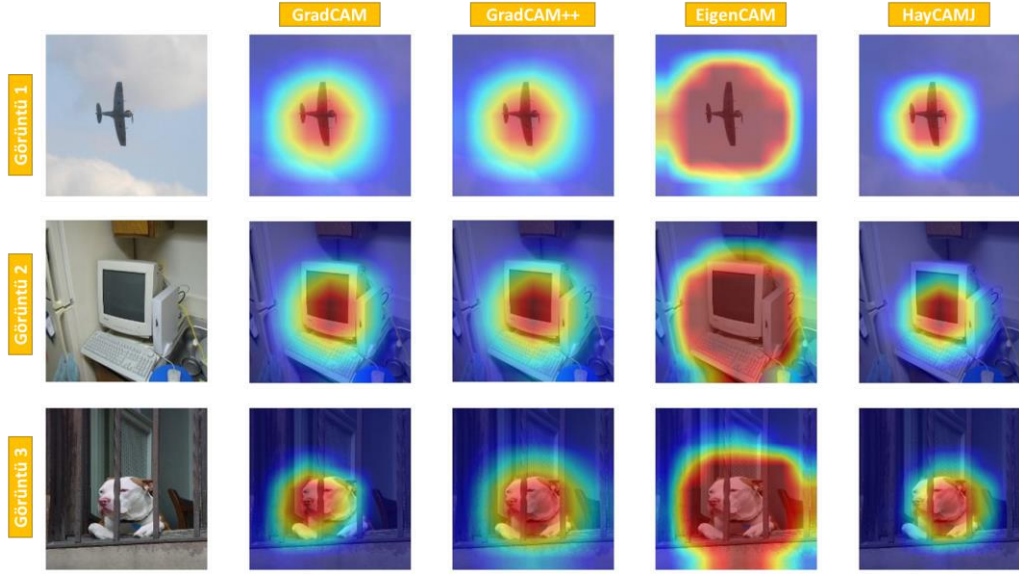
Çizelge 4.10. Resnet50 için Çeşitli nesne/görüntü oranlarına göre IoU değerleri

nesne/görüntü oranı <	GradCAM	GradCAM++	EigenCAM	HayCAMJ
0.95	0.4901	0.5007	0.4214	0.4597
0.90	0.4898	0.4987	0.4178	0.4596
0.85	0.4877	0.4951	0.4097	0.4605
0.80	0.4835	0.4896	0.4008	0.4604
0.75	0.4792	0.4830	0.3916	0.4588
0.70	0.4746	0.4764	0.3834	0.4571
0.65	0.4707	0.4710	0.3766	0.4562
0.60	0.4653	0.4637	0.3680	0.4534
0.55	0.4586	0.4555	0.3586	0.4497
0.50	0.4519	0.4474	0.3494	0.4457
0.45	0.4451	0.4398	0.3407	0.4416
0.40	0.4394	0.4394	0.3325	0.4376
0.35	0.4330	0.4255	0.3249	0.4329
0.30	0.4271	0.4189	0.3179	0.4274
0.25	0.4209	0.4122	0.3116	0.4236

Çizelge 4.11. Mobilenet için Çeşitli nesne/görüntü oranlarına göre IoU değerleri

nesne/görüntü oranı <	GradCAM	GradCAM++	EigenCAM	HayCAMJ
0.95	0.3831	0.4549	0.4441	0.3441
0.90	0.3786	0.4462	0.4309	0.3427
0.85	0.3758	0.4405	0.4214	0.3421
0.80	0.3732	0.4357	0.4142	0.3418
0.75	0.3724	0.4301	0.4062	0.3413
0.70	0.1982	0.4068	0.3566	0.3596
0.65	0.1968	0.3960	0.3434	0.3577
0.60	0.1846	0.3798	0.3250	0.3562
0.55	0.1721	0.3656	0.3087	0.3542
0.50	0.1622	0.3533	0.2952	0.3517
0.45	0.1539	0.3442	0.2838	0.3503
0.40	0.1480	0.3336	0.2743	0.3480
0.35	0.1426	0.3246	0.2654	0.3438
0.30	0.1374	0.3158	0.2574	0.3381
0.25	0.1321	0.3081	0.2499	0.3331

HayCAMJ yöntemi Resnet18 modeli üzerinde nesne/görüntü oranı %50'nin (Çizelge 4.8), Resnet34 modeli üzerinde nesne/görüntü oranı %40'ın (Çizelge 4.9), Resnet50 modeli üzerinde nesne/görüntü oranı %30'un (Çizelge 4.10) ve Mobilenet modeli üzerinde nesne/görüntü oranı %45'in (Çizelge 4.11) altına düştüğünde daha iyi performans göstermektedir. HayCAMJ yöntemi, Pascal VOC 2007 ve 2012 veri setlerinde nesne/görüntü oranı %30'un altında olan görüntülere uygulandığında Şekil 4.22'deki gibi aktivasyon haritaları meydana gelmektedir.



Şekil 4.22. Görsel AYZ yöntemleri ile elde edilen aktivasyon haritaları

Şekil 4.22’den görüldüğü gibi HayCAMJ, diğer yöntemlerle karşılaştırıldığında nesnelerin üzerinde odaklanmış SAH’lar oluşturulmasını sağlamaktadır. Pascal VOC veri setlerinin sınıflayıcı bir model tarafından öğrenilmesi için sınıflama performansı yüksek Resnet18, Resnet34, Resnet50 ve Mobilenet modelleri seçilmiştir. Modeller 30 kez eğilerek Çizelge 4.12 görülen doğruluk değerlerine ulaşmıştır.

