



T.C.
EGE ÜNİVERSİTESİ
Fen Bilimleri Enstitüsü



**DERİN ÖĞRENME KULLANILARAK GÖZ,
YÜZ VE ESNEME ÖZELLİKLERİNE DAYALI
MELEZ BİR YORGUNLUK TANIMA SİSTEMİ
GELİŞTİRİLMESİ**

Yüksek Lisans Tezi

Rabia KÖK

Bilgisayar Mühendisliği Anabilim Dalı

İzmir
2021

T.C.
EGE ÜNİVERSİTESİ
Fen Bilimleri Enstitüsü

**DERİN ÖĞRENME KULLANILARAK GÖZ,
YÜZ VE ESNEME ÖZELLİKLERİNE DAYALI
MELEZ BİR YORGUNLUK TANIMA SİSTEMİ
GELİŞTİRİLMESİ**

Rabia KÖK

Danışman: Prof. Dr. Aybars UĞUR

Bilgisayar Mühendisliği Anabilim Dalı
Bilgisayar Mühendisliği Yüksek Lisans Programı

İzmir
2021

EGE ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ

ETİK KURALLARA UYGUNLUK BEYANI

EÜ Lisansüstü Eğitim ve Öğretim Yönetmeliğinin ilgili hükümleri uyarınca Yüksek Lisans Tezi olarak sunduğum “**Derin Öğrenme Kullanılarak Göz, Yüz ve Esneme Özelliklerine Dayalı Melez Bir Yorgunluk Tanıma Sistemi Geliştirilmesi**” başlıklı bu tezin kendi çalışmam olduğunu, sunduğum tüm sonuç, doküman, bilgi ve belgeleri bizzat ve bu tez çalışması kapsamında elde ettiğimi, bu tez çalışmasıyla elde edilmeyen bütün bilgi ve yorumlara atıf yaptığımı ve bunları kaynaklar listesinde usulüne uygun olarak verdiğimi, tez çalışması ve yazımı sırasında patent ve telif haklarını ihlal edici bir davranışımın olmadığını, bu tezin herhangi bir bölümünü bu üniversite veya diğer bir üniversitede başka bir tez çalışması içinde sunmadığımı, bu tezin planlanmasından yazımına kadar bütün safhalarda bilimsel etik kurallarına uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul edeceğimi beyan ederim.

01 / 09 / 2021

Rabia KÖK

ÖZET**DERİN ÖĞRENME KULLANILARAK GÖZ, YÜZ VE ESNEME
ÖZELLİKLERİNE DAYALI MELEZ BİR YORGUNLUK TANIMA
SİSTEMİ GELİŞTİRİLMESİ**

KÖK, Rabia

Yüksek Lisans Tezi, Bilgisayar Mühendisliği Anabilim Dalı
Tez Danışmanı: Prof. Dr. Aybars UĞUR
Eylül 2021, 75 sayfa

Trafik kazaları genel olarak sürücü, araç ve çevresel etmenlerin etkileşimi sonucu oluşmakta; can ve mal güvenliğini ciddi oranda tehdit etmektedir. Bu kazalarda genellikle ana faktör insandır. Uzun süre ve/veya dinlenmeden araç kullanmak, sürücünün yorgun düşmesine ya da sürüş esnasında uyuklamasına sebep olmaktadır. Yorgun/uykusuz araç kullanmak, en az alkollü araç kullanmak kadar tehlikelidir.

Görüntü işleme tabanlı yorgunluk tanımaaya yönelik bu tez çalışmasında, literatürde yer alan, ücretsiz erişilebilen Cew ve YawDD veri setleri kullanılmıştır. Cew veri seti, kapalı ve açık göz görüntüleri ile kapalı ve açık gözlerin bulunduğu yüz görüntüleri içermektedir. YawDD veri seti ise esneme, konuşma, şarkı söyleme gibi davranışlarda bulunan sürücü görüntüleri barındıran video görüntülerinden oluşmaktadır. Öncelikle bu veri setinden, esneyen ve esnemeyen sürücülerin görüntü kareleri ayrıştırılmış ve ardından MTCNN yöntemi kullanılarak görüntülerdeki yüz, göz ve ağız görüntü parçaları çıkarılmıştır. Yüz, göz ve ağız görüntülerinden yorgunluk tanıma için derin öğrenme modelleri geliştirilmiştir. Ayrıca, i) yüz ve gözlerin; ii) yüz, gözler ve ağzın birlikte değerlendirildiği fonksiyonel modellerin gerçekleştirimi yapılmıştır. Çalışma, yorgunluk tanıma alanında YawDD veri seti üzerinde ve durağan yüz bölümleri görüntülerinin öz niteliklerini birleştirmek üzere geliştirilen ilk modeli içermesi yönü ile yenilik içermektedir. Deneysel çalışmalar, yöntemin belirli durumlarda sınıflandırma başarısını artırabildiğini göstermektedir.

Anahtar sözcükler: Yorgunluk Tanıma, Melez Sistem, Derin Öğrenme, Model Ortalama, Topluluk Öğrenmesi, Evrimsel Sinir Ağları

ABSTRACT**DEVELOPMENT OF HYBRID FATIGUE RECOGNITION SYSTEM
BASED ON EYE, FACE AND YAW FEATURES USING DEEP
LEARNING**

KÖK, Rabia

MSc in Computer Eng.
Supervisor: Prof. Dr. Aybars UĞUR
September 2021, 75 pages

Traffic accidents generally occur as a result of the interaction of driver, vehicle, and environmental factors; threat the safety of life and property seriously. People are usually the main factor in these accidents. Driving for a long time and/or without resting causes the driver to get tired or drowsy while driving. Tired/sleepless driving is just as dangerous as drunk driving.

In this thesis study on image processing-based fatigue recognition, the freely accessible Cew and YawDD datasets in the literature were used. The Cew dataset includes images of closed and open eyes and facial images with closed and open eyes. On the other hand, the YawDD dataset consists of video images of drivers engaging in behaviors such as yawning, speaking, and singing. First of all, image frames of yawning and non-yawning drivers were separated from this dataset and then the face, eye, and mouth parts of the images were extracted from the images using MTCNN method. Deep learning models have been developed for fatigue recognition using face, eye, and mouth images. In addition, functional models were implemented in which i) face and eyes, ii) face, eyes, and mouth were evaluated together. The study has a contribution in the field of fatigue recognition on the YawDD dataset in that includes the first model developed to combine the features of the images of stationary facial parts. Experimental studies show that the method can increase classification success in certain situations.

Keywords: Fatigue Recognition, Hybrid System, Deep Learning, Model Averaging, Ensemble Learning, Convolutional Neural Networks

ÖNSÖZ

Bu tez çalışmasında yorgunluk/uykusuzluk tanımada en belirgin özelliklerin gözlerin kapanması ve esneme olup olmamasına odaklanılmıştır. Bu özellikler kullanılarak yorgunluk/uykusuzluk tespiti için CNN mimarisine dayalı derin öğrenme modellerinin gerçekleştirimi yapılmış ve topluluk öğrenmesi yöntemi kullanılarak veri setlerinin birlikte değerlendirilmesi sağlanmıştır. Deneysel çalışmalar belirlenen iki veri seti üzerinde gerçekleştirilmiştir. Bu tez çalışmasında, yorgunluk/uykusuzluk tanıma konusunda farklı bir yaklaşımla farklı yönlerin ortaya çıkarılması amaçlanmıştır.

Bu tez çalışmasının hazırlanmasında bilgilerini ve tecrübelerini hiçbir zaman esirgemeyen ve çalışmanın her aşamasında kıymetli vaktini ayırarak bana yardımcı olan değerli hocam Prof. Dr. Aybars UĞUR’a teşekkürü bir borç bilirim.

İZMİR

01/09/2021

Rabia KÖK

İÇİNDEKİLERSayfa

İÇ KAPAK.....	i
KABUL VE ONAY SAYFASI.....	iii
ETİK KURALLARA UYGUNLUK BEYANI.....	v
ÖZET.....	vii
ABSTRACT.....	ix
ÖNSÖZ.....	xi
İÇİNDEKİLER.....	xiii
ŞEKİLLER DİZİNİ.....	xvi
TABLolar DİZİNİ.....	xx
KISALTMALAR DİZİNİ.....	xxii
1. GİRİŞ.....	1
2. KURAMSAL TEMELLER.....	3
2.1. Makine Öğrenmesi.....	3
2.1.1. Denetimli öğrenme.....	4
2.1.2. Denetimsiz öğrenme.....	4
2.2. Derin Öğrenme.....	5

İÇİNDEKİLER (devam)

Sayfa

2.2.1. Evrişimsel sinir ağları	6
2.3. Topluluk Öğrenmesi ve Model Ortalaması.....	11
2.4. Nirengi Noktası Bulma Yöntemi ve MTCNN.....	13
3. ÖNCEKİ ÇALIŞMALAR.....	17
4. DERİN ÖĞRENME KULLANILARAK GÖZ, YÜZ VE ESNEME ÖZELLİKLERİNE DAYALI MELEZ BİR YORGUNLUK TANIMA SİSTEMİ.....	21
4.1. Giriş.....	21
4.2. Veri Setlerinin Düzenlenmesi.....	23
4.3. CNN Mimarisine Dayalı Sınıflandırma Modeli.....	24
4.4. Topluluk Öğrenmesi Yöntemine Dayalı Sınıflandırma Modeli.....	31
5. DENEYSEL ÇALIŞMALAR.....	33
5.1. Giriş.....	33
5.2. Başarım Ölçütleri.....	33
5.3. Kullanılan Veri Setleri.....	36
5.3.1. Cew veri seti.....	36
5.3.2. YawDD veri seti.....	37

İÇİNDEKİLER (devam)

Sayfa

5.4. Veri Ön İşleme.....	38
5.5. CNN Mimarisine Dayalı Derin Öğrenme Modeli Gerçekleştirimi.....	38
5.6. Topluluk Öğrenmesine Dayalı Derin Öğrenme Modeli Gerçekleştirimi.....	41
5.7. CEW Veri Seti Üzerindeki Deneysel Sonuçlar.....	41
5.8. YawDD Veri Seti Üzerindeki Deneysel Sonuçlar.....	48
5.9. YawDD Veri Setinde Deneylerin Tekrarlanması ile Elde Edilen Deneysel Sonuçlar.....	58
5.10. Cew ve YawDD Veri Setlerinin Birleştirilmesi ile Elde Edilen Deneysel Sonuçlar.....	60
5.11. Önerilen Yöntemin Literatürdeki Çalışmalar ile Karşılaştırılması.....	63
6. SONUÇ VE ÖNERİLER.....	66
KAYNAKLAR DİZİNİ.....	68
TEŞEKKÜR.....	74
ÖZGEÇMİŞ.....	75
EKLER.....	

ŞEKİLLER DİZİNİ

<u>Şekil</u>	<u>Sayfa</u>
2.1. CNN Mimarisi.....	7
2.2. Evrişim Katmanında filtre dolaştırılması ile özellik haritası oluşturma işlemi...	8
2.3. Düzleştirilmiş doğrusal birim katmanının çıkış verisine etkisi.....	8
2.4. Girdi görüntüsüne 2x2 boyutunda filtre ile iki adım kaymalı uygulanan maksimum havuzlama işlemi.....	9
2.5. CNN ağına seyrekleştirme uygulanmadan önceki ağ yapısı.....	10
2.6. CNN ağına seyrekleştirme uygulandıktan sonra ağ yapısı.....	10
2.7. P-Net Ağ Yapısı.....	15
2.8. R-Net Ağ Yapısı.....	15
2.9. O-Net Ağ Yapısı.....	16
4.1. Önerilen Sistemin Akış Şeması.....	22
4.2. YawDD veri seti üzerinde ayrıştırılan yüz, göz ve ağız bölgeleri.....	23
4.3. Çalışma kapsamında tasarlanan sıralı derin öğrenme modeli.....	24
4.4. Çalışma kapsamında yüz, sol göz, sağ göz özniteliklerinin birleştirildiği fonksiyonel derin öğrenme modeli.....	25
4.5. Çalışma kapsamında yüz, sol göz, sağ göz ve ağız özniteliklerinin birleştirildiği fonksiyonel derin öğrenme modeli.....	26

ŞEKİLLER DİZİNİ (devam)

Şekil

Sayfa

4.6. Cew veri setinde yüz görüntü parçasına dayalı sıralı derin öğrenme modeli.....	28
4.7. Cew veri setinde göz görüntü parçasına dayalı sıralı derin öğrenme modeli.....	28
4.8. YawDD veri setinde yüz görüntü parçasına dayalı sıralı derin öğrenme modeli.....	29
4.9. YawDD veri setinde göz görüntü parçasına dayalı sıralı derin öğrenme modeli.....	29
4.10. YawDD veri setinde ağız görüntü parçasına dayalı sıralı derin öğrenme modeli.....	29
4.11. Cew veri setinde yüz, sol göz ve sağ göz görüntü parçalarına dayalı fonksiyonel derin öğrenme modeli	30
4.12. YawDD veri setinde yüz, sol göz, sağ göz görüntü parçalarına dayalı fonksiyonel derin öğrenme modeli.....	30
4.13. YawDD veri setinde yüz, sol göz, sağ göz ve ağız göz görüntü parçalarına dayalı fonksiyonel derin öğrenme modeli.....	31
5.1. 24x24 piksel boyutundaki göz görüntüleri için elde edilen karışıklık matrisi..	46
5.2. 48x48 piksel boyutundaki göz görüntüleri için elde edilen karışıklık matrisi..	46
5.3. 100x100 piksel boyutundaki yüz görüntüleri için elde edilen karışıklık matrisi.....	47

ŞEKİLLER DİZİNİ (devam)

Şekil

Sayfa

5.4. 24x24 piksel boyutundaki sol göz ve sağ göz ile 100x100 piksel boyutundaki yüz görüntülerinin özniteliklerinin birleşimi sonucunda elde edilen karışıklık matrisi.....	47
5.5. 48x48 piksel boyutundaki sol göz ve sağ göz ile 100x100 piksel boyutundaki yüz görüntülerinin özniteliklerinin birleşimi sonucunda elde edilen karışıklık matrisi.....	48
5.6. 24x24 piksel boyutundaki göz görüntüleri için elde edilen karışıklık matrisi.....	54
5.7. 48x48 piksel boyutundaki göz görüntüleri için elde edilen karışıklık matrisi.....	55
5.8. 48x48 piksel boyutundaki ağız görüntüleri için elde edilen karışıklık matrisi.....	55
5.9. 100x100 piksel boyutundaki yüz görüntüleri için elde edilen karışıklık matrisi.....	56
5.10. 24x24 piksel boyutundaki sol göz ve sağ göz ile 100x100 piksel boyutundaki yüz görüntülerinin özniteliklerinin birleşimi sonucunda elde edilen karışıklık matrisi.....	56
5.11. 48x48 piksel boyutundaki sol göz ve sağ göz ile 100x100 piksel boyutundaki yüz görüntülerinin özniteliklerinin birleşimi sonucunda elde edilen karışıklık matrisi.....	57
5.12. 48x48 piksel boyutundaki ağız ile 100x100 piksel boyutundaki yüz görüntülerinin özniteliklerinin birleşimi sonucunda elde edilen karışıklık matrisi.....	57

ŞEKİLLER DİZİNİ (devam)ŞekilSayfa

- 5.13. 24x24 piksel boyutundaki sol göz ve sağ göz, 48x48 piksel boyutundaki ağız ile 100x100 piksel boyutundaki yüz görüntülerinin özniteliklerinin birleşimi sonucunda elde edilen karışıklık matrisi.....58
- 5.14. Cew ve YawDD veri setlerinde benzer özelliklerin birleştirilmesi için geliştirilen model ortalaması modeli.....61



TABLOLAR DİZİNİ

<u>Tablo</u>	<u>Sayfa</u>
5.1. Karışıklık Matrisi.....	34
5.2. Cew veri setinde göz görüntülerine ait veri miktarları.....	36
5.3. Cew veri setinde yüz görüntülerine ait veri miktarları.....	37
5.4. YawDD veri setindeki veri miktarı ile MTCNN kullandıktan sonra elde edilen veri miktarları.....	38
5.5. CNN modellerinde kullanılan hiperparametre değerleri.....	41
5.6. Cew veri seti üzerindeki sıralı modellerin eğitim ve doğrulama verileri üzerindeki başarı değerleri.....	42
5.7. Cew veri seti üzerindeki fonksiyonel modellerin eğitim ve doğrulama verileri üzerindeki başarı değerleri.....	43
5.8. Cew veri seti test verileri üzerindeki başarı değerleri.....	44
5.9. YawDD veri seti üzerindeki sıralı modellerin eğitim ve doğrulama verileri üzerindeki başarı değerleri.....	49
5.10. YawDD veri seti üzerindeki fonksiyonel modellerin eğitim ve doğrulama verileri üzerindeki başarı değerleri.....	50
5.11. YawDD veri setindeki test verileri üzerinde elde edilen başarı değerleri....	52
5.12. YawDD veri setinin eğitim ve doğrulama verileri üzerindeki ortalama kayıp ve doğruluk başarı değerleri.....	59

TABLÖLAR DİZİNİ (devam)

<u>Tablo</u>	<u>Sayfa</u>
5.13. YawDD veri setinin test verileri üzerinde elde edilen ortalama başarımların değerleri.....	59
5.14. Cew ve YawDD veri setlerinin test görüntüleri üzerindeki model ortalama yöntemi sınıflandırma sonuçları.....	62
5.15. Önerilen sıralı modellerin literatürdeki çalışmalar ile karşılaştırılması.....	64
5.16. Önerilen YawDD veri seti için en başarılı fonksiyonel modelin literatürdeki çalışmalar ile karşılaştırılması.....	64

KISALTMALAR DİZİNİ

<u>Kısaltmalar</u>	<u>Açıklama</u>
CEW	Closed Eyes In The Wild (Uykusuzluk Veri Seti)
CLAHE	Kontrast Sınırlı Uyarlanabilir Histogram Eşitleme
CNN	Evrişimsel Sinir Ağları
ESR-Net	Göz Durum Tanıma Ağı
FN	Yanlış Negatif
FP	Yanlış Pozitif
LSTM	Uzun - Kısa Süreli Bellek
MAR	Ağız En Boy Oranı
MOSSE	Kare Hatasının Çıktı Toplamını En Aza İndirme
MSR-Net	Ağız Durum Tanıma Ağı
MTCNN	Çok Görevli Basamaklı Evrişimsel Ağlar
NMS	Maksimum Olmayan Bastırma
NTHU-DDD	Ulusal Tsuing Hua Üniversitesi Sürücü Uykululuk Tespiti
O-Net	Çıktı Ağı
P-Net	Teklif Ağı
R-Net	İyileştirme Ağı

KISALTMALAR DİZİNİ (devam)

<u>Kısaltmalar</u>	<u>Açıklama</u>
SHRP2 NDS	İkinci Stratejik Otoyol Araştırma Programı 2 Doğal Sürüş Çalışması
TN	Doğru Negatif
TP	Doğru Pozitif
YawDD	Yawning Detection Dataset (Esname Tanıma Veri Seti)

1. GİRİŞ

Yorgunluk, insan vücudunun durma, dinlenme veya uyuma ihtiyacını belirtme şeklidir. Kötü uyku düzeni yorgunluğun temel sebebidir. Yorgunluğun giderilmesi için yapılması gereken dinlenmek ve uyumaktır. Araç kullanmada da yorgunluk önemli bir risk faktörüdür. Kötü ve yetersiz uyku; halsizlik, dikkat azalması, iş üretkenliğinde azalma ve hata yapma olasılığını artırmakta ve bu sebepten ötürü kaza yapma riskini artırmaktadır. Yorgun ve uykusuz araç kullanılması, alkollü araç kullanılması kadar tehlikeli olmakla birlikte en önemli kaza nedenlerinden birini oluşturmaktadır. Yorgunluk, uykusuzluk, alkol, uyuşturucu gibi etmenler sürüş performansını olumsuz etkilemekte ve ağır trafik kazalarına neden olmaktadır (Baş, 2015).