Çizelge 4.12. HayCAMJ Nesne Tanıma IoU Değerleri Pascal VOC 2007

Veri	Resnet18	Resnet34	Resnet50	Mobilenet
Eğitim	0.9930	0.9950	0.9960	0.9116
Geçerleme	0.9050	0.9140	0.9250	0.6899
Test	0.9160	0.9301	0.9380	0.7091

Eğitilen Resnet18, Resnet34, Resnet50 ve Mobilenet sınıflayıcı modellerinin Pascal VOC 2007 ve 2012 veri setleri üzerindeki nesne tanıma performansları Çizelge 4.13’de ve Çizelge 4.14’te sırasıyla gösterilmiştir.

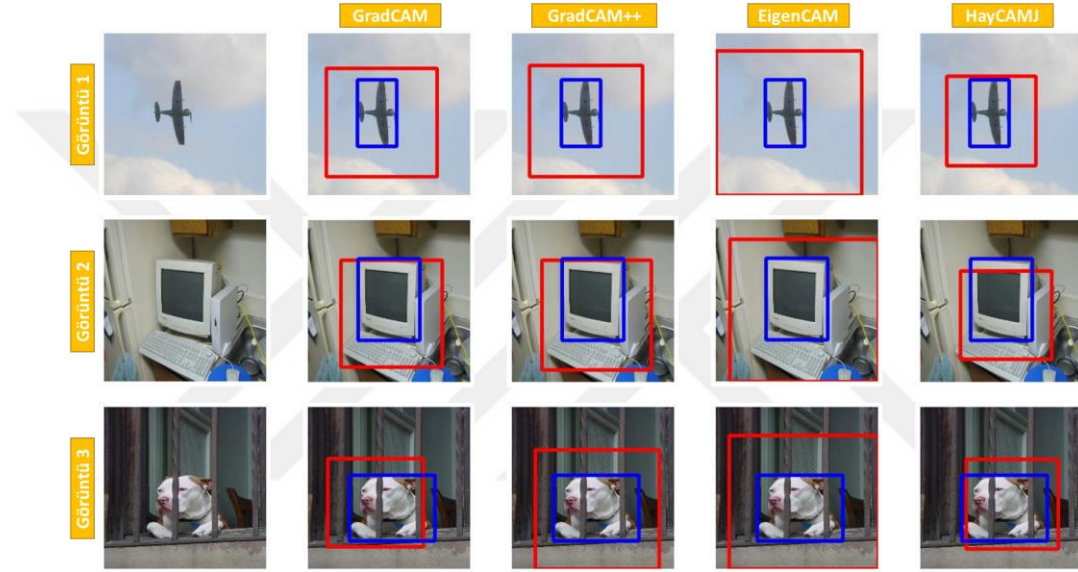
Çizelge 4.13. HayCAMJ Nesne Tanıma IoU Değerleri Pascal VOC 2007

Methods	Resnet18	Resnet34	Resnet50	Mobilenet
GradCAM	0.2738	0.2548	0.2771	0.1533
GradCAM++	0.2377	0.2296	0.2494	0.1503
EigenCAM	0.1559	0.1482	0.1278	0.1275
HayCAMJ	0.3277	0.3224	0.3107	0.1605

Çizelge 4.14. HayCAMJ Nesne Tanıma IoU Değerleri Pascal VOC 2012

Methods	Resnet18	Resnet34	Resnet50	Mobilenet
GradCAM	0.2452	0.2424	0.2639	0.1405
GradCAM++	0.2231	0.2212	0.2383	0.1475
EigenCAM	0.1541	0.1463	0.1262	0.1260
HayCAMJ	0.3101	0.3118	0.2902	0.1486

Çizelge 4.13 ve Çizelge 4.14'ten görüldüğü gibi HayCAMJ diğer SAH yöntemlerinden daha yüksek performans değeri sağlamaktadır. Bu SAH'lar kullanılarak gerçekleştirilen nesne tanıma işlemleri ise Şekil 4.23'te gösterilmiştir.



Şekil 4.23. HayCAMJ ile Nesne Tanıma.

4.5. Geliştirilen Yöntemlerin Pascal VOC 2007 ve 2012 Üzerinde Karşılaştırılması

Gerçekleştirilen tez çalışması ile birlikte dört farklı görsel AYZ yöntemi geliştirilmiştir. Bu yöntemler CodCAM, HayCAM, HayCAM+ ve HayCAMJ yöntemleridir. Bu yöntemlere ait özet karşılaştırma Çizelge 4.15'te verilmiştir.

Çizelge 4.15. Geliştirilen Yöntemlerin Karşılaştırılması

Yöntem	Veriseti	Model	Filtre Sayısı
CodCAM	1. Yeni doğan bebeklere ait termal görüntüler 2. Pascal VOC 2007 3. Pascal VOC 2012	1. Resnet18 2. Resnet34 3. Resnet50 4. Mobilenet	Tüm filtreler
HayCAM	1. Maske takan ve takmayan insanlara ait görüntüler 2. Pascal VOC 2007 3. Pascal VOC 2012	1. Resnet18 2. Resnet34 3. Resnet50 4. Mobilenet	19
HayCAM+	1. Maske takan ve takmayan insanlara ait görüntüler 2. Pascal VOC 2007 3. Pascal VOC 2012	1. Resnet18 2. Resnet34 3. Resnet50 4. Mobilenet	Otomatik
HayCAMJ	1. Pascal VOC 2007 2. Pascal VOC 2012	1. Resnet18 2. Resnet34 3. Resnet50 4. Mobilenet	1

HayCAM+ yönteminde kullanılacak filtre sayısının belirlenmesi için Pascal VOC veri setleri üzerinde çeşitli varyans eşik değerleri için IoU performansı hesaplanmıştır. Çizelge 4.16’de Resnet18, Resnet34, Resnet50 ve mobilenet modelleri için sonuçlar verilmiştir.