Ülkemizdeki trafik kazalarında her yıl 9 binden fazla can kaybı yaşanmakta, 200 bin kişi de yaralanmaktadır. Avrupa’da da trafik kazalarına bağlı olarak her yıl ortalama 40 bin kişi hayatını kaybetmektedir. Trafik kazaları, genel olarak sürücü, araç ve çevresel etkenlerin etkileşiminin bir sonucu olmakla birlikte trafik kaza analizlerinde trafik kazalarının %90’ına yakın bir kısmında ana faktör insandır. Amerika Birleşik Devletleri’nde trafik kazalarının %1-4’ünün, Norveç’te %4’ünün, Avustralya’da %6’sının ve İngiltere’de %16’sının sürücünün uyuyakalması veya aşırı uykululuğu nedeniyle meydana geldiği bildirilmiştir (Ursavaş ve Ege, 2004).

Ulusal Karayolu Trafik Güvenliği İdaresi raporuna göre, trafik kazalarının %22-24’ü sürücü yorgunluğundan kaynaklanmaktadır. Sürüş sırasındaki sürücü yorgunluğu, kaza riskini dört ila altı kat artırabilir. Trafik kazalarının sıklıkla meydana gelmesi, insanların hayatlarının ve mallarının güvenliğini ciddi şekilde tehdit etmektedir (Liu et al., 2019).

Trafik kazaları maddi ve manevi kayıplara sebep olmaktadır. Trafik kazalarının nedenlerinin belirlenip kaza sayısını en aza indirebilmek için yapılan çalışmalar sayesinde can kayıplarının ve sakatlanmaların önüne geçilebilmektedir. İnsan, doğası gereği ihtiyacı olan uyku gereksinimini tamamlayamadığında, uykusuzluk veya yorgunluğun sürücü davranışları üzerindeki olumsuz etkisi meydana gelmektedir. Dünyada ve ülkemizde trafik sayılarına ait istatistikler,

uykusuzluk/yorgunluk tanıma sistemlerinin hayati açıdan ne kadar önemli olabileceğini göstermektedir (Baş, 2015).

Kişilerin, kamera ve/veya donanım ekipmanları kullanılarak yorgun/uykusuz olup olmadığının tespit edilmesi işlemi, yorgunluk tanıma olarak nitelendirilmektedir. Donanım ekipmanları kullanılarak çeşitli sensörlerle kişilerin kalp ritim sinyalleri ile kişiler takibe alınmakta ve bu sinyallere göre yorgun veya yorgun değil olarak nitelendirilmektedir. Kamera görüntülerinden üzerinden genellikle kişilerin göz görüntüleri üzerinden, gözlerin kapalı açık olma durumuna göre kişilerin durumu yorgundur/uykusuzdur veya yorgun/uykusuz değildir olarak değerlendirilmektedir. Ayrıca baş pozisyonu ve/veya esneme durumu da göz önünde bulundurularak görüntüler üzerinden yorgunluk/uykusuzluk tespiti yapılabilmektedir.

Bu tez çalışmasında literatürde yer alan YawDD ve Cew veri setleri kullanılmıştır. Video kareleri içeren YawDD veri setinden, önce MTCNN (Çok Görevli Basamaklı Evrişimsel Ağlar) yöntemi kullanılarak yüz, sol göz, sağ göz ve ağız görüntüleri çıkarılmıştır. Durağan görüntüler üzerinde yorgunluk/uykusuzluk tanınması amacıyla CNN (Evrişimsel Sinir Ağı) mimarisine dayalı derin öğrenme modellerinin gerçekleştirimi yapılmıştır. Yüz, gözler ve ağız özelliklerinin her biri için sıralı; yüz, sol göz, sağ göz ve ağız özelliklerinin özniteliklerinin birleştirilmesi için fonksiyonel modeller üzerinde deneysel çalışmalar gerçekleştirilmiştir. Veri setlerinin birlikte değerlendirilmesi amacıyla model ortalama topluluk öğrenmesi yöntemi kullanılmıştır.

Bu tez çalışmasının ikinci bölümünde kuramsal temellere yer verilmektedir. Bu bölümde makine öğrenmesi, derin öğrenme ve nesne tespit yöntemleri anlatılmaktadır. Üçüncü bölümde, literatürdeki yorgunluk tanıma üzerine yapılan çalışmalara yer verilmektedir. Dördüncü bölümde bu tez çalışması kapsamında yorgunluk tanıma amacıyla geliştirilen CNN mimarisine dayalı derin öğrenme yöntemleri kullanılarak yapılan sınıflandırma yöntemi açıklanmaktadır. Beşinci bölümde bu tez kapsamında yapılan deneysel çalışmalar, sonuçlar ve değerlendirme bulunmaktadır. Altıncı bölümde ise sonuçlar ve gelecek çalışmalar için sunulan öneriler yer almaktadır.

2. KURAMSAL TEMELLER

Bu bölümde, tez çalışmasında kullanılan yöntemlerin bilimsel açıklaması yer almaktadır.

Makine öğrenmesinin alt dalları olan derin öğrenme ile topluluk öğrenmesi temel alınarak makine öğrenmesi yöntemlerine yer verilmektedir. Bununla birlikte, bu çalışmada kullanılan derin öğrenme ve topluluk öğrenmesi yöntemleri üzerinde daha ayrıntılı bir anlatım yapılmaktadır.

2.1. Makine Öğrenmesi

Makine öğrenmesi, veriden bilgi çıkarımı ile ilgilenir. İstatistik, yapay zeka ve bilgisayar biliminin kesiştiği bir araştırma alanıdır. Tahmine dayalı analitik veya istatistiksel öğrenme olarak da bilinir (Müller and Guido, 2016).

Bilgisayar bilimi, manuel olarak nasıl programlama yapılacağına; istatistik, öncelikle verilerden hangi sonuçların çıkarılabileceğine odaklanır. Makine öğrenmesi ise bilgisayarların kendilerini nasıl programlayacaklarına; verileri en etkili şekilde yakalamak, depolamak, indekslemek, almak ve birleştirmek için hangi hesaplama mimarilerinin ve algoritmalarının kullanılabileceğine, daha büyük bir sistemde birden fazla öğrenme alt görevinin nasıl düzenlenebileceğine odaklanır. Makine öğrenmesi ile ilgili alanlar ve Psikoloji, Sinirbilim de insan ve hayvan öğrenimini inceleyerek bilgisayarların nasıl öğrenebileceğine odaklanmaktadır (Mitchell, 2006).

Makine öğrenmesinde örnek verileri veya geçmiş deneyimleri kullanarak bir performans kriterini optimize etmek için bilgisayarlar programlanır. Bu amaçla bazı parametreler içeren ve öğrenen bir model tanımlanır. Model, geleceğe yönelik tahminler yapmak için öngörücü veya verilerden bilgi edinmek için tanımlayıcı yahut her ikisi de olabilir.

Makine öğrenmesi, matematiksel modeller oluştururken istatistik teorisini kullanır, çünkü asli görevi bir örnekten çıkarım yapmaktır. Bilgisayar biliminin rolü

iki yönlüdür: Birincisi, eğitimde, optimizasyon problemini çözmek ve genel olarak sahip olunan büyük miktarda veriyi depolamak ve işlemek için verimli algoritmalara ihtiyaç vardır. İkincisi, bir model bir kez öğrenildiğinde, onun temsili ve çıkarım için algoritmik çözümünün de verimli olması gerekir. Bazı uygulamalarda, öğrenme veya çıkarım algoritmasının verimliliği, yani uzay ve zaman karmaşıklığı, tahmin doğruluğu kadar önemli olabilir (Alpaydın, 2010).

Makine öğrenmesinde kullanılabilecek çeşitli algoritmalar vardır. Makine öğrenmesi algoritmaları genellikle iki öğrenme biçiminden birine girer: denetimli ya da denetimsiz öğrenme (Bhukya, 2021).

2.1.1. Denetimli öğrenme

Denetimli makine öğrenmesi, elde edilen sonuçlarda kabul görmüş bir gerçeklik kaynağına mümkün olduğunca yakın bir değer verme amacı taşımaktadır. Denetimli öğrenmede ilk olarak uzman insanlarca hedef değerleri bilinen bir veri grubu edinilir. Edinilen bu veri, eğitim ve doğrulama olmak üzere gruplara ayrılır. Sonuca dair çıktı oluşturan model parametrelerini belirleyen grup, eğitim verisi grubu olarak adlandırılır. Doğrulama grubu ise, model eğitiminde modelin doğrulama oranının test edilmesi amacıyla oluşturulur. Model, doğrulama grubuna göre bir çıktı üretir. Bu çıktılarla uzmanlarca belirlenen çıktılar karşılaştırılarak modelin doğruluğu belirlenir. Model doğruluğundan kayda değer bir sonuç elde edilmediği takdirde, veri sayısı ya da özellik belirleme yönteminde farklılığa gidilmesi gibi yöntemlerle hata azaltılabilir. Model doğrulama başarısını uzman çıktısı ile karşılaştırmada iyi bir sonuç elde edilinceye kadar veri, özellik ve algoritma üzerindeki revize işlemlerine devam edilir. Denetimli makine öğrenmesi, veriler üzerinde sınıflandırma ya da tahminleme yapmak amacıyla kullanılmaktadır (Iansiti and Lakhani, 2021).

2.1.2. Denetimsiz öğrenme

Denetimsiz makine öğrenmesinde az sayıda bilgi ve varsayım üzerinden sonuç üretilmektedir. Denetimsiz öğrenme algoritmaları, denetimli öğrenme algoritmalarına karşın az sayıda bilgi ve varsayımdan yola çıkmakla birlikte veriler

üzerinde bir uzman sınıflandırması olmaksızın veya beklenen değerler sağlanmadan çalışır. Asıl amacı denetçinin farkına varamayacağı farklı yapıları ortaya koymaktır. Oluşan bu farklı kalıplar yahut gruplar kişiler ya da algoritmalar ile etiketlenerek sonuca varılır. Denetimsiz öğrenmede, kalıplaşmış etiketler yerine istatistiksel gruplar önerilmekte, bu gruplara da kişiler karar vermekte veya algoritmalar ile karar verilmektedir (Iansiti and Lakhani, 2021).

2.2. Derin Öğrenme

Derin öğrenme, insana özgü yetenekleri taklit ederek gözlemlene, analiz etme, öğrenme ve karar verme gibi işlemleri yapabilen bir makine öğrenmesi yöntemidir. (Kayaalp ve Süzen, 2018) Derin öğrenme yöntemleri, evrimsel tabanlı sinir ağı modeline dayanmaktadır (Metlek ve Çetiner, 2021). Bu nedenle derin sinir ağları olarak da adlandırılırlar. Derin öğrenme mimarileri, çok sayıda katman ve parametre içermektedir (Shinde and Shah, 2018).

Derin öğrenmede, çoklu katmanlarda doğrusal olmayan işleme ile denetimli veya denetimsiz öğrenme olmak üzere iki temel faktör göz önünde bulundurulur (Vargas et al., 2017). Özellik çıkarma ve dönüştürme işlemleri için doğrusal olmayan işleme kullanılmaktadır. Derin öğrenme, alt katmanlardan üst katmanlara doğru daha derin özelliklerin öğrenildiği hiyerarşik bir mimari yapıya sahiptir. Alt katmanlar, veri girişine yakın katmanlardır ve bu katmanlarda basit özellikler öğrenilir. Üst katmanlarda ise alt katmanlardan gelen özelliklerden daha karmaşık özellikler öğrenilir. Derin öğrenmenin sahip olduğu bu hiyerarşik yapı, verilerden faydalı bilgilerin analiz edilmesine olanak sunmaktadır (Shinde and Shah, 2018).

Derin öğrenme algoritmaları, verilerden temsilleri (öznitelikleri) otomatik olarak çıkarması özelliği ile geleneksel makine öğrenme yöntemlerinden farklılık göstermektedir (Valliani et al., 2019). Geleneksel makine öğrenmesinde öncelikle özellik vektörünün oluşturulması gerekmektedir. Özellik vektörleri, alanında uzman kişilerce çıkarılmaktadır. Derin öğrenme, ham veriler üzerinden işlem gerçekleştirdiğinden ham veriyi işlerken katmanlar arası oluşturduğu temsillerle özellik çıkarımını kendisi gerçekleştirmektedir (İnik ve Ülker, 2017). Derin evrimsel sinir ağı içerisinde özellik çıkarımı yapılmakta, ağı katmanları

içerisinde tespit edilecek yapıya ait özelliklerle belirlenmektedir. (Doğan ve Türkoğlu, 2019) Hiyerarşik yapı içerisinde yer alan düşük ve yüksek seviyeli özellikler, temsil çıkarma gerekliliğini ortadan kaldırmakta ve yeni verilere daha iyi genelleştirme yapılmasına olanak tanımaktadır (Valliani et al., 2019).

Derin öğrenme, gelişmekte olan bir yaklaşımdır. Anlamsal ayrıştırma, transfer öğrenme, doğal dil işleme, bilgisayarla görme gibi geleneksel yapay zeka alanlarında yaygın olarak uygulanmaktadır (Guo et al., 2016). Evrişimsel Sinir Ağı, AlexNet ZFNet, GoogLeNet, Microsoft ResNet ve R-CNN mimarileri, derin öğrenmede kullanılan temel mimarilerdir (İnik ve Ülker, 2017).

2.2.1. Evrişimsel sinir ağları

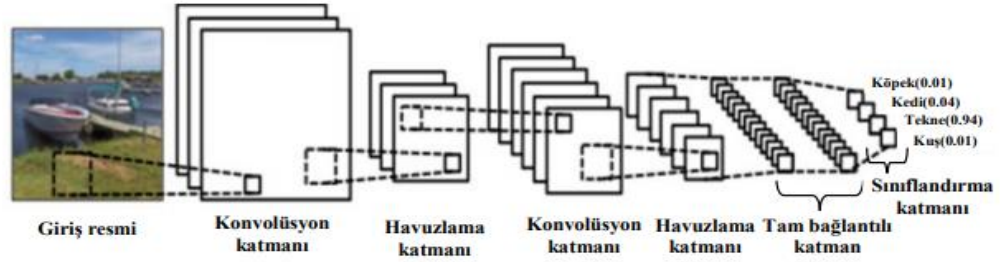
Derin öğrenme mimarileri içerisinde en çok kullanılan sinir ağı Evrişimsel Sinir Ağı yani CNN'dir. Genellikle görüntü analizlerinin yapılması için kullanılmaktadır. Filtrelemeye dayalı bir yapıya sahiptir. İlk olarak LeCun ve arkadaşları tarafından gradyan temelli bir ağ yapısı olarak tasarlanmıştır (Doğan ve Türkoğlu, 2019). Çok sayıda gizli katman barındırmaktadır. Görüntülerdeki farklı özellikler bu katmanlar sayesinde algılanmaktadır. Farklı özelliklerin algılanması, görüntülerdeki nitelik belirlemede önemli bir rol oynamaktadır (Kayaalp ve Süzen, 2018).

CNN eğitiminde ileri yönlü ve geri yönlü olmak üzere iki aşama söz konusudur. İleri yönlü aşamada asıl amaç, giriş görüntüsünü her katmanda ağırlık ve sapma olan mevcut parametrelerle temsil etmektir. Sonrasında temel doğruluk değerleriyle kayıp maliyeti hesaplamak için tahmin çıktısı kullanılmaktadır. Geri yönlü aşamada ise, kayıp maliyetine dayalı olarak zincir kuralları kullanılarak her parametreye ait gradyan değeri hesaplanmaktadır. Tüm parametrelerdeki gradyan değerleri güncellenerek parametreler, bir sonraki ileri hesaplama için hazırlanmaktadır. Ağ öğreniminin durması ileri ve geri yönlü aşamaların yeterli tekrarına bağlıdır (Guo et al., 2016).

Bir CNN mimarisinde temel olarak evrişimsel katman (konvolüsyon katmanı), havuzlama katmanı ve tam bağlantı katmanı olmak üzere üç sinir katmanı

bulunmaktadır. Her katmanda giriş hacminin çıkış hacmine dönüştürüldüğü sinir hücresi aktivasyonu kullanılmaktadır (Voulodimos et al., 2018).

Şekil 2.1’de CNN mimarisinin genel yapısı gösterilmektedir.



Şekil 2.1. CNN Mimarisi (İnik ve Ülker, 2017).

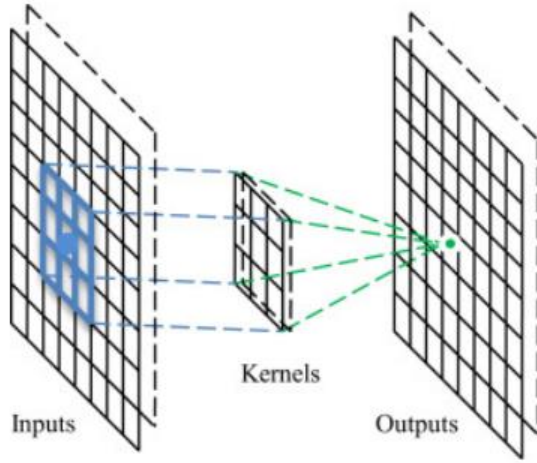
CNN’i oluşturan katmanlar:

Giriş katmanı

CNN’in ilk katmanıdır. Veriler ham olarak alınmaktadır. Bu katmana verilecek girdi görüntü boyutu; bellek ihtiyacı, eğitim süresi, görüntü başına düşen test süresi ve modelin başarısı açısından önem taşımaktadır (İnik ve Ülker, 2017). Verilerde boyutlandırma işlemleri yapılmışsa ya da boyut değişikliğine gereksinim yoksa veriler, girdi katmanı doğrudan ağıza verilebilmektedir. Fakat veriler üzerinde boyut değişikliği yapılacaksa yeniden boyutlandırma işlemleri bu katmanda gerçekleştirilmektedir (Metlek ve Çetiner, 2021).

Evrişim Katmanı

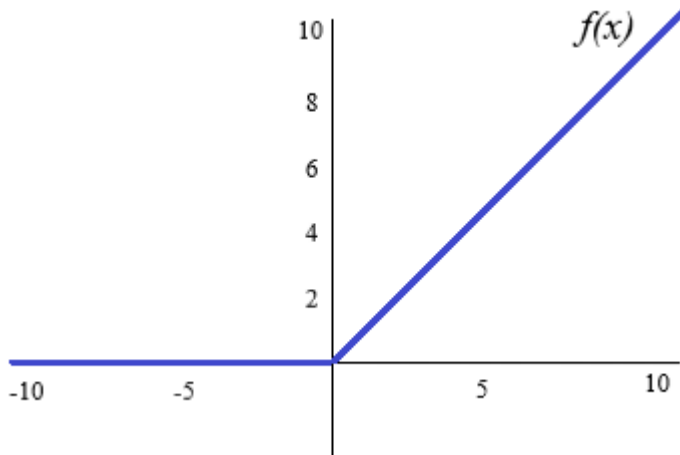
Dönüşüm katmanı olarak da bilinmektedir. CNN’in temelini oluşturmaktadır. Görüntü üzerinde filtre dolaştırılarak dönüşüm işlemi gerçekleştirilir. Filtreler farklı boyutlarda olabilir (2x2, 3x3, 5x5 gibi). Dönüşüm işlemi sonucunda her filtreye özgü niteliklerin ortaya çıktığı aktivasyon (özellik haritası) oluşturulmaktadır (Şekil 2.2) (Voulodimos et al., 2018).



Şekil 2.2. Evrişim Katmanında filtre dolaştırılması ile özellik haritası oluşturma işlemi (Guo et al., 2016).

Düzleştirilmiş Doğrusal Birim Katmanı (ReLU)

Aktifleştirme katmanı olarak da bilinmektedir. Evrişim katmanından sonra gelen bu katmanda, belirli matematiksel işlemleri yapıldığından doğrusal yapıda olan veri, doğrusal olmayan yapıya dönüştürülür. Giriş verisindeki negatif değerler sıfıra çekilmektedir (Şekil 2.3). Aktifleştirme katmanının kullanılması, ağın daha hızlı öğrenmesini sağlamaktadır (İnik ve Ülker, 2017).