Çizelge 4.16. Pascal VOC Üzerinde Çeşitli Varyans Eşik Değerleri İçin HayCAM+ Performansı

varyans eşik değeri	Resnet18	Resnet34	Resnet50	mobilenet
0.95	0.2887	0.3376	0.3279	0.2697
0.90	0.2969	0.3509	0.3278	0.2645
0.85	0.3183	0.3413	0.3411	0.2782
0.80	0.3094	0.3418	0.3533	0.2765
0.75	0.3207	0.3323	0.3496	0.2810
0.70	0.3207	0.3409	0.3559	0.2845
0.65	0.3378	0.3489	0.3599	0.2815
0.60	0.3331	0.3449	0.3558	0.2617
0.55	0.3555	0.3478	0.3651	0.2722
0.50	0.3707	0.3726	0.3688	0.2663
0.45	0.3691	0.3664	0.3797	0.2794
0.40	0.3800	0.3700	0.3842	0.2938
0.35	0.3926	0.3867	0.3972	0.3025
0.30	0.4145	0.4104	0.3940	0.3011
0.25	0.4226	0.4195	0.3937	0.2877

Çizelge 4.16'dan görüldüğü gibi HayCAM+ yöntemi Pascal VOC veri setleri üzerinden en iyi %25 ile %35 varyans eşik değerleri arasında çalışmaktadır. Tez kapsamında geliştirilen yöntemler Pascal VOC 2007 ve 2012 veri setleri üzerinde test edildiğinde ortaya çıkan sonuçlar Çizelge 4.17 ve Çizelge 4.18'de sırasıyla verilmiştir.

Çizelge 4.17. Geliştirilen Yöntemlerin Nesne Tanıma IoU Değerleri Pascal VOC 2007

Methods	Resnet18	Resnet34	Resnet50	Mobilenet
HayCAM	0.2309	0.2422	0.2230	0.1773
HayCAM+	0.3303	0.2421	0.3097	0.1737
HayCAMJ	0.3277	0.3224	0.3107	0.1605
CodCAM	0.4048	0.4079	0.3548	0.2132

Çizelge 4.18. Geliştirilen Yöntemlerin Nesne Tanıma IoU Değerleri Pascal VOC 2012

Methods	Resnet18	Resnet34	Resnet50	Mobilenet
HayCAM	0.2170	0.2484	0.2283	0.1472
HayCAM+	0.3101	0.2484	0.2941	0.1532
HayCAMJ	0.3101	0.3118	0.2902	0.1486
CodCAM	0.3519	0.3674	0.3322	0.1960

Çizelge 4.17'dan görüldüğü gibi HayCAM, Resnet18 ve Resnet34 için benzer sonuçlar üretirken, Resnet50 ve Mobilenet için daha düşük bir performansa sahiptir. HayCAM+, Resnet18 ve Resnet50 için daha yüksek bir başarı elde ederken, Resnet34 ve Mobilenet için benzer sonuçlar verir. HayCAMJ, Resnet18 ve Resnet34 için benzer bir performansa sahiptir, ancak Resnet50 ve Mobilenet için daha düşük sonuçlar sunar.

Burada CodCAM, diğer tüm yöntemlere göre daha yüksek bir başarı elde eder. CodCAM, özellikle Resnet18 ve Resnet34 için en iyi performansa sahiptir.

Çizelge 4.18'e gelindiğinde HayCAM yöntemi, tüm modellerde benzer bir performans gösterirken, Resnet34 ile en yüksek başarıyı elde eder. Resnet50 ve Mobilenet için daha düşük performansa sahiptir. HayCAM+ yöntemi, Resnet18 ve Resnet50 ile daha iyi sonuçlar elde ederken, Resnet34 ve Mobilenet için benzer sonuçlar verir. HayCAMJ, Resnet18 ve Resnet34 ile benzer bir başarıya sahiptir, ancak Resnet50 ve Mobilenet için daha düşük bir performansa sahiptir. CodCAM yöntemi, diğer tüm yöntemlere göre daha yüksek bir başarı elde etmiştir. Özellikle Resnet34 ve Resnet50 ile en yüksek başarıyı sağlar. Burada da CodCAM farklı yöntemlerin çıktılarını alarak ortak bir karar verilmesini sağladığından en iyi performans değerine sahiptir.

Zayıf danışmanlı nesne tespitinde bir nesneye ait dört koordinatın etiketlenmesi yerine, sadece ilgili resimde o nesnenin olup olmadığının söylenmesi nedeniyle IoU değerleri çok yüksek değerlere ulaşamamaktadır. Ancak bölümlerde gösterilen şekillere bakıldığında IoU değerleri düşük olsa bile ilgili nesnenin sınıflayıcı model tarafından bulunduğu görülmektedir.

5. SONUÇLAR VE ÖNERİLER

Bu bölümde, tez çalışması fikrinin ortaya çıkması, önerilen yöntemlerin gelişim süreci, elde edilen sonuçların önemi ve öneriler verilmiştir.

5.1 Sonuçlar

Öğrenebilen, diğer bir ifade ile veriden anlam çıkarabilen yapay zeka modelleri insan hayatının kolaylaştırılmasında büyük bir rol oynamaya başlamıştır. Günümüzde internet sitelerinde bir şey alırken önerilen ürün yapay zeka ile, gidilecek yol navigasyon ile, stok sayıları tahminleyici modeller ile ve hastalık tespiti sınıflayıcı modeller ile gerçekleştirilmektedir. Tüm bu gelişmelere karşın bir yapay zeka modelinin nasıl karar verdiği tam olarak açıklanamamaktadır. Örneğin bir görüntü yapay zeka tarafından hasta veya sağlıklı olarak sınıflandığı takdirde doktor neden böyle bir kararın verildiğini de anlamak istemektedir. Ancak yapay zeka modellerinin son kararı verebilecek hale gelmesi için daha fazla gelişmeye ihtiyaç vardır. Bu gelişmelerden bazıları şunlardır: veri kalitesi ve miktarı, daha iyi eğitim algoritmaları, duyarlılık ve etik, hukuki çerçeve ve daha iyi açıklanabilirlik.

Yapay zeka yaklaşımlarından biri olan derin öğrenme modelleri, içerdikleri milyonlarca parametre ve doğrusal olmayan fonksiyonlar sayesinde görüntü gibi verileri yüksek doğrulukla sınıflayabilmektedir. Ancak bu mimariler çok fazla parametre içerdiklerinden dolayı açıklanamamakta ve “Kara Kutu” modeller olarak adlandırılmaktadır. Gerçekleştirilen tez çalışmasında derin öğrenme modellerinden biri olan ESA’yı açıklanabilir hale getirmek için CodCAM, HayCAM, HayCAM+ ve HayCAMJ yöntemleri geliştirilmiştir.

CodCAM yöntemi, GradCAM, GradCAM++, LayerCAM ve EigenCAM yöntemlerinin aktivasyon haritalarını analiz ederek ortak bir karar verme yaklaşımına dayanmaktadır. Medikal termal görüntüler kullanılarak sınıflama işlemi gerçekleştirildiğinde, eğitilen modelin %95 üzerinde doğrulukla sınıflama gerçekleştirdiği ve bebeklerin kol, koltuk altı, baş, boyun, ayak ve vücut gibi belirgin anatomik bölgelerine odaklandığı gösterilmiştir. Bu sonuçlar, modelin sadece genel görüntüyü değil, aynı zamanda bebeklerin vücut yapılarını tanıdığını ve bu anatomik özellikleri kullanarak doğru sınıflandırmalar yaptığını göstermektedir.

Pascal VOC 2007 ve 2012 veri setleri üzerinde yapılan testler, CodCAM yönteminin diğer modellere kıyasla en yüksek performansı gösterdiğini ortaya koymuştur. Bu sonuçlar, bir yöntemin farklı veri setleri üzerinde başarılı bir şekilde performans gösterdiğinde, bu yöntemin medikal termal görüntüler üzerinde de etkili olabileceği ve görsel sonuçlara güvenilmesi gerektiği yönünde bir çıkarımı desteklemektedir.

CodCAM, SAH üretirken dört farklı yöntem kullandığından işlem maliyeti yüksek bir yöntemdir. GradCAM yöntemlerine bakıldığında, son evrişim katmanında bulunan tüm filtrelerin hesaplama katıldığı görülmektedir. Ancak ESA mimarisi boyunca var olan tüm evrişimler aynı derecede öneme sahip değildir. Bu nedenle evrişim filtrelerinin hepsini kullanmak yerine, seçilen filtrelerin kullanılması için HayCAM, HayCAM+ ve HayCAMJ yöntemleri geliştirilmiştir.

HayCAM, son evrişim katmanındaki filtre sayısını TBA ile sınırlı bir sayıya indiren bir AYZ yöntemidir. Maske tespit etme üzerine elde edilen sonuçlar, IoU değerlerinin; GradCAM için 0,1922; GradCAM++ için 0,2472; EigenCAM için 0,3386 ve HayCAM için 0,3487 olduğunu göstermiştir. Kullanılacak filtre sayısının otomatik olarak belirlenmesi üzerine gidildiğinde, varyans değerleri kullanılarak filtre sayısının belirlenebileceği bulunmuş ve HayCAM+ geliştirilmiştir. HayCAM+ yöntemi, HayCAM'de kullanılan aynı model ve aynı veri seti üzerinde uygulandığında IoU değerini %2.5 artırarak 0,3740'a ulaşmıştır.

HayCAMJ yöntemi HayCAM ve HayCAM+ yöntemlerinden farklı olarak sadece tek bir filtre kullanma üzerine geliştirilmiştir. HayCAMJ yönteminin değerlendirilmesi için Pascal VOC 2007 ve 2012 veri setleri seçilmiştir. Ancak modeller eğitildikten sonra nesne tanıma işlemi gerçekleştirilirken HayCAMJ'e ait IoU değerlerinin diğer yöntemlerden çok düşük olduğu fark edilmiştir. Durum incelendiğinde; Pascal VOC 2007 ve 2012 veri setleri içerisinde nesne/görüntü oranı %80'in bile üzerinde olan görüntüler bulunduğu fark edilmiştir. Bu nedenle ilgili görüntü sınıflayıcı modele girmeden önce nesne/görüntü oranı kontrol edilerek IoU değerleri hesaplanmıştır. HayCAMJ yöntemi Resnet18 modeli üzerinde nesne/görüntü oranı %50'nin, Resnet34 modeli üzerinde nesne/görüntü oranı %40'ın, Resnet50 modeli üzerinde nesne/görüntü oranı %30'un ve Mobilenet modeli üzerinde nesne/görüntü oranı %45'in altına düştüğünde daha iyi performans göstermiştir.

5.2 Öneriler

Açıklanabilir yapay zeka yaklaşımlarından görsel açıklamaların iki büyük faydası bulunmaktadır: (1) verilen kararların görselleştirilerek anlaşılmasının sağlanması (2) görselleştirilen çıktılar kullanılarak nesne tespiti uygulaması geliştirilmesi.