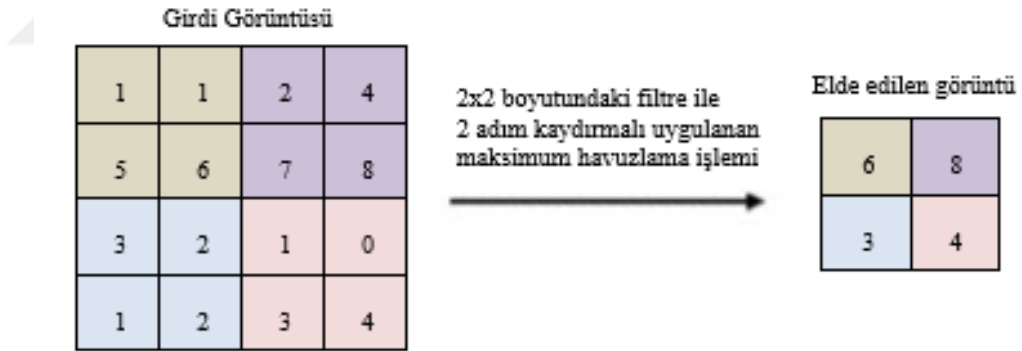


Şekil 2.3. Düzleştirilmiş doğrusal birim katmanının çıkış verisine etkisi (İnik ve Ülker, 2017).

Havuzlama Katmanı

Genellikle ReLu katmanından sonra gelen havuzlama katmanı, bir sonraki evrişim katmanı için girdi boyutunu azaltmakla görevlidir. Bu katmanda gerçekleştirilen işleme, Aşağı Örnekleme işlemi de denmektedir. Havuzlama katmanında gerçekleşen görüntü boyutundaki azalma, bilgi kaybına yol açmaktadır. Fakat bu kayıp ağ için sistemin ezberlemesini önlemek ve bir sonraki ağ katmanlarındaki hesaplama yükünü azaltmak açısından faydalıdır (Voulodimos et al., 2018).

Havuzlama katmanında da evrişim katmanındaki gibi görüntüler üzerinde filtreler gezdirilmektedir. Filtreler, görüntüler üzerindeki maksimum değeri ya da değerlerin ortalamasını alarak görüntü boyutunu azaltmaktadır (Şekil 2.4). Daha iyi sonuç vermesi nedeniyle genellikle maksimum havuzlama tercih sebebidir. Havuzlama katmanı isteğe bağlı olduğundan bazı mimarilerde bu katman kullanılmamaktadır (İnik ve Ülker, 2017).

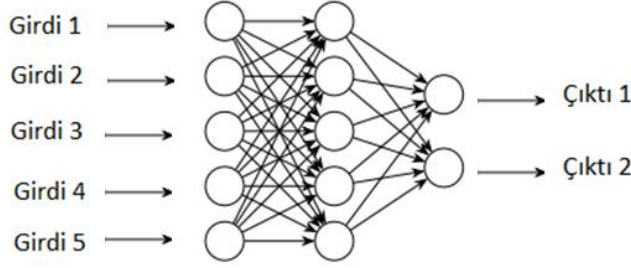


Şekil 2.4. Girdi görüntüsüne 2x2 boyutunda filtre ile iki adım kaymalı uygulanan maksimum havuzlama işlemi (Albawi et al., 2017).

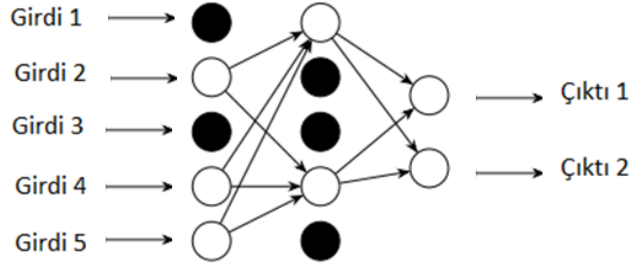
Seyrekleştirme Katmanı

CNN üzerinde büyük veriler ile işlem yapıldığı zaman ağ üzerinde ezberleme gerçekleşmesini önlemek amacıyla ağdaki bazı düğümlerin kaldırılması seyrekleştirme katmanının görevidir (Guo et al., 2016). Seyrekleştirme işleminden önceki ve seyrekleştirme uygulanmasından sonraki CNN ağ yapısının

karşılaştırmalı olarak gösterimi Şekil 2.5 ve 2.6’da sunulmaktadır (Metlek ve Çetiner, 2021).



Şekil 2.5. CNN ağına seyrekleştirme uygulanmadan önceki ağ yapısı (Labach et al., 2019).



Şekil 2.6. CNN ağına seyrekleştirme uygulandıktan sonra ağ yapısı (Labach et al., 2019).

Tam Bağlantı Katmanı

Sinir ağındaki üst düzey akıl yürütme işlemi tam bağlantı katmanı ile gerçekleştirilmektedir. Önceki katmanlardaki tüm aktivasyonlarla bağlantısı bulunmaktadır. Tam bağlantı katmanı, çok boyutlu özellik haritalarını tek boyutlu özellik vektörüne dönüştürmektedir. Oluşturulan özellik vektörü sınıflandırma işlemi için ya da sonraki işlemler için kullanılmaktadır (Voulodimos et al., 2018).

Sınıflandırma Katmanı

Tam bağlantı katmanından sonra gelerek sınıflandırma işlemini gerçekleştirmektedir. Katmandaki çıkış değeri, sınıf sayısına eşittir. Sınıflandırma işleminde her sınıf için 0-1 değer aralığında bir sayı üretilmektedir (Metlek ve Çetiner, 2021).

Geleneksel yapay sinir ağlarındaki sığ öğrenmeye kıyasla CNN’de daha soyut bilgiler öğrenilmesi avantaj oluştururken tam bağlantı katmanının çok sayıda parametre içermesi ezberleme sorununa yol açabilmektedir (Guo et al., 2016). CNN, daha çok sınıflandırma ve tespit işlemlerinde kullanılmakta, özellikle sınıflandırma işlemlerinde başarılı sonuçlar göstermektedir (Doğan ve Türkoğlu, 2019).

2.3. Topluluk Öğrenmesi ve Model Ortalaması

Topluluk öğrenmesi, farklı makine öğrenmesi algoritmalarını birleştirerek ortak bir görev oluşturmayı amaçlamaktadır. Her bir makine öğrenmesi algoritmasının tamamlayıcı bilgisi ile yeni bir model geliştirilmektedir. Amaç, modelin performansını iyileştirmek veya kötü bir modelin seçilme olasılığını azaltmaktır (Mahesh, 2020).

Klasik makine öğrenme algoritmaları, bilinmeyene yaklaşırken hipotez uzayı oluşturarak hangi hipotezin en iyi olduğunu belirlemek için bir öğrenme algoritması kullanılmaktadır. Eğitim veri değerlerinin hedef ile ne kadar eşleştiğini ya da tutarlı olduğunu değerlendirerek en iyi hipotezi belirlemektedir. Topluluk öğrenme algoritmaları, klasik makine öğrenme algoritmalarından farklı bir yaklaşıma sahiptir. En iyi hipotezi bulmak yerine, yeni veri etiketini tahmin etmek için oluşturdukları hipotezler üzerinde oylama yapmaktadırlar.

$\{h[1], \dots, h[k]\}$ hipotez kümesi, $\{w[1], \dots, w[k]\}$ ağırlık kümesi olmak üzere, oluşturulan hipotez kümesi ve seçilen ağırlık kümesini oluşturur ve bu değerleri Formül 2.1’te belirtildiği gibi kullanarak oylanan sınıflandırıcıyı oluşturur. Birleşik sınıflandırıcı H ’nin kararı, sıfıra eşit veya sıfırdan büyükse +1; değilse -1 değerini almaktadır (Dietterich, 2002).

$$H(\mathbf{x}) = w_1 h_1(\mathbf{x}) + \dots + w_K h_K(\mathbf{x}) \quad (2.1)$$

Topluluk öğrenmesinin avantajları:

- Ezberden kaçınma: Öğrenme algoritmaları, veri miktarı az olduğunda eğitimde olmayan veriler üzerinde zayıf tahminlemeler yapan, eğitim verilerinde ise bunun aksine güçlü tahminlemeler yapan hipotezler seçmeye meyillidir. Topluluk öğrenmesi, farklı hipotezlerin ortalamasını alarak yanlış bir hipotez seçme riskini azaltmaktadır. Böylece genel tahmin performansını iyileştirmektedir.
- Hesaplama avantajı: Birkaç öğrencinin birleştirilmesiyle elde edilen topluluk öğrenmesi, yerel minimumda takılma riskini azaltmaktadır.
- Temsil: Farklı modellerin birleştirilmesi sayesinde arama alanı genişlemekte ve veri alanına daha iyi bir uyum sağlanabilmektedir (Sagi and Rokach, 2018).

Topluluk öğrenmesi kapsamında tez çalışmasında kullanılan model ortalaması yöntemi hakkında sonraki paragraflarda bilgiler verilmiştir.

Topluluk öğrenmesi yöntemi, yeni veriler üzerinde tahminleme yapılması amacıyla kullanılmaktadır. Birden fazla eğitilmiş modellerin, test verileri üzerinde bulunduğu tahminlerin ortalaması alınarak bir sonuç üretilmektedir. Model ortalaması yönteminde amaç, model üzerindeki varyansı azaltarak performansındaki güvenilirliği artırmaktır.

D verisi üzerinde ilgili parametrelerdeki sonsal dağılımı β olan, $M = \{ M_1, M_2, \dots, M_K \}$ gibi eğitilmiş modeller dizisinin ağırlıklı ortalaması Formül 2.2 ile elde edilmektedir.

$$Pr(\beta|D) = \sum_{j=1}^K Pr(\beta|M_j, D) Pr(M_j|D) \quad (2.2)$$

Model ortalaması yönteminde yeniden bir eğitim yapılmasına gerek yoktur. Veriler üzerinde yapılan testlerde, eğitilmiş olan modellerin tahmin değerleri birleştirilmektedir. Bu yöntemle bazı modellerdeki en iyi sonuçlar göz ardı edilse de daha doğru yayılımda bir çıktı üretilmektedir (Amini and Parmeter, 2012; Brownlee, 2018).

2.4. Nirengi Noktası Bulma Yöntemi ve MTCNN

Görüntü üzerinde yüze ait önemli noktaların konumlarını bulmak için kullanılmaktadır. Yüz üzerindeki göz, ağız, kaş, burun, çene gibi nirengi noktalarının konumlarını otomatik olarak tespit edilmesini amaçlayan bir yöntemdir (Mahpod et al., 2021). Yüz doğrulama, yüz çıkarımı, yüz tanıma gibi birçok alanda bu yöntemden yararlanılmaktadır. Yüz ifadelerindeki farklılıklar, çevresel koşullar gibi etkenler yüz nirengi noktası tespitini zorlaştırmaktadır (Wu and Ji, 2019).

Nirengi noktası bulma yöntemlerinden tez çalışmasında kullanılan MTCNN yöntemi hakkında sonraki paragraflarda bilgiler verilmiştir.

MTCNN, yüz algılama ve yüz hizalama işlemlerini gerçekleştirmek için geliştirilen bir çerçevedir (Zhang et al., 2016). Üç adet CNN ağı içermektedir. Yapısında bulunan birleşik basamaklı CNN'leri kullanarak çok görevli öğrenim gerçekleştirmektedir (Li et al., 2020).

MTCNN ağı, yüz sınıflandırması, sınırlayıcı kutu entegrasyonu ve yüz yer işareti lokalizasyonu olmak üzere üç çıktıya sahiptir.

Yüz sınıflandırılmasında, çapraz entropi kaybı Formül 2.3 kullanılarak ikili sınıflandırma gerçekleştirilmektedir.

$$L_i^{\text{det}} = -(y_i^{\text{det}} \log(p_i) + (1 - y_i^{\text{det}})(1 - \log(p_i))) \quad (2.3)$$

Sınırlayıcı kutu regresyonunda, Öklid kaybı Formül 2.4 kullanılarak her aday pencere için aday ile en yakın gerçek arasındaki fark hesaplanmaktadır.

$$L_i^{\text{box}} = \|\hat{y}_i^{\text{box}} - y_i^{\text{box}}\|_2^2 \quad (2.4)$$

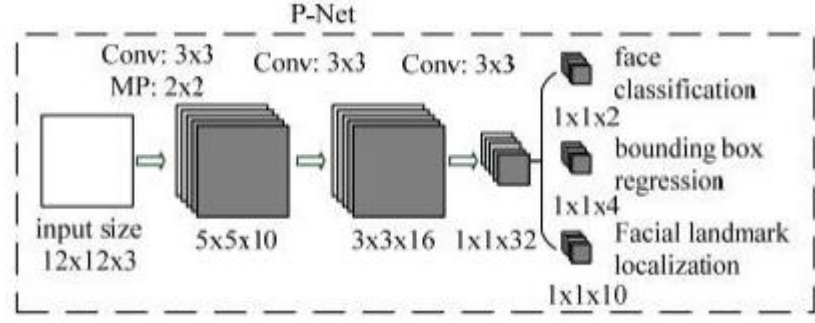
Yüz yer işareti lokalizasyonunda ise, sol göz, sağ göz, burun, sol ve sağ ağız köşesi belirlenmektedir (Zhang et al., 2016).

MTCNN ağında süreç, evrişimsel ağların üç aşamasından oluşmaktadır. İlk aşamada sığ bir CNN kullanılarak aday pencereler üretilmektedir. Sonraki aşamada, önerilen aday pencereler işlenmektedir. Son aşamada ise yüzdeki kilit noktaları çıkarılmakta ve iyileştirmeler gerçekleştirilmektedir. İlk aşamada daha basit bir CNN mimarisi bulunurken sonraki aşamalardaki ağlar bir önceki aşamadakinden daha karmaşık bir yapıdadır (Ji et al., 2019).

MTCNN ağ yapısını oluşturan ağlar: Teklif Ağı (P-Net), İyileştirme Ağı (R-Net) ve Çıktı Ağı (O-Net) olmak üzere üç alt ağdan oluşmaktadır (Li et al., 2020).

Teklif Ağı (P-Net)

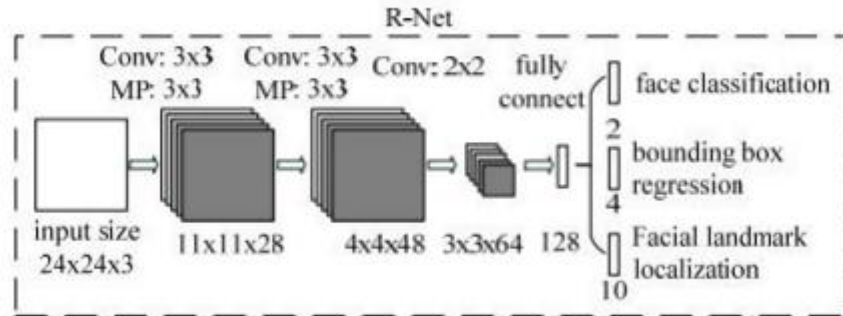
MTCNN ağının ilk aşaması olan P-Net ağı, tam evrişimli bir ağıdır (Zhang et al., 2016). Orijinal görüntüyü birden çok çözünürlükte işlemektedir. Aday pencereler, sınır kutusunun regresyon vektörü tarafından kalibre edilmektedir. Kalibre işleminden sonra örtüşen aday çerçeveler Maksimum olmayan bastırma (NMS) algoritması tarafından filtrelenmektedir (Khabarlak and Koriashkina, 2021). Filtreleme işleminden sonra çok sayıda kaba yüz bölgelerinin sınır kutuları tahminlerini vermektedir. Çok ölçekli hedef sorununu çözmek amacıyla farklı boyutlardaki görüntüler kademeli ağına gönderilmektedir (Ji et al., 2019). Sınırlayıcı kutu vektörleri elde edildikten sonra örtüşen bölgelerin birleştirilmesi amacıyla iyileştirmeler yapılmaktadır. Yapılan iyileştirmelerden sonra çıktı, tüm aday pencerelerdir (Zhang et al., 2016). Ağ yapısı Şekil 2.7’de gösterilmektedir.



Şekil 2.7. P-Net Ağ Yapısı (Shams and Nasim, 2019).

İyileştirme Ağı (R-Net)

R-Net ağı, P-Net ağı tarafından oluşturulan tüm aday pencerelerden beslenmektedir. Bu ağda aday sayısı daha da azaltılarak sınırlayıcı kutu regresyonu ile kalibrasyon gerçekleştirilmektedir. Kalibrasyon aşamasından sonra örtüşen aday pencereler, NMS algoritması ile birleştirilmektedir (Zhang et al., 2016). P-Net ağına kıyasla R-Net ağ yapısı, yanlış pozitif örnekleri daha iyi dışlayabilmektedir (Ji et al., 2019). Ayrıca, tüm görüntüyü yeniden işlemeksizin öngörülen kutuları iyileştirmesi sebebiyle hesaplama süresi bakımından tasarruf sağlamaktadır (Khabarлак and Koriashkina, 2021). Ağ yapısı Şekil 2.8’de gösterilmektedir.

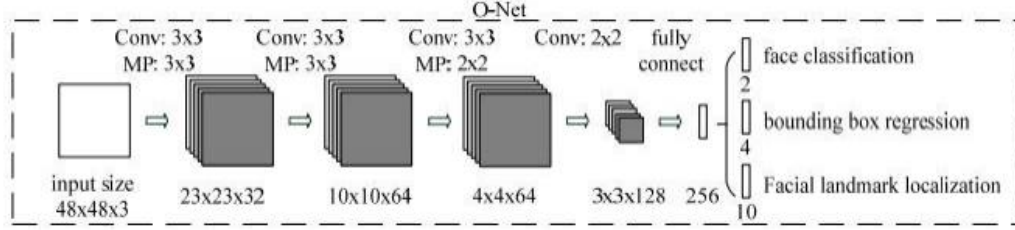


Şekil 2.8. R-Net Ağ Yapısı (Shams and Nasim, 2019).

Çıktı Ağı (O-Net)

MTCNN mimarisindeki en yavaş ağıdır (Khabarлак and Koriashkina, 2021). O-Net ağında gerçekleşen işlemler, R-Net ağındaki işlemlerle benzerlik

göstermektedir (Zhang et al., 2016). R-Net'ten farklı olarak bu ağda, R-Net tarafından filtrelenen yüz aday pencereleri ayrıca filtrelenmekte ve her bir yüz bölgesindeki kilit noktaların konumları belirlenmektedir (Ji et al., 2019). MTCNN ağındaki son düzeltmeler yapılmaktadır (Khabarлак and Koriashkina, 2021). Ağ yapısı Şekil 2.9'da gösterilmektedir.



Şekil 2.9. O-Net Ağ Yapısı (Shams and Nasim, 2019).

3. ÖNCEKİ ÇALIŞMALAR

Trafik kazalarına neden olan insan faktörü üzerinde, sürücülerin yorgunluk/uykusuzluğuna bağlı kazaların önlenmesi amaçlı literatürde pek çok çalışma yer almaktadır. Bu alandaki çalışmalarda en çok göz, yüz ve esneme parametreleri üzerinden yorgunluk/uykusuzluk durumu tespit edilmektedir.

Choi et al. (2016) yaptıkları çalışmada, sürücünün baktığı yöne bağlı olarak uyku olup olmadığını tespit etmek için, öncelikle Haar ve MOSSE takipçisi ile sürücünün yüzünü belirlemektedirler. Evrişimsel Sinir Ağı (CNN) kullanarak geliştirdikleri yöntemle sürücünün, sürüş süresince baktığı bölgeler sınıflandırılmaktadır. Çalışmada kendi oluşturdukları veri seti kullanılarak ortalama %95 başarı elde edilmiştir.