Sınıflayıcı modelin iyi bir şekilde seçilmesi, ilk adım olarak değerlendirilmelidir. Eğer kullanılan sınıflayıcı eğitim ve test veri seti üzerinde iyi bir şekilde çalışmıyorsa, SAH'ların üretilmesi için kötü bir başlangıç olacaktır. Sınıflayıcı model olarak Resnet modelleri kullanışlı modellerdir çünkü hem aşırı öğrenmeye karşı dayanıklıdırlar hem de derin yapıları sayesinde veriyi iyi bir şekilde öğrenirler.

Model eğitildikten sonra SAH'ların oluşturulması için GradCAM ve EigenCAM gibi herhangi bir yöntem kullanılabilir. Bu tez çalışması ile beraber kullanılacak 4 yeni yöntem daha bulunmaktadır. SAH'lar elde edildikten sonra nesne tanıma yapmak için 0 ile 1 değeri arasında oluşan değerlerin işlenmesi gerekmektedir. Bu adımlar Bölüm 4.1'de anlatıldığı gibi: eşikleme yapmak için OpenCV'nin "inRange" özelliği, eşikleme yaptıktan sonra kenarların bulunması için "findContours" özelliği ve kenarları içeren en iyi kutunun tespit edilmesi için "boundingRect" özelliği kullanılmalıdır. Burada kullanılan parametreler: `inRange(0.3)`, `findContours(retr_tree, chain_approx_simple)`, `boundingRect(edges)` olarak seçilmelidir.

KAYNAKLAR

- Adadi, A. ve Berrada, M., 2018, Peeking inside the black-box: a survey on explainable artificial intelligence (XAI), *IEEE access*, 6, 52138-52160.
- Amari, S.-i., 1993, Backpropagation and stochastic gradient descent method, *Neurocomputing*, 5 (4-5), 185-196.
- Amit, Y., Felzenszwalb, P. ve Girshick, R., 2020, Object detection, *Computer Vision: A Reference Guide*, 1-9.
- Antoniadi, A. M., Du, Y., Guendouz, Y., Wei, L., Mazo, C., Becker, B. A. ve Mooney, C., 2021, Current challenges and future opportunities for XAI in machine learning-based clinical decision support systems: a systematic review, *Applied Sciences*, 11 (11), 5088.
- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D. ve Benjamins, R., 2020, Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI, *Information fusion*, 58, 82-115.
- Bae, W., Noh, J. ve Kim, G., 2020, Rethinking class activation mapping for weakly supervised object localization, *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XV* 16, 618-634.
- Bany Muhammad, M. ve Yeasin, M., 2021, Eigen-CAM: Visual explanations for deep convolutional neural networks, *SN Computer Science*, 2, 1-14.
- Basha, S. S., Dubey, S. R., Pulabaigari, V. ve Mukherjee, S., 2020, Impact of fully connected layers on performance of convolutional neural networks for image classification, *Neurocomputing*, 378, 112-119.
- Cao, L., 2022, Ai in finance: challenges, techniques, and opportunities, *ACM Computing Surveys (CSUR)*, 55 (3), 1-38.
- Castelvecchi, D., 2016, Can we open the black box of AI?, *Nature News*, 538 (7623), 20.
- Chattopadhyay, A., Sarkar, A., Howlader, P. ve Balasubramanian, V. N., 2018, Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks, *2018 IEEE winter conference on applications of computer vision (WACV)*, 839-847.
- Chen, L., Chen, J., Hajimirsadeghi, H. ve Mori, G., 2020, Adapting grad-cam for embedding networks, *proceedings of the IEEE/CVF winter conference on applications of computer vision*, 2794-2803.
- Cheng, G., Yang, J., Gao, D., Guo, L. ve Han, J., 2020, High-quality proposals for weakly supervised object detection, *IEEE Transactions on Image Processing*, 29, 5794-5804.
- Das, A. ve Rad, P., 2020, Opportunities and challenges in explainable artificial intelligence (xai): A survey, *arXiv preprint arXiv:2006.11371*.
- Das, K., Jiang, J. ve Rao, J., 2004, Mean squared error of empirical predictor.
- Datta, A., Sen, S. ve Zick, Y., 2016, Algorithmic transparency via quantitative input influence: Theory and experiments with learning systems, *2016 IEEE symposium on security and privacy (SP)*, 598-617.
- Eitel, F., Ritter, K. ve Initiative, A. s. D. N., 2019, Testing the robustness of attribution methods for convolutional neural networks in MRI-based Alzheimer's disease classification, *Interpretability of Machine Intelligence in Medical Image Computing and Multimodal Learning for Clinical Decision Support: Second*

- International Workshop, iMIMIC 2019, and 9th International Workshop, ML-CDS 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 17, 2019, Proceedings 9*, 3-11.
- Espadoto, M., Martins, R. M., Kerren, A., Hirata, N. S. ve Telea, A. C., 2019, Toward a quantitative survey of dimension reduction techniques, *IEEE transactions on visualization and computer graphics*, 27 (3), 2153-2173.
- Everingham, M., Van Gool, L., Williams, C. K. I., Winn, J. ve Zisserman, A., 2007, The PASCAL Visual Object Classes Challenge 2007, <http://host.robots.ox.ac.uk/pascal/VOC/voc2007/>:
- Everingham, M., Van Gool, L., Williams, C. K. I., Winn, J. ve Zisserman, A., 2012, Visual Object Classes Challenge 2012 (VOC2012), <http://host.robots.ox.ac.uk/pascal/VOC/voc2012/>:
- Forrest, E. ve Hoanca, B., 2015, Artificial intelligence: Marketing's game changer, *Trends and innovations in marketing information systems*, 45-64.
- Gardner, M. W. ve Dorling, S., 1998, Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences, *Atmospheric environment*, 32 (14-15), 2627-2636.
- Gunning, D. ve Aha, D., 2019, DARPA's explainable artificial intelligence (XAI) program, *AI magazine*, 40 (2), 44-58.
- Hamet, P. ve Tremblay, J., 2017, Artificial intelligence in medicine, *Metabolism*, 69, S36-S40.
- Hase, P. ve Bansal, M., 2020, Evaluating explainable AI: Which algorithmic explanations help users predict model behavior?, *arXiv preprint arXiv:2005.01831*.
- Henry, E. ve Hofrichter, J., 1992, [8] Singular value decomposition: Application to analysis of experimental data, In: *Methods in enzymology*, Eds: Elsevier, p. 129-192.
- Hilpisch, Y., 2020, *Artificial Intelligence in Finance*, O'Reilly Media, p.
- Holzinger, A., Langs, G., Denk, H., Zatloukal, K. ve Müller, H., 2019, Causability and explainability of artificial intelligence in medicine, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9 (4), e1312.
- Huawei, 2023, M2150-10-EI, <https://support.huawei.com/enterprise/en/intelligentvision/m2150-10-ei-pid-250673491>:
- Infratec, 2023, Infrared HD Camera, <https://www.infratec-infrared.com/thermography/camera-series-variocamr-high-definition/>: [2023].
- Jiang, P.-T., Zhang, C.-B., Hou, Q., Cheng, M.-M. ve Wei, Y., 2021, Layercam: Exploring hierarchical class activation maps for localization, *IEEE Transactions on Image Processing*, 30, 5875-5888.
- Joseph, L. P., Joseph, E. A. ve Prasad, R., 2022, Explainable diabetes classification using hybrid Bayesian-optimized TabNet architecture, *Computers in biology and medicine*, 151, 106178.
- Jung, H. ve Oh, Y., 2021, Towards better explanations of class activation mapping, *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 1336-1344.
- kaggle, 2020a, face mask detection dataset, <https://www.kaggle.com/datasets/wobotintelligence/face-mask-detection-dataset/discussion>:
- kaggle, 2020b, Face Mask Detection ~12K Images Dataset, <https://www.kaggle.com/datasets/ashishjangra27/face-mask-12k-images-dataset>:

- Kang, D., Park, S. ve Paik, J., 2020, SdBAN: Salient object detection using bilateral attention network with dice coefficient loss, *IEEE access*, 8, 104357-104370.
- Kaplan, K., Kaya, Y., Kuncan, M. ve Ertunç, H. M., 2020, Brain tumor classification using modified local binary patterns (LBP) feature extraction methods, *Medical hypotheses*, 139, 109696.
- Khaire, U. M. ve Dhanalakshmi, R., 2022, Stability of feature selection algorithm: A review, *Journal of King Saud University-Computer and Information Sciences*, 34 (4), 1060-1073.
- Khalid, S., Khalil, T. ve Nasreen, S., 2014, A survey of feature selection and feature extraction techniques in machine learning, *2014 science and information conference*, 372-378.
- Kim, H. E., Cosa-Linan, A., Santhanam, N., Jannesari, M., Maros, M. E. ve Ganslandt, T., 2022, Transfer learning for medical image classification: a literature review, *BMC medical imaging*, 22 (1), 69.
- Koh, P. W. ve Liang, P., 2017, Understanding black-box predictions via influence functions, *International conference on machine learning*, 1885-1894.
- Kramer, O. ve Kramer, O., 2016, Scikit-learn, *Machine learning for evolution strategies*, 45-53.
- Lau, M. M. ve Lim, K. H., 2018, Review of adaptive activation function in deep neural network, *2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, 686-690.
- Lecué, G. ve Lerasle, M., 2020, Robust machine learning by median-of-means: theory and practice.
- LeCun, Y., Bengio, Y. ve Hinton, G., 2015, Deep learning, *nature*, 521 (7553), 436-444.
- Liu, S. ve Deng, W., 2015, Very deep convolutional neural network based image classification using small training sample size, *2015 3rd IAPR Asian conference on pattern recognition (ACPR)*, 730-734.
- Maathuis, C., 2022, On explainable ai solutions for targeting in cyber military operations, *International Conference on Cyber Warfare and Security*, 166-175.
- Marino, M., Li, Y., Rueschman, M. N., Winkelman, J. W., Ellenbogen, J., Solet, J. M., Dulin, H., Berkman, L. F. ve Buxton, O. M., 2013, Measuring sleep: accuracy, sensitivity, and specificity of wrist actigraphy compared to polysomnography, *Sleep*, 36 (11), 1747-1755.
- Marreiros, A. C., Daunizeau, J., Kiebel, S. J. ve Friston, K. J., 2008, Population dynamics: variance and the sigmoid activation function, *Neuroimage*, 42 (1), 147-157.
- Mateen, M., Wen, J., Nasrullah, Song, S. ve Huang, Z., 2018, Fundus image classification using VGG-19 architecture with PCA and SVD, *Symmetry*, 11 (1), 1.
- Minh, D., Wang, H. X., Li, Y. F. ve Nguyen, T. N., 2022, Explainable artificial intelligence: a comprehensive review, *Artificial Intelligence Review*, 1-66.
- Montavon, G., Binder, A., Lapuschkin, S., Samek, W. ve Müller, K.-R., 2019, Layer-wise relevance propagation: an overview, *Explainable AI: interpreting, explaining and visualizing deep learning*, 193-209.
- Moore, R. C. ve DeNero, J., 2011, L1 and L2 regularization for multiclass hinge loss models.
- Morgan, F. E., Boudreaux, B., Lohn, A. J., Ashby, M., Curriden, C., Klima, K. ve Grossman, D., 2020, Military applications of artificial intelligence: ethical