Shakeel et al. (2019) yaptıkları çalışmada, sürücünün gözlerinin açık ya da kapalı olmasına bakarak Android ortamında MobileNET ile gerçekleştirmektedirler. Video akışı üzerinden sürücünün yüz ve göz bölgeleri tespit edilmektedir. Eğer sürücünün gözlerinin ikisi de kapalıysa sürücü sesli olarak uyarılmaktadır. Dört farklı veri seti ile yapılan bu çalışmanın başarısı %84'tür.

Yalnızca göz görüntüleri üzerinden Evrişimsel Sinir Ağı kullanılarak gerçekleştirilen Tümen vd.'nin (2018) yaptığı çalışmada CEW veri seti kullanılarak sürücünün uykusuzluğu tespit edilmektedir. Sistemin, veri setindeki dört sınıfı (sol göz kapalı, sol göz açık, sağ göz kapalı, sağ göz açık) tanımadaki başarısı %92,99-%96,5 düzeyinde ölçülmüştür.

Yüzün belirleyici noktaları kullanılarak Jabbar et al.'ın (2018) çalışmalarında, video görüntüleri üzerinden çıkartılan karelerde sürücünün yüzü tespit edilmektedir. Evrişimsel Sinir Ağları kullanılan bu çalışma, Android ortamında sürücünün uykusuzluğunu tespit etmeye yöneliktir. NTHU veri seti kullanılan sistemde %81 düzeyinde başarı elde edilmiştir.

Uykusuzluk/yorgunluğu esneme üzerinden tespit eden Xie et al. (2018) çalışmalarında, Evrişimsel Sinir Ağı ve LSTM (Uzun Kısa Süreli Bellek)

modellerini birlikte kullanmışlardır. YawDD, NTHU-DDD ve özel olarak düzenlenen esneme veri setleri üzerinde yapılan deneyler sonucunda eğitim verileri üzerinde doğruluk değeri %100, test verileri üzerinde doğruluk ve duyarlılık değerleri sırasıyla %98,4 ve %97,36 olarak elde edilmiştir.

Literatürde yorgunluk/uykusuzluğun belirlenmesi için birden fazla özelliğin bir arada değerlendirildiği çalışmaların genelde son birkaç yılda yapıldığı görülmektedir.

Dwivedi et al. (2014), sürücünün uykululuk durumunu belirleyip uyarmak için Evrişimsel Sinir Ağı kullanarak yüze ait sanal öznitelikler belirlemektelerdir ve sürücünün durumunu uykulu ya da uykulu değil olarak sınıflandırmaktadırlar. Video görüntülerinden elde edilen kareler üzerinde yüz tespiti yapıp Temsili Öğrenme kullandıkları bu çalışmada eğitim verilerinde doğrulama başarı düzeyi %92,33 iken test verisinde ortalama başarı düzeyi %88 olarak ölçülmüştür.

Yadawadkar et al. (2018) yaptıkları çalışmada istatistiksel öznitelik çıkarımı içeren Uzun Kısa Süreli Bellek kullanılmaktadır ve Evrişimsel Sinir Ağları ile video sınıflandırılması gerçekleştirilmektedir. SHRP2NDS veri seti kullanılarak yapılan bu çalışmada %70 başarı elde edilmiştir.

Celona et al. (2018) gerçekleştirdikleri çalışmada, sürücünün uykusuzluğunu tespit etmek için baş pozisyonu, gözlerinin ve ağzının durumunu analiz ederek sürücünün uykusuzluk düzeyini belirlemektedirler. Deneysel çalışmalar NTHU-DDD veri seti üzerinde gerçekleştirilmiştir ve %75,73 başarı elde edilmiştir. Daha sonra sistem üzerinde yapılan iyileştirme çalışmaları sonucunda başarı düzeyi %83,44'e ulaşmıştır.

Liu et al. (2019) yaptıkları çalışmada yorgunluğu tespit etmek amacıyla göz ve ağız görüntülerini birlikte değerlendirmektedirler. Görüntü özellikleri statik ve dinamik olarak çıkartılarak Evrişimsel Sinir Ağı ile geliştirilen sistem, NTHU - DDD veri seti üzerinde %97,06 başarı düzeyi ile sonuçlanmıştır.

Laouz (2019) tarafından sürücünün hareketlerine göre gerçek zamanlı olarak araç kamerası üzerinden yorgunluk/uykusuzluğunu tespit eden sistem tasarlanmıştır. Bu amaçla sürücü görüntüsünden yüz, göz ve ağız bölgeleri 68 noktalı nirengi noktası bulma yöntemi kullanılarak belirlenmektedir. Göz görüntüleri 3 evrişim katmanlı CNN modeli ile eğitilmiştir. Öklid uzaklığına bağlı olarak hesaplama yapan MAR hesaplaması kullanılarak esneme tespiti gerçekleştirilmektedir. Göz görüntüsünde gözün kapalı / açık olma durumu ile ağız görüntüsünde ağzın geniş açık ve normal olma durumuna göre uyanık, esneyen ve uyuyan olarak sınıflandırma yapılmaktadır. MRL Eyes ve YawDD veri setleri kullanılarak gerçekleştirilen bu çalışmada, gözlerin açık olması ile esneme olmaması uyanıklık durumu olarak; esneme oranı ile göz kırpma oranı olarak esneme durumu ve gözlerin uzun süreli kapalı kalması ile başı çokça oynatmama hali olarak da uyuma durumu şeklinde değerlendirme yapılmaktadır.

Ji et al. (2019) tarafından yapılan bu çalışmada yorgunluk durumu tespiti için Cew, YawDD ve kendi oluşturdukları veri setleri olmak üzere 3 farklı veri seti kullanılmaktadır. MTCNN yöntemi ile görüntüler üzerinde yüz, gözler ve ağız bölgeleri belirlenmiştir. Göz görüntülerini eğitmek için ESR-Net olarak adlandırdıkları bir CNN ağı ve ağız görüntülerini eğitmek için tasarladıkları MSR-Net olarak adlandırdıkları bir CNN ağı tasarlamışlardır. Görüntüler ESR Net'te, kapalı göz ve açık göz; MSR-Net'te ise kapalı ağız ve açık ağız olarak sınıflandırılmaktadır. Çalışma kapsamında gözlerin 1 saniyeden fazla kapalı kalması tehlikeli, saniye başı göz kapanma süresinin 0,15 saniyeden büyük veya ağzın 4 saniye boyunca açık olması uyarılmalı ve diğer durumlar normal olarak nitelendirilmektedir. ESR-Net üzerindeki test başarısı Cew veri setinde %97,87; YawDD veri setinde %98,2 ve kendi oluşturdukları veri setinde %97,17 olmak üzere göz veri setindeki test ortalama başarısı %98,17 doğruluk değeri göstermiştir. MSR-Net üzerindeki test başarısı ise YawDD veri setinde %97,4 ve kendi oluşturdukları veri setinde %98,13 olmak üzere ağız veri setindeki test ortalama başarısı %97,77 doğruluk değeri göstermiştir. Veri setleri üzerinde çalışmalarda doğrulama ve test başarısı olarak; göz veri seti üzerindeki ortalama başarısı %98,42; ağız veri seti üzerindeki başarısı %97,93 doğruluk göstermiştir.

Kameradan alınan görüntü kareleri üzerinden uykusuzluk tespit eden Alparslan et al. (2020) çalışmalarında Eye Blink ve Cew olmak üzere iki farklı veri seti kullanılmaktadır. Eye Blink veri seti ile göz modeli, Cew veri seti ile yüz modeli geliştirmişlerdir. Göz modelinde %95,84; yüz modelinde %80,01 başarı elde edilmiştir. İyileştirme amaçlı yüz modeli üzerinde rakipli eğitim ve veri artırımı ve çoğaltma kullanarak başarı %6 artırılmıştır. Göz modelinde ise gürültü eklenerek başarıyı artırma amaçlı çalışmalar gerçekleştirilmiştir. Gürültü kullanılarak ve gürültü kullanılmadan gerçekleştirilen yüz ve göz model eğitimlerinden sonra göz modeli, test verileri üzerinde %94 başarı gösterirken yüz modeli %2 daha az başarı göstermiştir. Gürültü eklenen yüz model eğitiminde %4-%16 kötüleşme gösterirken test verisinde %3-%13 kötüleşme göstermiş fakat göz kapağı algılamasında duyarlık değerinde %5 artış göstermiştir.

Shen et al. (2020) tarafından önerilen sistem dört aşamadan oluşmaktadır. İlk aşamada, Optical Flow yöntemi ile sürücü görüntüsü üzerinde önemli bölgeler belirlenmektedir. İkinci aşamada, CLAHE Histogram eşitleme yöntemi ile farklı aydınlatma koşullarının etkisi önlenmeye çalışılmaktadır. Üçüncü aşamada, her bir özellik için iki akışın birleştirildiği tasarım gerçekleştirilmektedir. Son aşamada ise üç alt ağ yapısı birleştirilerek yüz, göz ve ağız birlikte eğitilmekte ve modelin tam bağlantı katmanında sürücünün durumu değerlendirilmektedir. NTHU-DDD video veri seti üzerinde yapılan bu çalışmada %94,46 başarı elde edilmiştir.

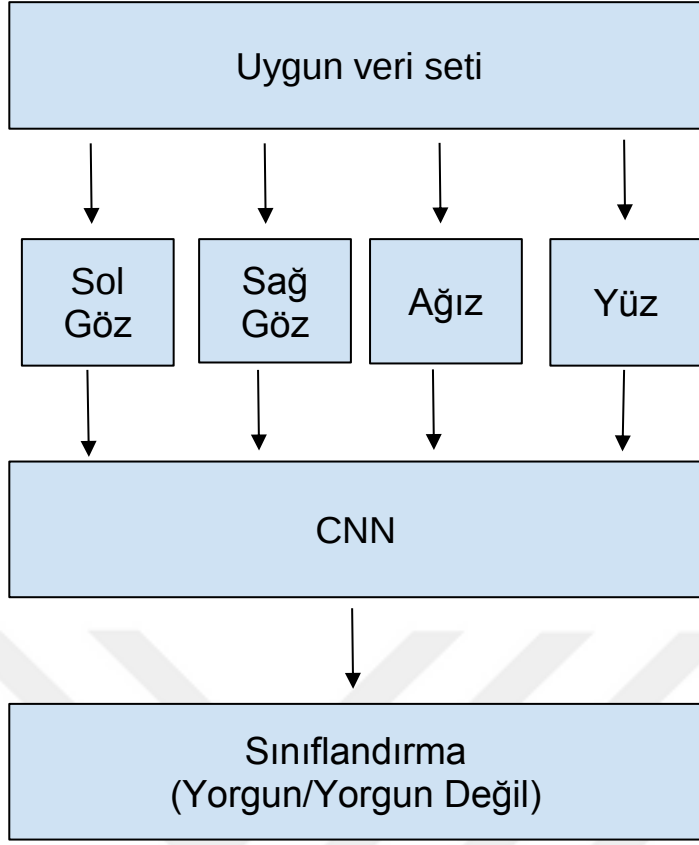
4. DERİN ÖĞRENME KULLANILARAK GÖZ, YÜZ VE ESNEME ÖZELLİKLERİNE DAYALI MELEZ BİR YORGUNLUK TANIMA SİSTEMİ

Bu bölümde, tez çalışmasında göz, yüz ve esneme özelliklerine dayalı olarak tasarlanan yorgunluk tanıma sistemi ve modelleri anlatılmaktadır.

4.1. Giriş

Tez çalışması süresince gerçekleştirilen sistem, derin öğrenme ve topluluk öğrenme yöntemleri kullanılarak kişilerin yorgunluk/uykusuzluğuna bağlı oluşabilecek sorunların önlenmesi amacıyla geliştirilmiştir. Yorgunluk belirtilerinden olan esneme, gözlerin kapanması veya kapalıya yakın olmasından faydalanılarak durağan görüntüler üzerinde çalışma yapılmıştır. Geliştirilen sistemde gri tonlamalı görüntüler kullanılmıştır.

Yorgunluk tanıma alanında yapılmış olan önceki çalışmalardan farklı olarak, hem durağan görüntüler üzerinde çalışılmış hem de CNN mimarisine dayalı derin öğrenme modeline girdi olarak farklı görüntü parçaları verilip bu görüntü parçalarının öznelikleri birleştirilmiştir. Farklı türden girdilerin durağan görüntüler üzerinde, aynı modelde eğitilip sınıflandırılması ile alana özgü çalışmalara farklı bir açıdan yaklaşılarak literatüre yeni bir yorgunluk tanıma yöntemi kazandırılmıştır. Tez çalışması kapsamında önerilen genel sistemin akış şeması Şekil 4.1’de sunulmuştur.



Şekil 4.1. Önerilen Sistemin Akış Şeması.

Bu bölümde tez çalışması kapsamında izlenen prosedüre yer verilmektedir. İlk olarak literatürde yorgunluk/uykusuzluk tespiti üzerine yapılan çalışmalar incelenmiş ve uygun veri setleri tespit edilmiştir. Yüz, göz özelliklerine dayalı 1 veri seti (Cew), yüz ve göz ve ağız özelliklerine dayalı 1 veri seti (YawDD) olmak üzere toplamda 2 veri seti belirlenmiştir. İkinci adımda yüz, göz ve ağız özelliklerine dayalı olan veri seti üzerinde MTCNN yöntemi kullanılarak yüz görüntüsü üzerinde beş belirleyici noktadan (sol göz, sağ göz, burun ucu, sol ve sağ kenar ağız ucu) yararlanılarak yüz, gözler ve ağız görüntüleri ayrıştırılmış ve tez çalışmasında planlanan şekilde kullanılabilir hale getirilmiştir. Son adımda, belirlenen veri setleri (Cew ve YawDD) üzerinde 3 farklı deneysel çalışma grubu gerçekleştirilmiştir:

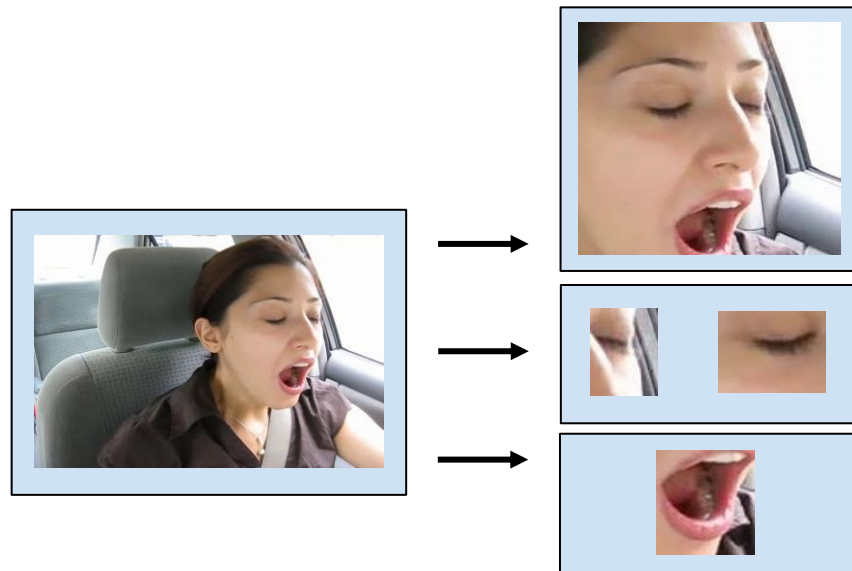
- 1) Yüz, göz ve ağız görüntü parçalarının her biri için sıralı model üzerinde ayrı başarımlar elde edilmiştir.

- 2) Aynı yüz görüntüsüne ait yüz bölümlerinin girdi olarak verilip öz niteliklerinin birleştirildiği fonksiyonel model üzerinde deneysel sonuçlar elde edilmiştir.
- 3) Model ortalaması topluluk öğrenmesi yöntemi kullanılarak farklı veri setlerinin benzer özelliklerinin birleştirilmesi ile deneysel sonuçlar elde edilmiştir.

4.2. Veri Setlerinin Düzenlenmesi

Tez çalışmasında, yorgunluk belirtilerinin en belirgin özellikleri olarak dikkate alınan esneme ve gözlerin kapalı/açık olma durumlarına bağlı olarak yorgunluk tespitinde gözlerin ve ağzın kapalı/açık olma durumlarından yararlanılmıştır. Her bir bölgenin özelliğini dikkate alabilmek amaçlı olarak belirlenen veri seti (YawDD) üzerinde MTCNN kullanılarak yüz, göz ve ağız bölgeleri belirlenmiştir. Belirlenen her bir bölgenin görüntüsü ilgili sınıfa ait olarak etiketlenerek yüz, göz ve ağız veri setleri oluşturulmuştur.

Yüz, göz ve esneme özelliklerinden yararlanılan Raju (2020) tarafından düzenlenen YawDD veri setinde video kareleri üzerinden ayrıştırılan yüz, göz ve ağız bölgeleri Şekil 4.2’de bir örnek üzerinde gösterilmektedir.

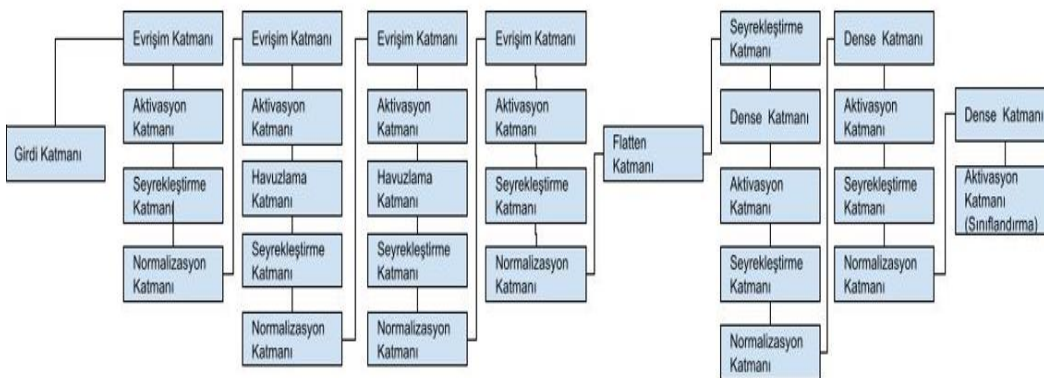


Şekil 4.2. YawDD veri seti üzerinde ayrıştırılan yüz, göz ve ağız bölgeleri.

4.3. CNN Mimarisine Dayalı Sınıflandırma Modeli

Yorgunluk/uykusuzluk tespitinde belirleyici olan özellikler gözlerin kapanması ve esneme gibi özelliklerdir. Derin öğrenme kapsamındaki CNN mimarisi, veriler üzerinden özellikleri öğrenmeye ve veriler arasında farklı özelliklerin birleştirilmesine olanak sağlamaktadır.

Bu tez çalışması kapsamında CNN mimarisi kullanılarak derin öğrenme modelleri geliştirilmiştir. Sequential (sıralı) ve functional (fonksiyonel) olarak farklı derin öğrenme modelleri tasarlanmıştır. Tasarlanan sıralı modellerde 4 adet evrişim katmanı, 2 adet tam bağlantı katmanı, 1 adet çıktı katmanı kullanılmaktadır (Şekil 4.3). Veri setlerinden alınan yüz, sol göz ve sağ göz özelliklerine ait katmanlar birleştirilerek tasarlanan CNN mimarisine dayalı fonksiyonel yapıdaki derin öğrenme modelinde 12 adet evrişim katmanı, 2 adet tam bağlantı katmanı ve 1 adet çıktı katmanı kullanılmaktadır (Şekil 4.4). Geliştirilen diğer bir fonksiyonel modelde ise veri setinden alınan yüz, sol göz, sağ göz ve ağız özelliklerine ait katmanlar birleştirilerek 16 adet evrişim katmanı, 2 adet tam bağlantı katmanı ve 1 adet çıktı katmanı kullanılmaktadır (Şekil 4.5). Tasarlanan modellerde girdi katmanına parametre olarak veri setlerindeki görüntü boyutları verilmiştir. Çıktı katmanında ise sınıflandırma yapılacağından ve sınıf sayısı kadar yapay sinir hücresi bulunması gerektiğinden, çıktı katmanına parametre olarak sınıf sayısı sunulmuştur.