- concerns in an uncertain world, *RAND PROJECT AIR FORCE SANTA MONICA CA SANTA MONICA United States*.
- Navada, A., Ansari, A. N., Patil, S. ve Sonkamble, B. A., 2011, Overview of use of decision tree algorithms in machine learning, *2011 IEEE control and system graduate research colloquium*, 37-42.
- Niu, S., Liu, Y., Wang, J. ve Song, H., 2020, A decade survey of transfer learning (2010–2020), *IEEE Transactions on Artificial Intelligence*, 1 (2), 151-166.
- Ornek, A. H., Ceylan, M. ve Ervural, S., 2019, Health status detection of neonates using infrared thermography and deep convolutional neural networks, *Infrared Physics & Technology*, 103, 103044.
- Ornek, A. H. ve Ceylan, M., 2022, Medical thermograms' classification using deep transfer learning models and methods, *Multimedia Tools and Applications*, 81 (7), 9367-9384.
- Ornek, A. H. ve Ceylan, M., 2023, CodCAM: A new ensemble visual explanation for classification of medical thermal images, *Quantitative InfraRed Thermography Journal*, 1-25.
- Owotogbe, J., Ibiyemi, T. ve Adu, B., 2019, Edge detection techniques on digital images-a review, *Int. J. Innov. Sci. Res. Technol*, 4 (11), 329-332.
- Örnek, A., Celik, M. ve Ceylan, M., 2021a, Explainable artificial intelligence: How face masks are detected via deep neural networks, *International Journal of Innovative Science and Research Technology*, 6 (9), 1104-1112.
- Örnek, A. H. ve Ceylan, M., 2020, Explainable features in classification of neonatal thermograms, *2020 28th Signal Processing and Communications Applications Conference (SIU)*, 1-4.
- Örnek, A. H. ve Ceylan, M., 2021, Explainable artificial intelligence (XAI): Classification of medical thermal images of neonates using class activation maps, *Traitement du Signal*.
- Örnek, A. H., Çelik, M. ve Ceylan, M., 2021b, Mask detection from face images using deep learning and transfer learning, *2021 15th Turkish National Software Engineering Symposium (UYMS)*, 1-4.
- Örnek, A. H. ve Ceylan, M., 2022a, A Novel Approach for Visualization of Class Activation Maps with Reduced Dimensions, *2022 Innovations in Intelligent Systems and Applications Conference (ASYU)*, 1-5.
- Örnek, A. H. ve Ceylan, M., 2022b, HayCAM: A Novel Visual Explanation for Deep Convolutional Neural Networks, *Traitement du Signal*.
- Panwar, H., Gupta, P., Siddiqui, M. K., Morales-Menendez, R., Bhardwaj, P. ve Singh, V., 2020, A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images, *Chaos, Solitons & Fractals*, 140, 110190.
- Papanastasopoulos, Z., Samala, R. K., Chan, H.-P., Hadjiiski, L., Paramagul, C., Helvie, M. A. ve Neal, C. H., 2020, Explainable AI for medical imaging: deep-learning CNN ensemble for classification of estrogen receptor status from breast MRI, *Medical imaging 2020: Computer-aided diagnosis*, 228-235.
- Pereira, S., Meier, R., Alves, V., Reyes, M. ve Silva, C. A., 2018, Automatic brain tumor grading from MRI data using convolutional neural networks and quality assessment, *Understanding and Interpreting Machine Learning in Medical Image Computing Applications: First International Workshops, MLCN 2018, DLF 2018, and iMIMIC 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16-20, 2018, Proceedings 1*, 106-114.

- Praveen, S. P., Srinivasu, P. N., Shafi, J., Wozniak, M. ve Ijaz, M. F., 2022, ResNet-32 and FastAI for diagnoses of ductal carcinoma from 2D tissue slides, *Scientific Reports*, 12 (1), 20804.
- Purves, R. D., 1992, Optimum numerical integration methods for estimation of area-under-the-curve (AUC) and area-under-the-moment-curve (AUMC), *Journal of pharmacokinetics and biopharmaceutics*, 20, 211-226.
- Ramaswamy, H. G., 2020, Ablation-cam: Visual explanations for deep convolutional network via gradient-free localization, *proceedings of the IEEE/CVF winter conference on applications of computer vision*, 983-991.
- Remeseiro, B. ve Bolon-Canedo, V., 2019, A review of feature selection methods in medical applications, *Computers in biology and medicine*, 112, 103375.
- Riedl, M. O. ve Zook, A., 2013, AI for game production, *2013 IEEE Conference on Computational Intelligence in Games (CIG)*, 1-8.
- Ringnér, M., 2008, What is principal component analysis?, *Nature biotechnology*, 26 (3), 303-304.
- Robert, C., 2014, Machine learning, a probabilistic perspective, Taylor & Francis.
- Rocchi, L., Chiari, L. ve Cappello, A., 2004, Feature selection of stabilometric parameters based on principal component analysis, *Medical and Biological Engineering and Computing*, 42, 71-79.
- Sayres, R., Taly, A., Rahimy, E., Blumer, K., Coz, D., Hammel, N., Krause, J., Narayanaswamy, A., Rastegar, Z. ve Wu, D., 2019, Using a deep learning algorithm and integrated gradients explanation to assist grading for diabetic retinopathy, *Ophthalmology*, 126 (4), 552-564.
- Schmidt-Hieber, J., 2020, Nonparametric regression using deep neural networks with ReLU activation function.
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D. ve Batra, D., 2017, Grad-cam: Visual explanations from deep networks via gradient-based localization, *Proceedings of the IEEE international conference on computer vision*, 618-626.
- Sharma, S., Sharma, S. ve Athaiya, A., 2017, Activation functions in neural networks, *Towards Data Sci*, 6 (12), 310-316.
- Singh, A., Sengupta, S. ve Lakshminarayanan, V., 2020, Explainable deep learning models in medical image analysis, *Journal of Imaging*, 6 (6), 52.
- Sliwinski, J., Strobel, M. ve Zick, Y., 2017, A characterization of monotone influence measures for data classification, *Proceedings of the Workshop on Explainable AI (XAI); International Joint Conferences on Artificial Intelligence (IJCAI), Melbourne, Australia*, 19-25.
- Szegedy, C., Ioffe, S., Vanhoucke, V. ve Alemi, A., 2017, Inception-v4, inception-resnet and the impact of residual connections on learning, *Proceedings of the AAAI conference on artificial intelligence*.
- Van der Waa, J., Nieuwburg, E., Cremers, A. ve Neerincx, M., 2021, Evaluating XAI: A comparison of rule-based and example-based explanations, *Artificial Intelligence*, 291, 103404.
- Van Molle, P., De Strooper, M., Verbelen, T., Vankeirsbilck, B., Simoens, P. ve Dhoedt, B., 2018, Visualizing Convolutional Neural Networks to Improve Decision Support for Skin Lesion Classification, *Cham*, 115-123.
- Vidal, R., Ma, Y., Sastry, S. S., Vidal, R., Ma, Y. ve Sastry, S. S., 2016, Principal component analysis, *Generalized principal component analysis*, 25-62.
- Wang, H., Wang, Z., Du, M., Yang, F., Zhang, Z., Ding, S., Mardziel, P. ve Hu, X., 2020a, Score-CAM: Score-weighted visual explanations for convolutional

- neural networks, *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, 24-25.
- Wang, Q., Ma, Y., Zhao, K. ve Tian, Y., 2020b, A comprehensive survey of loss functions in machine learning, *Annals of Data Science*, 1-26.
- Weiss, K., Khoshgoftaar, T. M. ve Wang, D., 2016, A survey of transfer learning, *Journal of Big data*, 3 (1), 1-40.
- Winston, P. H., 1984, Artificial intelligence, Addison-Wesley Longman Publishing Co., Inc., p.
- Yang, H.-L., Kim, J. J., Kim, J. H., Kang, Y. K., Park, D. H., Park, H. S., Kim, H. K. ve Kim, M.-S., 2019a, Weakly supervised lesion localization for age-related macular degeneration detection using optical coherence tomography images, *PloS one*, 14 (4), e0215076.
- Yang, K., Li, D. ve Dou, Y., 2019b, Towards precise end-to-end weakly supervised object detection network, *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 8372-8381.
- Yu, X., Wang, J., Hong, Q.-Q., Teku, R., Wang, S.-H. ve Zhang, Y.-D., 2022, Transfer learning for medical images analyses: A survey, *Neurocomputing*, 489, 230-254.
- Zeiler, M. D., Krishnan, D., Taylor, G. W. ve Fergus, R., 2010, Deconvolutional networks, *2010 IEEE Computer Society Conference on computer vision and pattern recognition*, 2528-2535.
- Zhang, D., Han, J., Cheng, G. ve Yang, M.-H., 2021, Weakly supervised object localization and detection: A survey, *IEEE transactions on pattern analysis and machine intelligence*, 44 (9), 5866-5885.
- Zhao, H., Zhan, Z.-H., Lin, Y., Chen, X., Luo, X.-N., Zhang, J., Kwong, S. ve Zhang, J., 2019, Local binary pattern-based adaptive differential evolution for multimodal optimization problems, *IEEE transactions on cybernetics*, 50 (7), 3343-3357.
- Zhou, B., Khosla, A., Lapedriza, A., Oliva, A. ve Torralba, A., 2016, Learning deep features for discriminative localization, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2921-2929.
- Zhou, D., Fang, J., Song, X., Guan, C., Yin, J., Dai, Y. ve Yang, R., 2019, Iou loss for 2d/3d object detection, *2019 international conference on 3D vision (3DV)*, 85-94.
- Ziou, D. ve Tabbone, S., 1998, Edge detection techniques-an overview, *Pattern Recognition and Image Analysis C/C of Raspoznavaniye Obrazov I Analiz Izobrazhenii*, 8, 537-559.
- Zupan, J., 1994, Introduction to artificial neural network (ANN) methods: what they are and how to use them, *Acta Chimica Slovenica*, 41, 327-327.

EKLER**EK-1 Tez Çalışmasında Kullanılan Termal Görüntülerin Alınması için Etik Kurul Kararı**

T.C.
SELÇUK ÜNİVERSİTESİ
TIP FAKÜLTESİ DEKANLIĞI

GİRİŞİMSEL OLMAYAN KLİNİK ARAŞTIRMALAR ETİK KURULU KARARLARI

Toplantı Sayısı: 2015/1

Toplantı Tarihi : 06.01.2015

Karar Sayısı 2015/16 S.Ü. Mühendislik Fakültesi Elektrik-Elektronik Mühendisliği Anabilim Dalı Öğretim Üyesi Yrd.Doç.Dr. Murat CEYLAN'ın, "Termal Görüntüleme Tabanlı Yeni Doğan Destek Ünitesi" başlıklı araştırmasının değerlendirilme talebi ile ilgili 25.12.2014 tarihli dilekçesi ve ekleri görüşüldü.

Yapılan inceleme ve görüşmelerden sonra; Yrd.Doç.Dr. Murat CEYLAN'ın, "Termal Görüntüleme Tabanlı Yeni Doğan Destek Ünitesi" adlı araştırmanın kabulüne, TÜBİTAK desteği alındıktan sonra protokolün dosyaya ilave edilmek üzere Etik Kurul sekreteryasına teslim edilmesine oy birliği ile karar verildi.