Şekil 4.3. Çalışma kapsamında tasarlanan sıralı derin öğrenme modeli.



Şekil 4.4. Çalışma kapsamında yüz, sol göz, sağ göz özneliklerinin birleştirildiği fonksiyonel derin öğrenme modeli.



Şekil 4.5. Çalışma kapsamında yüz, sol göz, sağ göz ve ağız öz niteliklerinin birleştirildiği fonksiyonel derin öğrenme modeli.

Evrişim katmanlarında, özelliklerin öğrenilmesi için 3x3'lük filtreler kullanılmaktadır. Bu öğrenme işlemi için her bir katman üzerinde filtre, görüntünün tamamında dolaştırılmaktadır. Filtredeki ağırlıklar, piksel değerleri ile çarpılıp toplanmaktadır. Bu işlemler sonucunda görüntüde tespit edilen özelliklerde aktif olan filtreler öğrenilmektedir. Her katmanda öğrenilen özellikler hiyerarşik yapıdadır. Temel özellikler ilk katmanlarda, alt seviye özellikler ve bunlara dayalı üst seviye özellikler ise daha derin katmanlarda öğrenilmektedir.

Evrişim katmanlarının ardından parametre sayısını küçültmek amaçlı havuzlama katmanı kullanılmaktadır. Bu katmanda görüntü üzerinde 2x2'lik filtrelerin gezdirilmektedir. Filtre kullanımı ile parametre sayısı azaltılmakla

birlikte girdilerdeki önemli özelliklerin de korunması amacıyla maksimum havuzlama katmanı kullanılmaktadır.

Tasarlanan modellerde kullanılan aktivasyon fonksiyonu olarak ReLu aktivasyon fonksiyonu seçilmiştir. Tercih edilen fonksiyon, negatif değerleri sıfıra eşitler, hızlı çalışır ve türevi kolay alınır. Bu tez çalışmasında sınıflandırma yapıldığı için çıktı katmanında, verilen her bir girdinin bir sınıfa ait olma olasılığını gösteren softmax aktivasyon fonksiyonu tercih edilmiştir.

Geliştirilen derin öğrenme modellerinde, evrişim katmanlarından elde edilen özelliklerin tam bağlantı katmanına girdi olarak verilebilmesi için flatten katmanı kullanılmaktadır. Flatten katmanı çok boyutlu özellik haritalarını tek boyutlu vektörlere dönüştürerek özelliklerin tam bağlantı katmanına uygun girdi formatına dönüştürülmesini sağlamaktadır.

Düğüm seyrekleştirme katmanında, eğitim esnasında ezberlemeyi önlemek amacıyla eğitim öncesinde, katmanlarda rastgele seçilen nöronların bağları koparılır. Koparılan bağların ağırlıkları eğitimden etkilenmez. Eğitimden sonra orijinal ağırlıklar korunduğundan test sonuçlarının alınmasında bu ağırlık değerleri kullanılmaktadır. Bu sayede düğümler eğitim aşamasında birbirlerinin ağırlık değişimlerinden daha az etkilenmekte ve belli bir yüzdedeki ağırlık verilerden etkilenmediği için ağırlık genelleme yeteneği artmaktadır. Bu amaçla tasarlanan modellerde, eşik değeri 0,2 olarak verilen düğüm seyrekleştirme (dropout) katmanı kullanılmaktadır.

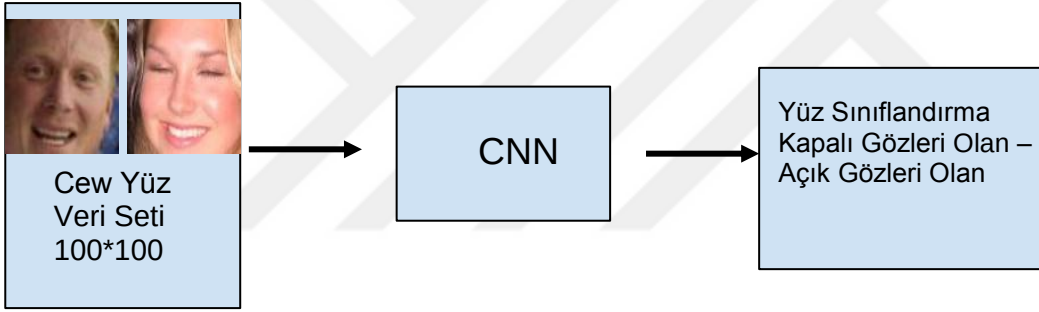
Oluşturulan sıralı modelde parçaların öznelilikleri birleştirilip evrişim, havuzlama, aktivasyon katmanlarından geçirilip tam bağlantı katmanına girdi olarak verilmekte ve sınıflandırılmaktadır. Fonksiyonel modellerde ise, veri setinin tüm özellikleri ayrı birer girdi olarak verilmektedir. Her bir girdi özelliği ayrı ayrı evrişim, havuzlama, aktivasyon katmanlarından geçirildikten sonra birleştirilip tam bağlantı katmanına girdi olarak sunulmaktadır.

İki gizli katman içeren tam bağlantı katmanı ve bir tam bağlantılı katman türü olan çıktı katmanı ile model sonlandırılmıştır. Gizli katmanlar sırasıyla 512 ve 128

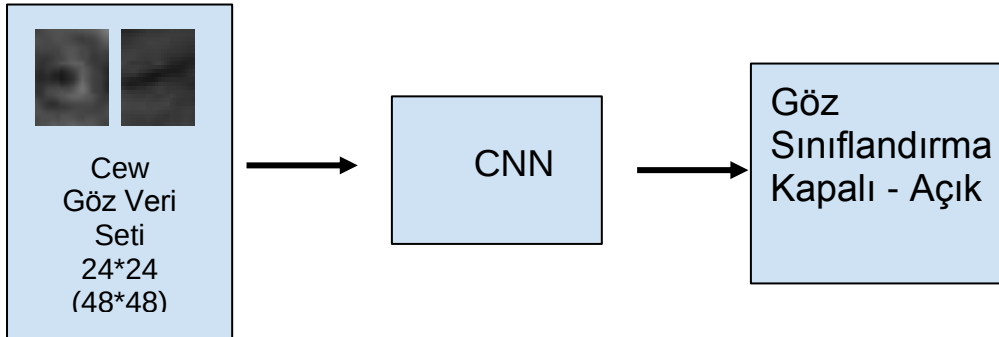
yapay sinir hücresi içermektedir. Çıktı katmanında ise 2 sınıf kullanıldığından 2 yapay sinir hücresi bulunmaktadır.

Gerçekleştirilen sıralı ve fonksiyonel modellerin gösterildiği Şekil 4.6, Şekil 4.7 ve Şekil 4.11’de yer alan Cew veri setine ait görüntüler, Song et al. (2014) tarafından düzenlenen veri setinden alınmıştır. Şekil 4.8 – Şekil 4.10 arasındaki şekillerde ve Şekil 4.12’de yer alan YawDD veri setine ait görüntüler, Raju (2020) tarafından düzenlenen veri setinden alınmıştır.

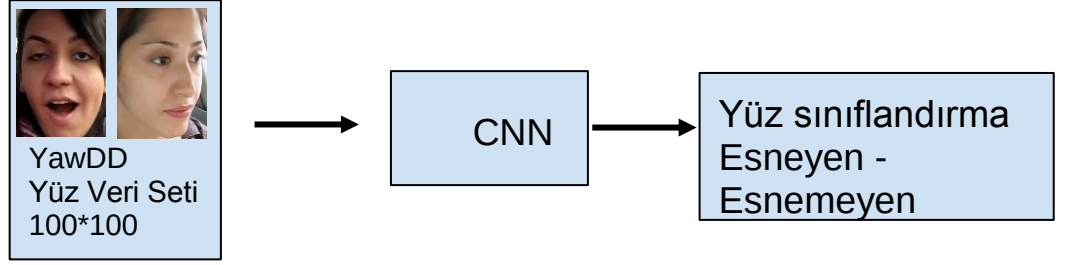
Cew ve YawDD veri setinde tasarlanan sıralı modeller; Cew veri setinde yüz görüntüleri için Şekil 4.6, göz görüntüleri için Şekil 4.7, YawDD veri setinde yüz görüntüleri için Şekil 4.8, göz görüntüleri için 4.9, ağız görüntüleri için 4.10’da sunulmuştur.



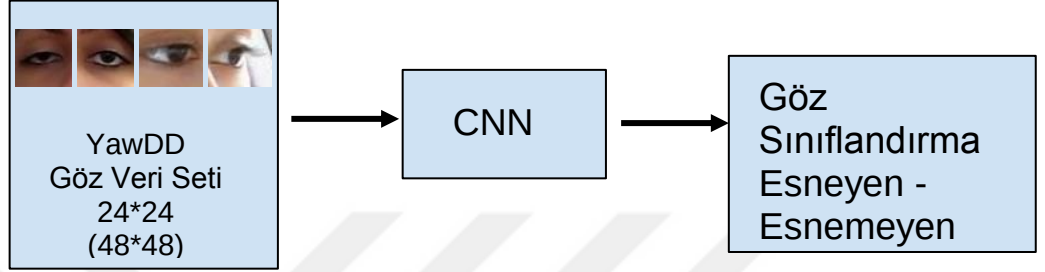
Şekil 4.6. Cew veri setinde yüz görüntü parçasına dayalı sıralı derin öğrenme modeli.



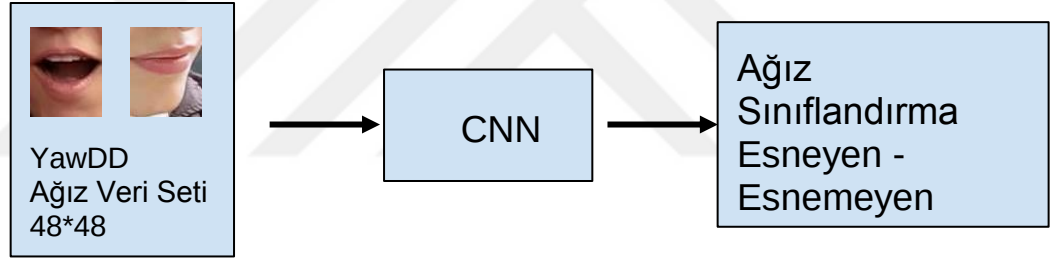
Şekil 4.7. Cew veri setinde göz görüntü parçasına dayalı sıralı derin öğrenme modeli.



Şekil 4.8. YawDD veri setinde yüz görüntü parçasına dayalı sıralı derin öğrenme modeli.

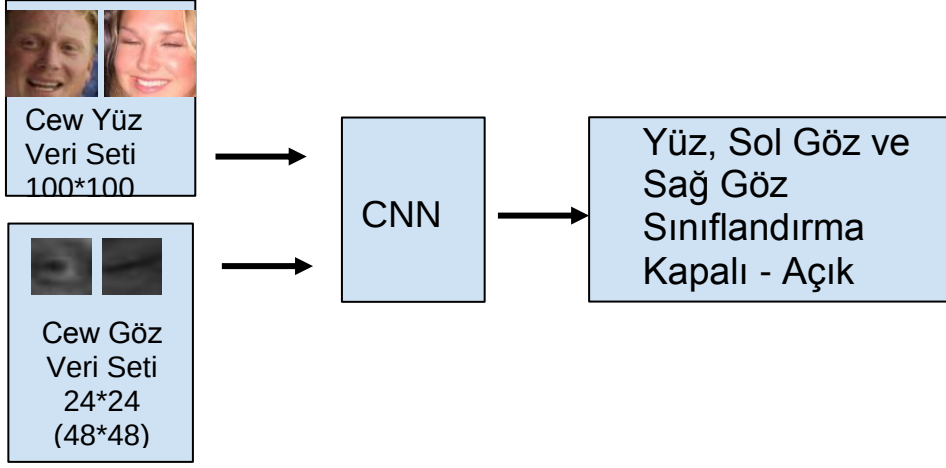


Şekil 4.9. YawDD veri setinde göz görüntü parçasına dayalı sıralı derin öğrenme modeli.

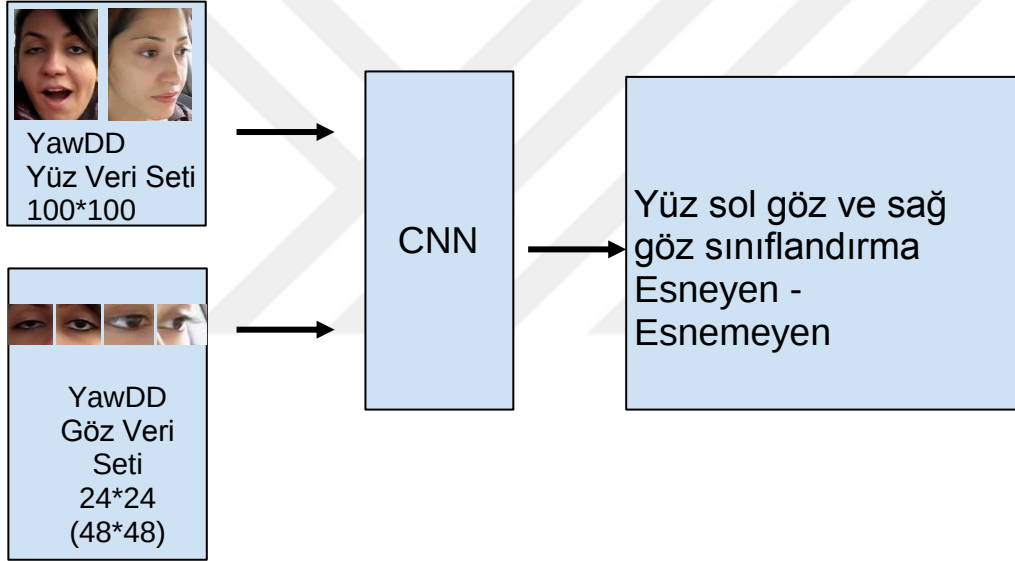


Şekil 4.10. YawDD veri setinde ağız görüntü parçasına dayalı sıralı derin öğrenme modeli.

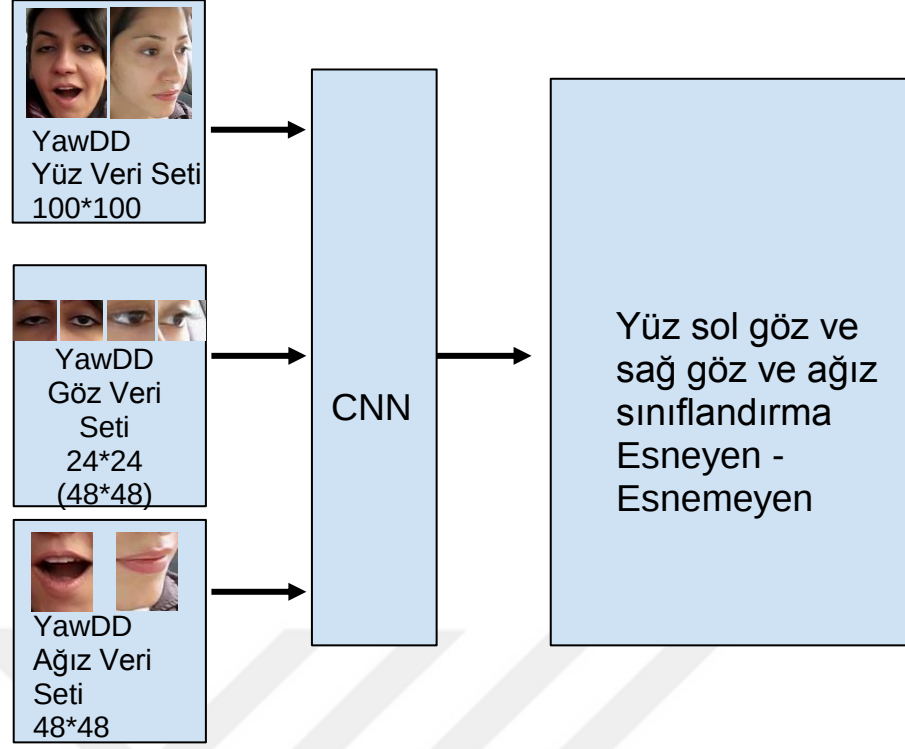
Cew ve YawDD veri setinde tasarlanan fonksiyonel modeller; Cew veri setinde yüz, sol göz ve sağ göz görüntüleri için Şekil 4.11, YawDD veri setinde yüz, sol göz ve sağ göz görüntüleri için Şekil 4.12, yüz, sol göz, sağ göz ve ağız görüntüleri için Şekil 4.13'te sunulmuştur.



Şekil 4.11. Cew veri setinde yüz, sol göz ve sağ göz görüntü parçalarına dayalı fonksiyonel derin öğrenme modeli.



Şekil 4.12. YawDD veri setinde yüz, sol göz ve sağ göz görüntü parçalarına dayalı fonksiyonel derin öğrenme modeli.



Şekil 4.13. YawDD veri setinde yüz, sol göz, sağ göz ve ağız görüntü parçalarına dayalı fonksiyonel derin öğrenme modeli.

4.4. Topluluk Öğrenmesi Yöntemine Dayalı Sınıflandırma Modeli

Topluluk öğrenmesi, modeller dizisi içerisinde en yüksek performans gösteren modeli elde etmek yerine modelleri oylamayı tercih etmektedir. Bu sayede klasik sınıflandırma yöntemlerine kıyasla sınıflandırmada daha iyi bir başarı sergilemektedir. Hesaplama karmaşıklığını artırmayı göze alarak daha etkin bir performans sergilemeyi amaçlamaktadır.

Bu tez çalışmasında yorgunluk tanıma alanında CNN mimarisine dayalı derin öğrenme yöntemi kullanılarak tasarlanan sıralı ve fonksiyonel modeller üzerinde eğitimler tamamlandıktan sonra uygun olan eğitilmiş modellerin sonuçları, model ortalaması topluluk öğrenmesi yöntemi ile birleştirilmiştir. Bu yöntem kullanılarak iki farklı veri seti üzerinde, sıralı modellerde yüz eğitim modelleri ve göz eğitim modelleri, fonksiyonel modellerde de yüz ve göz eğitim modelleri olarak geliştirilen modeller kendi arasında birleştirilerek topluluk öğrenmesinin sağladığı avantajlardan yararlanılmıştır.

Bu topluluk modelinde herhangi bir yeni öğrenme işlemi gerekmemektedir. Eğitilmiş olan modeller birleştirilerek geliştirilen topluluk öğrenmesi modelinin test verileri üzerinde test edilmesi ile sonuçların elde edilmesi yeterli olmaktadır.

Literatürde çok sayıda topluluk öğrenmesi yöntemi bulunmakla birlikte diğer topluluk öğrenmesi yöntemlerinin kullanımı tez kapsamı dışında tutulmuştur. Farklı veri setleri, yüz, göz, ağız gibi farklı görüntü parçaları ve verilerdeki görüntülerin farklı boyutlarda olması topluluk öğrenmesi yaklaşımlarının uygulanmasında zorluklara yol açmaktadır. İşin doğasına daha uygun olduğu için, tezin genelinde bir görüntüdeki farklı görüntü parçalarının özniteliklerinin birleştirildiği yapılar üzerine odaklanılmıştır. Önceki alt bölümlerde açıklandığı gibi, yüz, göz ve ağız görüntüleri girdiler olarak verilerek öznitelikleri, derin öğrenme yapısı içinde melezleştirilmiştir.

5. DENEYSEL ÇALIŞMALAR

5.1. Giriş

Bu tez çalışmasında geliştirilen yorgunluk tanıma sistemi iki farklı veri seti üzerinde eğitilmekte ve test edilmektedir. Veri setleri üzerinde yüz, göz ve ağız bölgeleri ayrıştırılarak değerlendirilmektedir.

Tez kapsamında Cew ve YawDD veri setleri üzerinde gerçekleştirilen 3 farklı deneysel çalışma grubu yer almaktadır:

- 1) Yüz, göz ve ağız görüntü parçalarının her biri için sıralı model üzerinde ayrı başarımlar elde edilmiştir.
- 2) Aynı yüz görüntüsüne ait yüz bölümlerinin ayrı girdi olarak verilip özneliklerinin birleştirildiği fonksiyonel model üzerinde deneysel sonuçlar elde edilmiştir.
- 3) Model ortalaması topluluk öğrenmesi yöntemi kullanılarak farklı veri setlerinin benzer özelliklerinin birleştirilmesi ile deneysel sonuçlar elde edilmiştir.

Tensorflow kütüphanesi üzerinde çalışan üst seviyeli Keras kütüphanesi kullanılarak Python programlama dili üzerinde derin öğrenme modelleri oluşturulmuştur. Görüntülerin ön işleme aşamasında; OpenCV, Scikit-image, Scikit-learn ve Numpy kütüphaneleri kullanılmıştır. Deneysel çalışmalardan alınan sonuçların görselleştirilmesinde Matplotlib ve Scikit-learn kütüphanesinden yararlanılmıştır.

5.2. Başarım Ölçütleri

Model başarısı değerlendirilirken kullanılan kavramlar model başarım ölçütleri olarak nitelendirilmektedir. Bu kavramlar hata oranı, kesinlik, duyarlılık ve F-ölçütüdür. Modelin başarısı, doğru sınıfa ve yanlış sınıfa atılan örnek sayısına

bağlı olarak belirlenmektedir. Test sonucunda elde edilen sonuçların başarı bilgileri karışıklık matrisi ile ifade edilmektedir. Karışıklık matrisinin genel gösterimi Tablo 5.1’de yer almaktadır (Coşkun ve Baykal, 2011).

Tablo 5.1. Karışıklık Matrisi.

		Tahminlenen Sınıf	
		Sınıf (1)	Sınıf (0)
Doğru Sınıf	Sınıf (1)	TP (Doğru Pozitif)	FN (Yanlış Negatif)
	Sınıf (0)	FP (Yanlış Pozitif)	TN (Doğru Negatif)

Karışıklık matrisleri, sınıflandırma işlemlerinde hataların değerlendirilmesinin bir yoludur. Her bir sınıf için hataları incelemek amacıyla kullanılmaktadır. Bu karışıklık matrisinden, TP ve TN, doğru şekilde sınıflandırılan pozitif ve negatif örneklerin sayısını göstermektedir. Karışıklık matrisleri, genel olarak sınıflandırıcı performansını değerlendirmek için kullanılmaktadır. Bu matriste yanlış sınıflandırılan örnekler FP, FN ile; doğru sınıflandırılan örnekler TP ve TN ile gösterilmektedir (Beauxis-Aussalet and Hardman, 2014; Hossin and Sulaiman, 2015).

Bu tez çalışmasında doğruluk (accuracy), kesinlik (precision), duyarlılık (recall) ve F-ölçütü (F-score) değerlendirme ölçütleri tercih edilmiştir. Bunlara ek olarak categorical crossentropy maliyet fonksiyonu ile elde edilen kayıp (loss) değeri de performans değerlendirmesinde kullanılmıştır.

Formül 5.1 ile Formül 5.5 arasında yer alan formüller Coşkun ve Baykal (2011)’den alınmıştır.

Doğruluk, sınıflandırılan tüm örnekler arasında alınan hem pozitif hem de negatif gerçek örneklerin toplam örnek sayısına oranıdır (Dalianis, 2018).

Doğruluk değeri, Formül 5.1’de gösterildiği gibi belirlenmektedir.

$$\text{Doğruluk} = \frac{TP + TN}{TP + FP + FN + TN} \quad (5.1)$$

Hata Oranı, Formül 5.2'deki belirtildiği şekilde bulunur. Yanlış sınıflandırılmış örnek sayısının, toplam örnek sayısına oranıdır.

$$\text{Hata Oranı} = \frac{FP + FN}{TP + FP + FN + TN} \quad (5.2)$$

Kesinlik, sınıfı 1 olarak tahminlenmiş TP örnek sayısının, sınıfı 1 olarak tahminlenmiş tüm örnek sayısına oranıdır (Hossin and Sulaiman, 2015).

Kesinlik değeri, Formül 5.3 ile elde edilir.

$$\text{Kesinlik} = \frac{TP}{TP + FP} \quad (5.3)$$

Duyarlılık, doğru sınıflandırılmış pozitif örnek sayısının toplam pozitif örnek sayısına oranıdır.

Duyarlılık hesabı, Formül 5.4'te sunulmaktadır.

$$\text{Duyarlılık} = \frac{TP}{TP + FN} \quad (5.4)$$

F-Ölçütü, kesinlik ve duyarlılık ölçütleri arasındaki harmonik ortalamadır. Kesinlik ve duyarlılığa farklı ağırlık veren farklı indislere sahip olabilir (Hossin and Sulaiman, 2015).

F-ölçütü, Formül 5.5'te sunulan işlem ile hesaplanmaktadır.

$$F - \text{Ölçütü} = \frac{2 \times \text{Duyarlılık} \times \text{Kesinlik}}{\text{Duyarlılık} + \text{Kesinlik}} \quad (5.5)$$

5.3. Kullanılan Veri Setleri

Bu tez çalışmasında yorgunluk tanıma için herhangi bir ücrete tabi olmadan erişime sunulan iki farklı veri seti kullanılmaktadır. Hem yüz hem göz görüntüleri içeren Cew, yüz görüntülerinden oluşan YawDD veri setleri üzerinde deneysel çalışmalar yapılmıştır.

5.3.1. Cew veri seti

Cew veri seti, farklı kişiler üzerinde, farklı aydınlatma koşulları ve bulanıklıkta çekilen kapalı ve açık göz görüntüleri ile gözleri kapalı ve açık olan yüz görüntülerinden oluşmaktadır (Song et al., 2014). Göz görüntülerinin her bir türü (sol ve sağ) 2'şer, yüz görüntüleri 2 sınıf içermektedir. Açık sağ göz olarak 1231, açık sol göz olarak 1231, kapalı sağ göz olarak 1192, kapalı sol göz olarak 1192; toplamda 4846 göz görüntüsü içermektedir. Her bir sınıfa ait veri miktarları Tablo 5.2'de sunulmuştur.

Tablo 5.2. Cew veri setinde göz görüntülerine ait veri miktarları.

Göz Sınıfı		Veri Miktarları
Sol	Kapalı	1192
	Açık	1231
Sağ	Kapalı	1192
	Açık	1231
Toplam		4846

Gözleri kapalı 1192, gözleri açık 1231; toplamda 2423 yüz görüntüsü içermektedir. Her bir sınıfa ait veri miktarı Tablo 5.3'te sunulmuştur.

Tablo 5.3. Cew veri setinde yüz görüntülerine ait veri miktarları.

Yüz Sınıfı	Veri Miktarları
Gözleri Kapalı	1192
Gözleri Açık	1231
Toplam	2423

5.3.2. YawDD veri seti

YawDD veri seti, gerçek bir araçta farklı etnik kökenden erkek ve kadın sürücülerin gözlüklü, güneş gözlüklü, gözlüksüz, konuşan, şarkı söyleyen, sessiz ve esneyen durumlardaki hallerinin araç içi kamera kullanılarak video görüntüsü olarak kaydedilmiş verilerden oluşmaktadır (Abtahi et al., 2014).

Bu tez çalışmasında, YawDD veri setindeki video görüntülerinden çıkarılmış görüntü kareleri ile oluşturulmuş ve eğitim ile test verisi olarak ayrıştırılmış görüntülerden oluşan YawDD veri seti kullanılmıştır (Raju, 2020). Bu veri setinde eğitim verisi olarak 617 esneyen, 616 esnemeyen görüntü olmak üzere toplam 1233; test verisi olarak esneyen 106, esnemeyen 109 olmak üzere toplam 215 görüntü; eğitim ve test verisi olarak toplam 1448 görüntü yer almaktadır. Bu veri seti üzerinde MTCNN yöntemi kullanılarak yüz, sol göz, sağ göz ve ağız görüntülerinin ayrıştırılması sonucunda eğitim verisinde esneyen 384, esnemeyen 426 olmak üzere toplam 810; test verisinde esneyen 61, esnemeyen 66 olmak üzere toplam 127; eğitim ve test verisi olarak toplam 937 görüntü elde edilmiştir.

YawDD veri setindeki (Raju, 2020) eğitim ve test görüntü sayısı ile bu görüntülerden MTCNN kullanılarak ayrıştırılarak oluşturulan veri setindeki eğitim ve test görüntü sayısı Tablo 5.4'te sunulmuştur.

Tablo 5.4. YawDD veri setindeki veri miktarı ile MTCNN kullandıktan sonra elde edilen veri miktarları.

Veri Sınıfı		YawDD veri setindeki veri miktarları	YawDD veri seti üzerinde MTCNN kullanılarak yüz, sol göz, sağ göz ve ağız görüntüleri ayrıştırıldıktan sonra elde edilen veri miktarları
Esneyen	Eğitim	617	384
	Test	106	61
Esnemeyen	Eğitim	616	426
	Test	109	66
Toplam		$1233 + 215 = 1448$	$810 + 127 = 937$

5.4. Veri Ön İşleme

Bu tez çalışmasında, hem yüz hem göz görüntüleri üzerinde çalışma yapılan CEW veri setinde görüntülerden aynı kişilere ait göz ve yüz görüntülerinin aynı grupta (göz görüntüleri eğitimde ise yüz görüntüsü de eğitimde gibi) olmasına dikkat edilerek eğitimde %70, doğrulamada %15, testte %15 oranında olacak şekilde ayrılmıştır. Derin öğrenme modelinde göz görüntüleri orijinal boyutu dışında eğitilirken 48x48 olarak yeniden boyutlandırılmıştır.

Yüz, sol göz, sağ göz ve ağız görüntüleri üzerinde çalışma yapılan YawDD veri setindeki tüm görüntülerden MTCNN kullanılarak yüz, sol ve sağ göz ile ağız görüntüleri elde edilmiştir. Eğitim ve test olarak ayrılmış olan verilerin aynı kişiye ait olmayan yüz, göz ve ağız görüntüleri ayıklanmıştır. Derin öğrenme modellerinde görüntüler eğitilirken gri tonlamalı görüntüye dönüştürülmüştür. Sol ve sağ göz için görüntü boyutları 24x24 ve 48x48 olarak, ağız için görüntü boyutu 48x48 olarak, yüz için görüntü boyutu 100x100 olarak yeniden boyutlandırılmıştır.

5.5. CNN Mimarisine Dayalı Derin Öğrenme Modeli Gerçekleştirimi

Bu tez çalışmasında geliştirilen CNN mimarisine dayalı derin öğrenme modelleri, Python programlama dili ile Tensorflow kütüphanesi üzerinde çalışan Keras kütüphanesi kullanılarak gerçekleştirilmiştir. Keras, derin öğrenme model

yapısı olarak hem sıralı (sequential) hem de fonksiyonel (functional) model yapıları sunmaktadır. Sıralı model yapısında tek girdi katmanı bulunmaktadır. Diğer katmanlar birbiri ardınca sıralanan bir yapıya sahiptir. Her bir katmanın çıktısı bir sonraki katmanın girdisi konumundadır. Fonksiyonel model yapısı, farklı girdi katmanlarının kullanılabilirdiği ve çıktı katmanlarının da istenilen şekilde birleştirilmesine olanak tanıyan, daha esnek bir yapıya sahiptir. Bu tez çalışması kapsamında hem sıralı hem fonksiyonel model yapıları kullanılarak derin öğrenme modelleri geliştirilmiştir.

Geliştirilen CNN mimarisine dayalı derin öğrenme modellerinin girdi katmanında veri boyutlarının parametre olarak kullanıldığı Keras kütüphanesine ait olan Input fonksiyonu kullanılmıştır. Çalışma kapsamında yüz görüntüleri için 100x100x1, göz görüntüleri için 24x24x1 ve 48x48x1, ağız görüntüleri için 48x48x1 boyutları parametre değerleri olarak verilmiştir. Evrişim katmanlarında filtre boyutu ve sayısının parametre değeri olarak kullanıldığı Keras kütüphanesinin Conv2D fonksiyonu kullanılmıştır. Sıralı model yapısında kullanılan evrişim katmanlarında her bir katman çıktısı bir sonraki katmanın girdisi olacak şekilde tanımlanmıştır. Fonksiyonel model yapısında da sıralı model yapısına benzer şekilde her bir katman çıktısı bir sonraki katmanın girdisi olarak verilmektedir. Farklı olarak tüm girdilerden elde edilen flatten katmanları birleştirilerek dense katmanına verilmekte ve sınıflandırma yapılmaktadır.

Her bir evrişim katmanının ardından bir aktivasyon katmanı, bir seyrekleştirme katmanı ve bir normalizasyon katmanı eklenmiştir. Bazı evrişim katmanlarından sonra maksimum havuzlama katmanı da eklenmiştir. Aktivasyon fonksiyonu olarak ReLu aktivasyon fonksiyonunun tercih edilmesi sebebiyle Keras kütüphanesinin Aktivasyon fonksiyonuna parametre değeri olarak 'relu' değeri verilmiştir. Seyrekleştirme katmanında bağlantıları engellenmek istenen yapay sinir hücresi sayısının katmandaki toplam sinir sayısına oranını parametre değeri olarak alan Dropout fonksiyonu kullanılmıştır. Parametre değerini [0,1] aralığında alan fonksiyon için parametre değeri 0,2 tercih edilmiştir. Normalizasyon katmanında BatchNormalization fonksiyonu kullanılmıştır. Havuzlama katmanında parametre değerleri olarak filtre boyutu ve stride değeri alan MaxPooling2D fonksiyonu kullanılmıştır. Parametre değerleri olarak filtre boyutu

2x2, stride değeri 2 olan ön tanımlı değerler kullanılmıştır. Çıktı katmanında sınıflandırma işlemi gerçekleştirildiğinden bu katmanda Dense fonksiyonuna parametre olarak sınıf sayısı, Activation fonksiyonuna da parametre olarak 'softmax' verilmiştir. Fonksiyonel model yapısında tüm bu katmanlara ek olarak çıktı katmanından önce girdilere ait tüm katmanların birleştirilmesi amacıyla her bir girdiye ait tanımlanan katmanların tek boyutlu vektörlerini parametre değeri olarak alan concatenate fonksiyonu kullanılmıştır.

Sıralı ve fonksiyonel model yapılarında Keras kütüphanesinin ilgili fonksiyonları kullanılarak tüm katmanlarının tanımlanmasından sonra Model fonksiyonuna parametre değerleri olarak girdi ve çıktı katmanlarının verilmesiyle tasarlanan model oluşturulmaktadır. Tasarlanan modellerde yalnızca evrişim tam bağlantı katmanları dikkate alınarak katman sayısı belirlenmektedir. Bu çalışmada tasarlanan sıralı modeller 4 adet evrişim ve 2 adet tam bağlantı katmanı olmak üzere 6 adet katman içermektedir. Fonksiyonel modellerden yüz, sol göz ve sağ göz girdi görüntülerinin birleştirildiği model için 2 adet tam bağlantı katmanı ve her bir özellik için 4 adet evrişim katmanı olan 12 adet evrişim katmanı olmak üzere toplamda 14 adet; yüz, sol göz, sağ göz ve ağız girdi görüntülerinin birleştirildiği model için 2 adet tam bağlantı katmanı ve her bir özellik için 4 adet evrişim katmanı olan 16 adet evrişim katmanı olmak üzere toplamda 18 adet katman içermektedir.

Geliştirilen sıralı modellerin her bir katmanına ilişkin parametre değerleri Ekler bölümünde Şekil A.1 - Şekil A.7 arasında gösterilmiştir.

Geliştirilen fonksiyonel modellerin her bir katmanına ilişkin parametre değerleri Ekler bölümünde Şekil A.8 - Şekil A.13 arasında gösterilmiştir.

Geliştirilen modellerin eğitiminde kullanılan hiperparametre değerleri Tablo 5.5'te gösterilmiştir.

Tablo 5.5. CNN modellerinde kullanılan hiperparametre deęerleri.

Hiperparametre	Deęeri
Maliyet Fonksiyonu (loss function)	Categorical Crossentropy
Optimizasyon algoritması	Adam
Batch size	16
Epoch	100

5.6. Topluluk Öğrenmesine Dayalı Derin Öğrenme Modeli Gerçekleřtirmesi

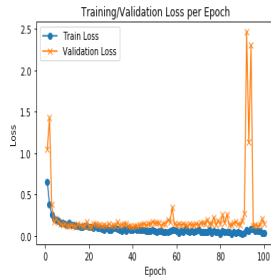
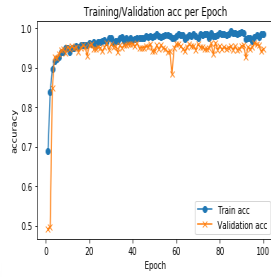
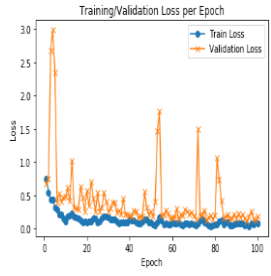
Bu tez çalışmasında keras kütüphanesinin sıralı model yapısı kullanılarak eğitilen CNN mimarisine dayalı derin öğrenme modelleri, model ortalaması topluluk öğrenmesi yöntemi ile birleştirilmiştir. Cew ve YawDD olmak üzere iki farklı veri seti üzerinde eğitilen sıralı modeller, topluluk öğrenmesi modelinin girdi katmanlarını oluşturmaktadır. mean() fonksiyonuna sıralı modeller girdi olarak verilmektedir. Test verisi üzerinde, girdi modellerinin mean() fonksiyonu kullanılarak elde edilen tahminleme ortalaması, modelin çıktısını oluşturmaktadır. Veri setleri üzerinde girdi katmanına yüz görüntülerinin verildięi sıralı modeller üzerinde, göz görüntülerinin verildięi sıralı modeller üzerinde ve yüz, sol göz, saę göz görüntülerinin verildięi fonksiyonel modeller üzerinde olmak üzere farklı model ortalaması topluluk öğrenme modelleri oluşturulmuştur.

5.7. CEW Veri Seti Üzerindeki Deneysel Sonuçlar

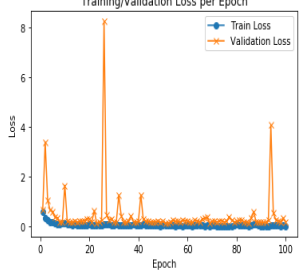
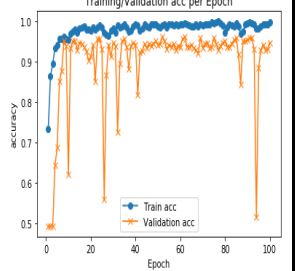
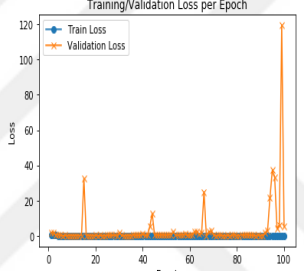
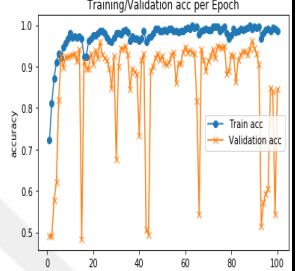
Bu tez çalışmasında önerilen yorgunluk tanıma yöntemlerinin Cew veri seti üzerinde eğitilmesi ve test edilmesi ile elde edilen sonuçlar, bu bölümde yer almaktadır. Önerilen yöntemde yüz ve gözlerden her bir özellięe ait görüntülerin tek başlarına kullanılması ile elde edilen sınıflandırma sonuçları ve özelliklerin birlikte kullanılması ile elde edilen sınıflandırma sonuçları bu bölümde verilmiştir.

Yapılan eğitimlerdeki eğitim ve doğrulama verileri üzerindeki kayıp ve doğruluk başarı değerleri ile bu değerlere ait grafikler Tablo 5.6 ve 5.7’de sunulmuştur.

Tablo 5.6. Cew veri setinde sıralı modellerin eğitim ve doğrulama verileri üzerindeki başarı değerleri.

Veri Seti				Kayıp Grafiği	Doğruluk Grafiği
Cew Gözler	Görüntü Boyutu (piksel)		24		
	Eğitim	Kayıp	0,0354		
		Doğruluk (%)	98,58		
	Doğru lama	Kayıp	0,1496		
		Doğruluk (%)	94,78		
	Cew Gözler	Görüntü Boyutu (piksel)			
Eğitim		Kayıp	0,0864		
		Doğruluk (%)	96,84		
Doğru lama		Kayıp	0,4894		
		Doğruluk (%)	92,99		
Cew Yüz		Görüntü Boyutu (piksel)		100	
	Eğitim	Kayıp	0,0739		
		Doğruluk (%)	97,23		
	Doğru lama	Kayıp	0,1845		
		Doğruluk (%)	93,68		

Tablo 5.7. Cew veri setinde fonksiyonel modellerin eğitim ve doğrulama verileri üzerindeki başarı değerleri.

Veri Seti				Kayıp Grafiği	Doğruluk Grafiği
Cew Sol göz, Sağ göz, Yüz	Görüntü Boyutu (piksel)	24			
		24			
		100			
	Eğitim	Kayıp	0,0101		
		Doğruluk (%)	99,71		
	Doğrulama	Kayıp	0,1875		
		Doğruluk (%)	94,78		
Cew Sol göz, Sağ göz, Yüz	Görüntü Boyutu (piksel)	48			
		48			
		100			
	Eğitim	Kayıp	0,0349		
		Doğruluk (%)	98,70		
	Doğrulama	Kayıp	5,3150		
		Doğruluk (%)	84,62		

Yalnızca yüz girdi görüntüleri ile eğitilen model üzerindeki eğitim ve doğrulama verilerindeki doğruluk değerleri sırasıyla %97,23 ve %93,68 olarak elde edilmiştir.

Yalnızca göz girdi görüntüleri ile eğitilen model üzerindeki eğitim ve doğrulama verilerindeki doğruluk değerleri sırasıyla 24x24 piksel olarak boyutlandırıldığında %98,58 ve %94,78; 48x48 piksel olarak boyutlandırıldığında %96,84 ve %92,99 olarak elde edilmiştir.

Yüz, sol göz ve sağ göz görüntülerinin birlikte girdi olarak verildiği model üzerindeki eğitim ve doğrulama verilerindeki doğruluk değerleri sırasıyla gözler 24x24 piksel olarak boyutlandırıldığında %99,71 ve %94,78; gözler 48x48 piksel

olarak boyutlandırıldığında %98,70 ve %84,62 olarak elde edilmiştir. Yüz ile 24x24 piksel boyutlu sol ve sağ göz girdi görüntülerinin birlikte girdi olarak verildiği model, Cew veri setinde tüm diğer modellerden daha yüksek bir eğitim verisi başarı değerine ulaşmıştır. Doğrulama verisi başarısında da en yüksek başarı değerine, 24x24 piksel boyutunda göz görüntüsünün girdi olarak verildiği tekil modelde ulaşılmıştır.

Yapılan testlerdeki başarımların değerleri Tablo 5.8’de sunulmuştur.

Tablo 5.8. Cew veri seti test verileri üzerindeki başarımların değerleri.

Veri Seti	Görüntü Boyutu (piksel)	Sınıf	Kesinlik	Duyarlılık	F-Ölçütü	Doğruluk (%)
Cew Yüz	100	Kapalı	0,92	0,99	0,95	95,33
		Açık	0,99	0,91	0,95	
Cew Gözler	24	Kapalı	0,98	0,99	0,98	98,35
		Açık	0,99	0,98	0,98	
Cew Gözler	48	Kapalı	0,99	0,95	0,97	97,12
		Açık	0,95	0,99	0,97	
Cew Sol Göz, Sağ Göz, Yüz	24 24 100	Kapalı	0,99	0,95	0,97	96,98
		Açık	0,95	0,99	0,97	
Cew Sol Göz, Sağ Göz, Yüz	48 48 100	Kapalı	0,77	0,98	0,86	84,89
		Açık	0,98	0,72	0,83	

Yalnızca yüz girdi görüntüleri ile eğitilen model üzerindeki test verilerindeki doğruluk değeri %95,33 olarak elde edilmiştir.

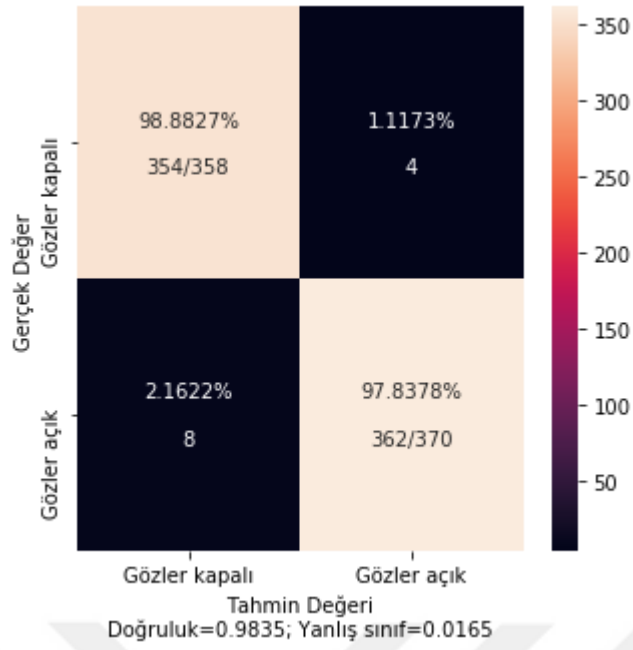
Yalnızca göz girdi görüntüleri ile eğitilen model üzerindeki test verilerindeki doğruluk değeri 24x24 piksel olarak boyutlandırıldığında %98,35; 48x48 piksel olarak boyutlandırıldığında %97,12 olarak elde edilmiştir.

Yüz, sol göz ve sağ göz girdilerinin birlikte girdi olarak verildiği model üzerindeki test verilerindeki doğruluk değeri gözler 24x24 piksel olarak boyutlandırıldığında %96,98 olarak elde edilmiştir.

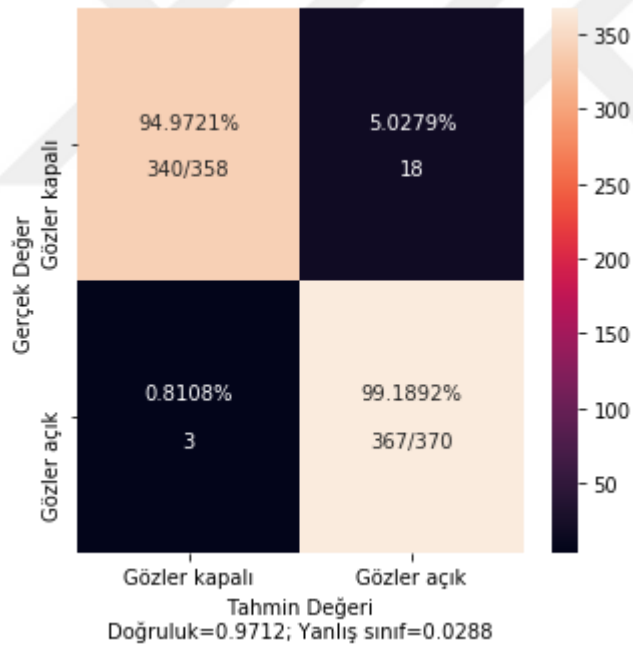
Yüz, sol göz ve sağ göz girdilerinin birlikte girdi olarak verildiği model üzerindeki test verilerindeki doğruluk değeri gözler 48x48 piksel olarak boyutlandırıldığında %84,89 olarak elde edilmiştir.

Tablo 5.6’da görüldüğü gibi Cew veri seti üzerinde sıralı modellerde eğitim başarısı olarak en iyi sonuç %98,58 doğruluk değeri ile göz görüntülerinde gerçekleştirilen eğitimden elde edilmiştir. Aynı veri setinde fonksiyonel model eğitimlerinde en iyi başarı değeri Tablo 5.7’de gösterildiği gibi veri setindeki orijinal görüntü boyutlarının kullanıldığı sol göz, sağ göz ve yüz görüntülerinin özneteliklerinin birleştirildiği modelin gerçekleştirilmesi ile elde edilmiştir. Tablo 5.8’de bu veri setinde test verileri üzerindeki en iyi sınıflandırma sonucunun da eğitim verileri üzerinde en iyi başarı elde edilen modellerden alındığı görünmektedir.

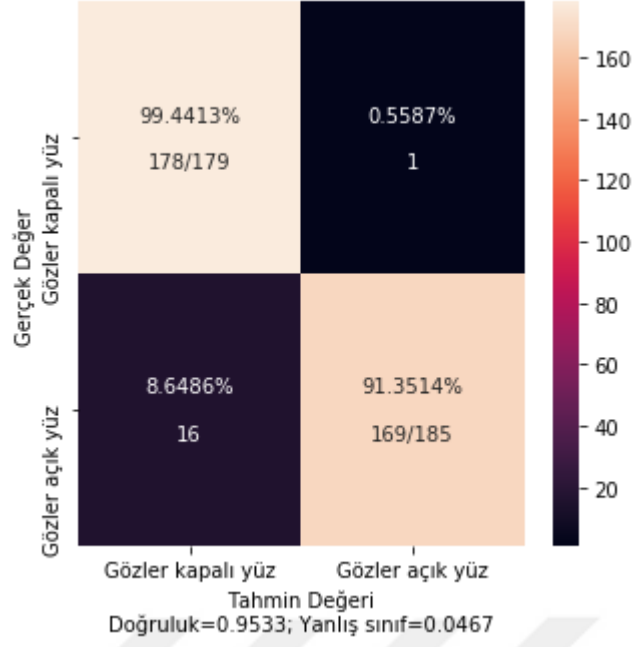
Cew veri seti üzerinde yapılan deneysel çalışmalarda elde edilen karışıklık matrisleri ve ardından başarı değerleri sunulmuştur. Karışıklık matrisleri (24x24, 48x48 boyutlarındaki göz; 100x100 boyutlarındaki yüz; sırasıyla 24x24, 24x24, 100x100 boyutlarındaki sol göz, sağ göz ve yüz; sırasıyla 48x48, 48x48, 100x100 boyutlarındaki sol göz, sağ göz ve yüz için) Şekil 5.1 - 5.5 arasında yer almaktadır.



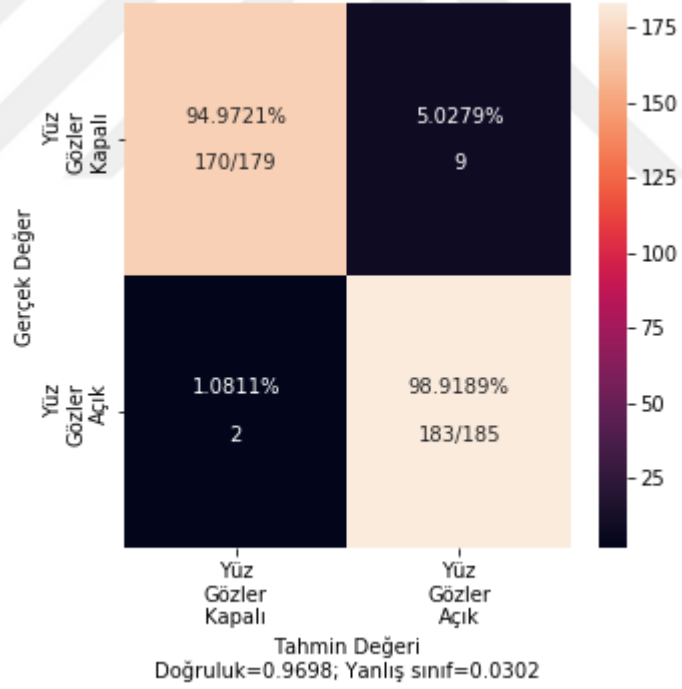
Şekil 5.1. 24x24 piksel boyutundaki göz görüntüleri için elde edilen karışıklık matrisi.



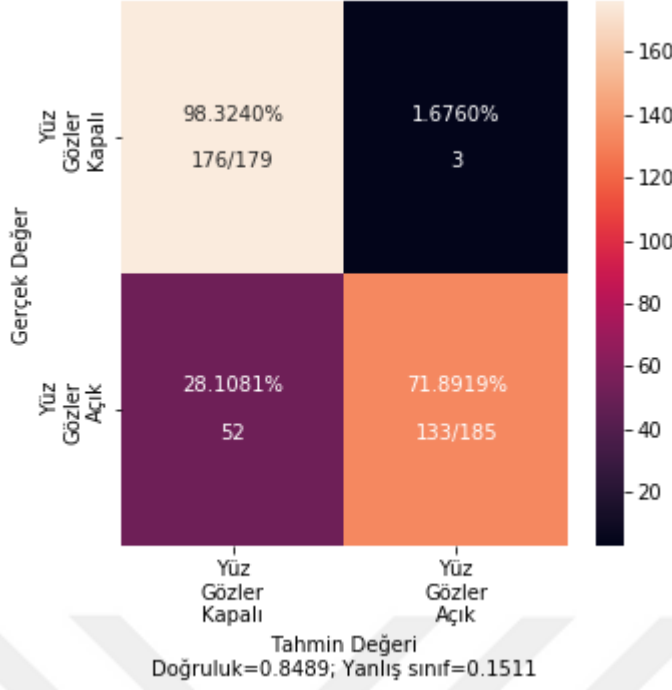
Şekil 5.2. 48x48 piksel boyutundaki göz görüntüleri için elde edilen karışıklık matrisi.



Şekil 5.3. 100x100 piksel boyutundaki yüz görüntüleri için elde edilen karışıklık matrisi.



Şekil 5.4. 24x24 piksel boyutundaki sol göz ve sağ göz ile 100x100 piksel boyutundaki yüz görüntülerinin özneliliklerinin birleşimi sonucunda elde edilen karışıklık matrisi.



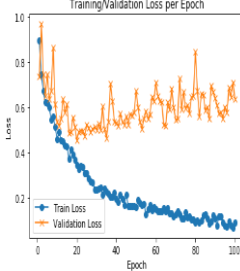
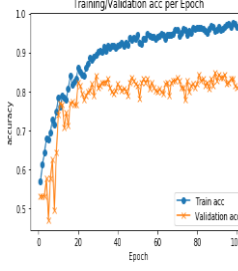
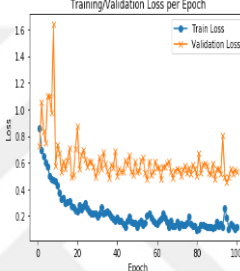
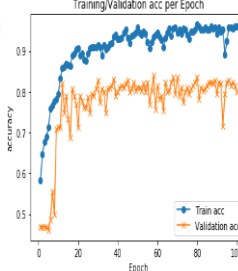
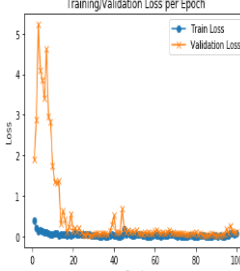
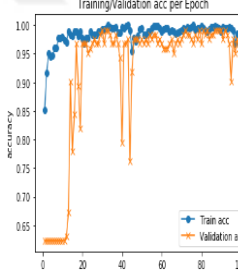
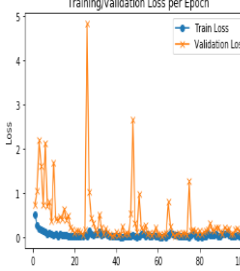
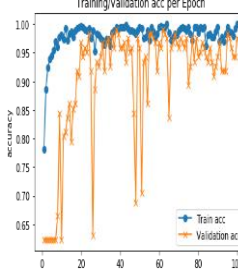
Şekil 5.5. 48x48 piksel boyutundaki sol göz ve sağ göz ile 100x100 piksel boyutundaki yüz görüntülerinin öznelitliklerinin birleşimi sonucunda elde edilen karışıklık matrisi.

5.8. YawDD Veri Seti Üzerindeki Deneysel Sonuçlar

Bu tez çalışmasında önerilen yorgunluk tanıma yöntemlerinin YawDD veri seti üzerinde eğitilmesi ve test edilmesi ile elde edilen sonuçlar bu bölümde yer almaktadır. Önerilen yöntemde yüz, ağız ve gözlerden her bir özelliğe ait görüntülerin tek başlarına kullanılması ile elde edilen sınıflandırma sonuçları ve özelliklerin birlikte kullanılması ile elde edilen sınıflandırma sonuçları bu bölümde verilmiştir.

Yapılan eğitimlerdeki eğitim ve doğrulama verileri üzerindeki kayıp ve doğruluk başarı değerleri ile bu değerlere ait grafikler Tablo 5.9 ve 5.10'da sunulmuştur.

Tablo 5.9. YawDD veri seti üzerindeki sıralı modellerin eğitim ve doğrulama verileri üzerindeki başarı değerleri.

Veri Seti				Kayıp Grafiği	Doğruluk Grafiği
YawDD Gözler	Görüntü Boyutu (piksel)		24		
	Eğitim	Kayıp	0,0919		
		Doğruluk (%)	96,51		
	Doğru lama	Kayıp	0,6370		
		Doğruluk (%)	81,07		
YawDD Gözler	Görüntü Boyutu (piksel)		48		
	Eğitim	Kayıp	0,1160		
		Doğruluk (%)	95,53		
	Doğru lama	Kayıp	0,5272		
		Doğruluk (%)	79,84		
YawDD Ağız	Görüntü Boyutu (piksel)		48		
	Eğitim	Kayıp	0,0686		
		Doğruluk (%)	97,97		
	Doğru lama	Kayıp	0,1081		
		Doğruluk (%)	95,08		
YawDD Yüz	Görüntü Boyutu (piksel)		100		
	Eğitim	Kayıp	0,0088		
		Doğruluk (%)	100,00		
	Doğru lama	Kayıp	0,1251		
		Doğruluk (%)	97,54		

Tablo 5.10. YawDD veri seti üzerindeki fonksiyonel modellerin eğitim ve doğrulama verileri üzerindeki başarı değerleri.

Veri Seti				Kayıp Grafiği	Doğruluk Grafiği
YawDD Sol göz, Sağ göz, Yüz	Görüntü Boyutu (piksel)		24 24 100		
	Eğitim	Kayıp	0,0121		
		Doğruluk (%)	99,42		
	Doğru lama	Kayıp	0,1334		
		Doğruluk (%)	95,90		
YawDD Sol göz, Sağ göz, Yüz	Görüntü Boyutu (piksel)		48 48 100		
	Eğitim	Kayıp	0,0091		
		Doğruluk (%)	99,56		
	Doğru lama	Kayıp	0,1045		
		Doğruluk (%)	97,54		
YawDD Ağız, Yüz	Görüntü Boyutu (piksel)		48 100		
	Eğitim	Kayıp	0,0034		
		Doğruluk (%)	99,85		
	Doğru lama	Kayıp	0,0324		
		Doğruluk (%)	98,36		
YawDD Sol göz, Sağ göz, Ağız, Yüz	Görüntü Boyutu (piksel)		24 24 48 100		
	Eğitim	Kayıp	0,0070		
		Doğruluk (%)	99,85		
	Doğru lama	Kayıp	0,1172		
		Doğruluk (%)	97,54		

Yalnızca yüz girdi görüntüleri ile eğitilen model üzerindeki eğitim ve doğrulama verilerindeki doğruluk değerleri sırasıyla %100,00 ve %97,54 olarak elde edilmiştir.

Yalnızca göz girdi görüntüleri ile eğitilen model üzerindeki eğitim ve doğrulama verilerindeki doğruluk değerleri sırasıyla 24x24 piksel olarak boyutlandırıldığında %96,51 ve %81,07; 48x48 piksel olarak boyutlandırıldığında %95,53 ve %79,84 olarak elde edilmiştir.

Yalnızca ağız girdi görüntüleri ile eğitilen model üzerindeki eğitim ve doğrulama verilerindeki doğruluk değerleri sırasıyla %97,97 ve %95,08 olarak elde edilmiştir.

Yüz, sol göz ve sağ göz görüntülerinin birlikte girdi olarak verildiği model üzerindeki eğitim ve doğrulama verilerindeki doğruluk değerleri sırasıyla gözler 24x24 piksel olarak boyutlandırıldığında %99,42 ve %95,90; gözler 48x48 piksel olarak boyutlandırıldığında %99,56 ve %97,54 olarak elde edilmiştir.

Yüz ve ağız görüntülerinin birlikte girdi olarak verildiği model üzerindeki eğitim ve doğrulama verilerindeki doğruluk değerleri sırasıyla %99,85 ve %98,36 olarak elde edilmiştir.

Yüz, sol göz, sağ göz ve ağız görüntülerinin birlikte girdi olarak verildiği model üzerindeki eğitim ve doğrulama verilerindeki doğruluk değerleri sırasıyla %99,85 ve %97,54 olarak elde edilmiştir.

Yapılan testlerdeki başarımların değerleri Tablo 5.11’de sunulmuştur.

Tablo 5.11. YawDD veri setindeki test verileri üzerinde elde edilen başarımların değerleri.

Veri Seti	Görüntü Boyutu (piksel)	Sınıf	Kesinlik	Duyarlılık	F-Ölçütü	Doğruluk (%)
YawDD Yüz	100	Esneyen	0,97	0,95	0,96	96,06
		Esne-meyen	0,96	0,97	0,96	
YawDD Göz	24	Esneyen	0,84	0,80	0,82	83,07
		Esne-meyen	0,82	0,86	0,84	
YawDD Göz	48	Esneyen	0,82	0,74	0,78	79,53
		Esne-meyen	0,78	0,85	0,81	
YawDD Ağız	48	Esneyen	1,00	0,97	0,98	98,43
		Esne-meyen	0,97	1,00	0,99	
YawDD Sol Göz, Sağ Göz, Yüz	24 24 100	Esneyen	1,00	0,92	0,96	96,06
		Esne-meyen	0,93	1,00	0,96	
YawDD Sol Göz, Sağ Göz, Yüz	48 48 100	Esneyen	1,00	0,95	0,97	97,64
		Esne-meyen	0,96	1,00	0,98	
YawDD Ağız, Yüz	48 100	Esneyen	1,00	0,98	0,99	97,64
		Esne-meyen	0,99	1,00	0,99	
YawDD Sol Göz, Sağ Göz, Ağız, Yüz	24 24 48 100	Esneyen	1,00	0,98	0,99	99,21
		Esne-meyen	0,99	1,00	0,99	

Yalnızca yüz girdi görüntüleri ile eğitilen model üzerindeki test verilerindeki doğruluk değeri %96,06 olarak elde edilmiştir. Bu model, YawDD veri setinde diğer tüm modellerden daha yüksek bir eğitim verisi başarı değerine ulaşmıştır.

Yalnızca göz girdi görüntüleri ile eğitilen model üzerindeki test verilerindeki doğruluk değeri 24x24 piksel olarak boyutlandırıldığında %83,07; 48x48 piksel olarak boyutlandırıldığında %79,53 olarak elde edilmiştir.

Yalnızca ağız girdi görüntüleri ile eğitilen model üzerindeki test verilerindeki doğruluk değeri %98,43 olarak elde edilmiştir.

Yüz, sol göz ve sağ göz görüntülerinin birlikte girdi olarak verildiği model üzerindeki test verilerindeki doğruluk değeri gözler 24x24 piksel olarak boyutlandırıldığında %96,06 olarak elde edilmiştir.

Yüz, sol göz ve sağ göz görüntülerinin birlikte girdi olarak verildiği model üzerindeki test verilerindeki doğruluk değeri gözler 48x48 piksel olarak boyutlandırıldığında %97,64 olarak elde edilmiştir.

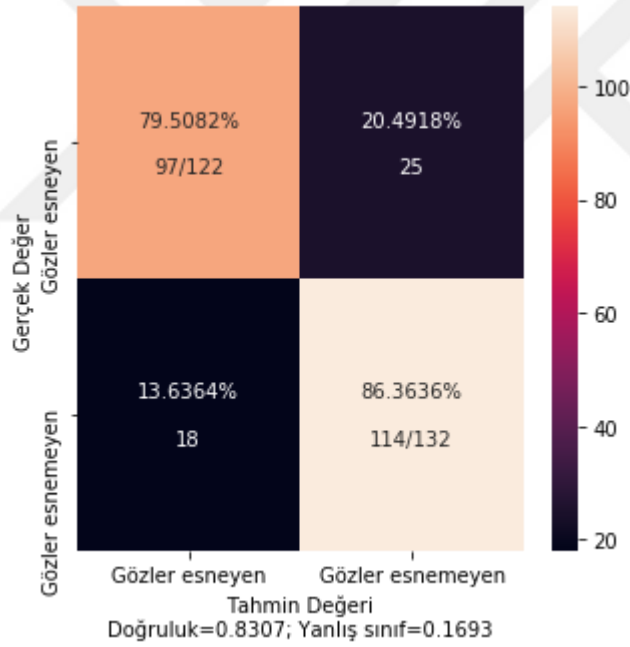
Yüz ile ağız görüntülerinin birlikte girdi olarak verildiği model üzerindeki test verilerindeki doğruluk değeri %97,64 olarak elde edilmiştir. Bu model, YawDD veri setinde diğer tüm modellerden daha yüksek bir doğrulama verisi başarı değerine ulaşmıştır.

Yüz, sol göz ve sağ göz ile ağız görüntülerinin birlikte girdi olarak verildiği model üzerindeki test verilerindeki doğruluk değeri %99,21 olarak elde edilmiştir.

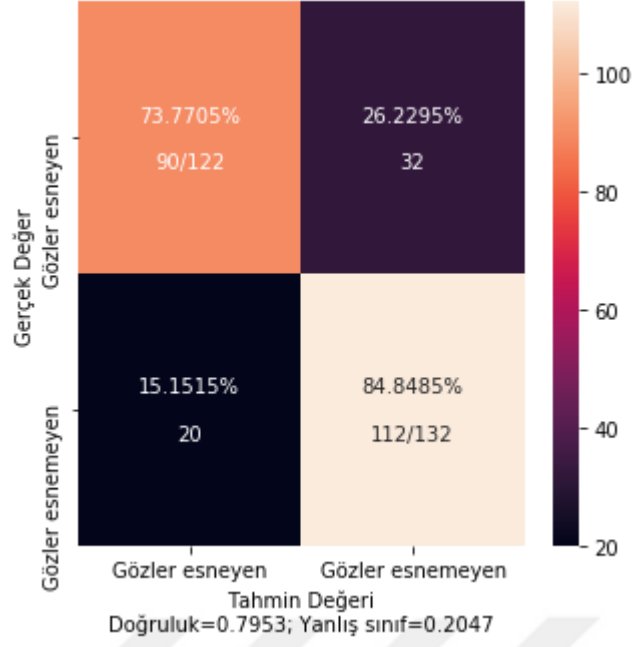
Tablo 5.9’da görüldüğü gibi YawDD veri seti üzerinde sıralı modellerde eğitim başarısı olarak en iyi sonuç %100 doğruluk değeri ile yüz görüntülerinde gerçekleştirilen eğitimden elde edilmiştir. Aynı veri setinde fonksiyonel model eğitimlerinde en iyi başarı değeri Tablo 5.10’da gösterildiği gibi sırasıyla 48x48, 100x100 piksel boyutlarında ağız ve yüz görüntülerinin ve sırasıyla 24x24, 24x24, 48x48, 100x100 piksel boyutlarında sol göz, sağ göz, ağız ve yüz görüntülerinin öz niteliklerinin birleştirildiği modellerin gerçekleştirilmesi ile %99,85 doğruluk değeri elde edilmiştir. Bu iki modelin eğitim verileri üzerindeki doğruluk değerleri aynı olmasına rağmen ağız ve yüz özelliklerinin birlikte eğitildiği fonksiyonel modelin kayıp değeri daha düşüktür. Tablo 5.11’de ise bu veri setinde test verileri üzerindeki en iyi sınıflandırma sonucunun sıralı modellerden ağız görüntüleri için

gerçekleştirilen modelden, fonksiyonel modellerden ise sol göz, sağ göz, ağız ve yüz görüntülerine ait öznelitliklerin birleştirildiği modelden elde edildiği görülmektedir. Sonuçta test verileri üzerinde tek deneme yapıldığında, en yüksek başarı sol göz, sağ göz, ağız ve yüz görüntülerinin birleşiminden elde edilmiştir.

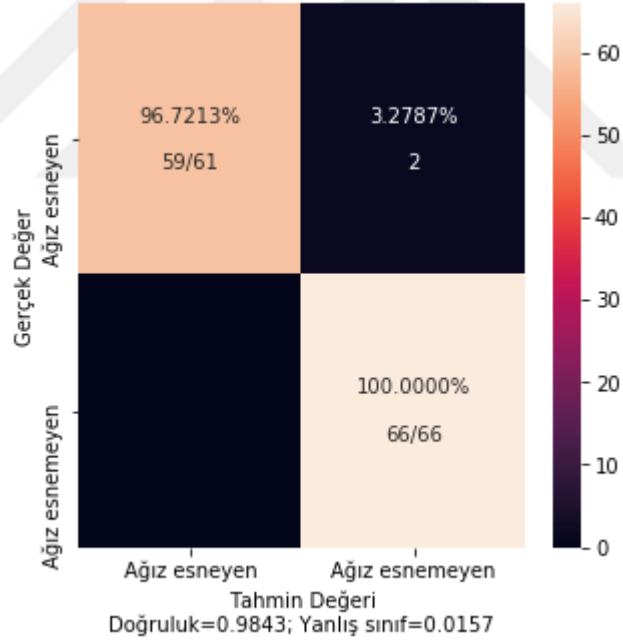
YawDD veri seti üzerinde yapılan deneysel çalışmalarda elde edilen karışıklık matrisleri ve ardından başarı değerleri sunulmuştur. Karışıklık matrisleri (24x24, 48x48 boyutlarındaki göz; 48x48 boyutlarındaki ağız; 100x100 boyutlarındaki yüz; sırasıyla 24x24, 24x24, 100x100 boyutlarındaki sol göz, sağ göz ve yüz; sırasıyla 48x48, 48x48, 100x100 boyutlarındaki sol göz, sağ göz ve yüz; 48x48, 100x100 boyutlarındaki ağız ve yüz; 24x24, 24x24, 48x48, 100x100 boyutlarındaki sol göz, sağ göz, ağız ve yüz için) Şekil 5.6 - 5.13 arasında yer almaktadır.



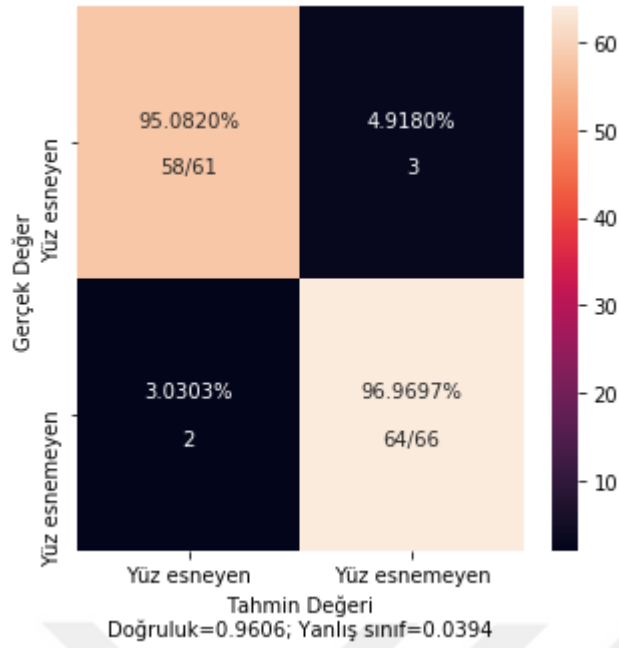
Şekil 5.6. 24x24 piksel boyutundaki göz görüntüleri için elde edilen karışıklık matrisi.



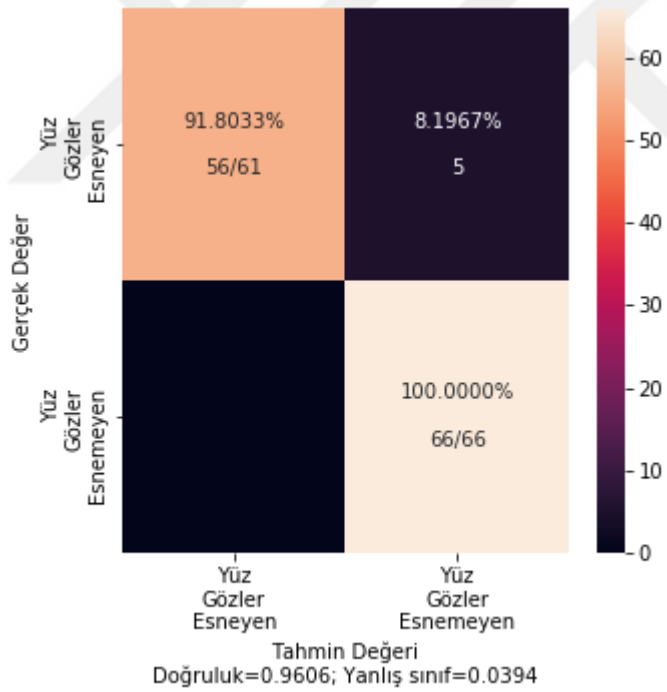
Şekil 5.7. 48x48 piksel boyutundaki göz görüntüleri için elde edilen karışıklık matrisi.



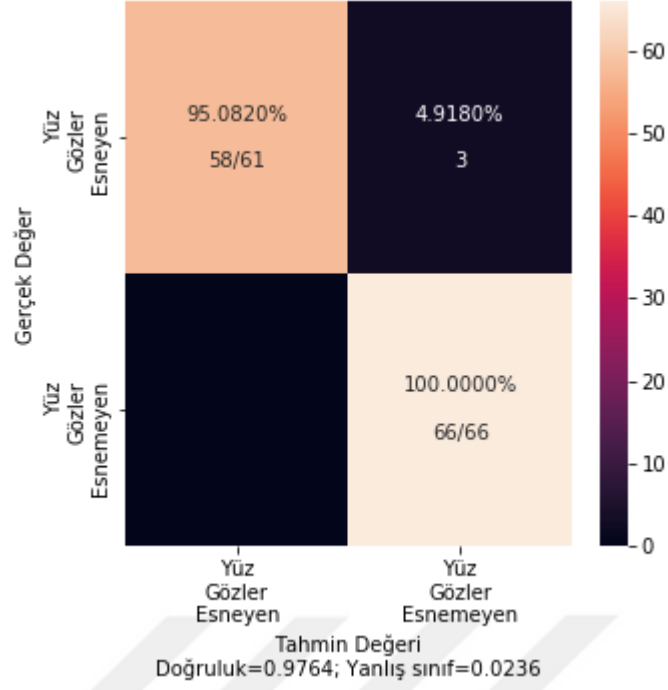
Şekil 5.8. 48x48 piksel boyutundaki ağız görüntüleri için elde edilen karışıklık matrisi.



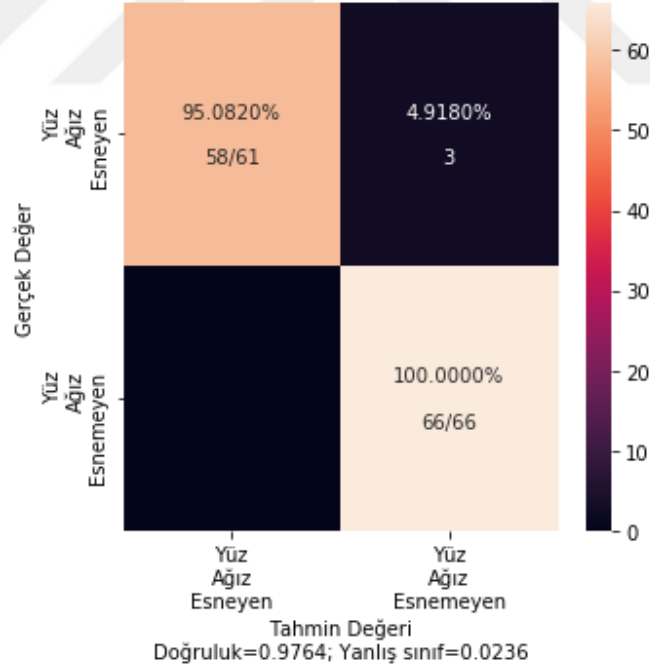
Şekil 5.9. 100x100 piksel boyutundaki yüz görüntüleri için elde edilen karışıklık matrisi.



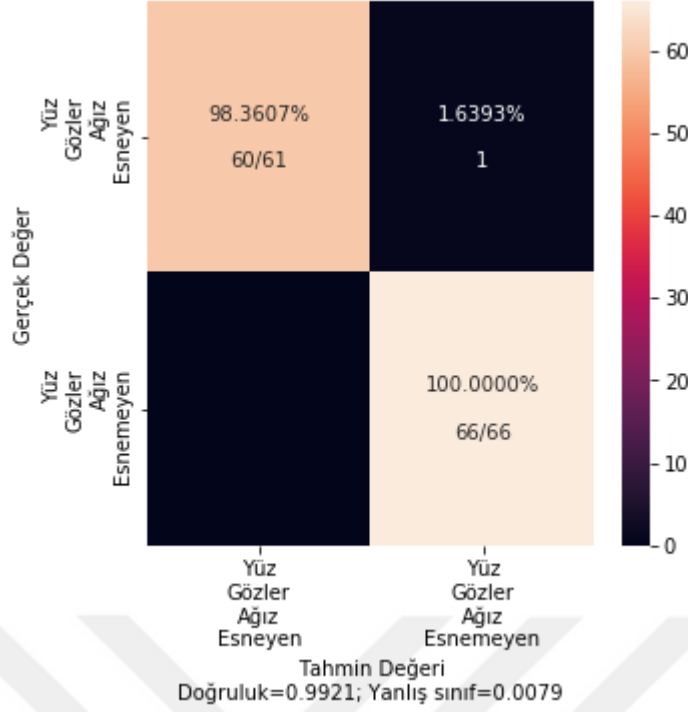
Şekil 5.10. 24x24 piksel boyutundaki sol göz ve sağ göz ile 100x100 piksel boyutundaki yüz görüntülerinin özneliliklerinin birleşimi sonucunda elde edilen karışıklık matrisi.



Şekil 5.11. 48x48 piksel boyutundaki sol göz ve sağ göz ile 100x100 piksel boyutundaki yüz görüntülerinin öz niteliklerinin birleşimi sonucunda elde edilen karışıklık matrisi.



Şekil 5.12. 48x48 piksel boyutundaki ağız ile 100x100 piksel boyutundaki yüz görüntülerinin öz niteliklerinin birleşimi sonucunda elde edilen karışıklık matrisi.



Şekil 5.13. 24x24 piksel boyutundaki sol göz ve sağ göz, 48x48 piksel boyutundaki ağız ile 100x100 piksel boyutundaki yüz görüntülerinin öznelitliklerinin birleşimi sonucunda elde edilen karışıklık matrisi.

5.9. YawDD Veri Setinde Deneylerin Tekrarlanması ile Elde Edilen Deneysel Sonuçlar

Tez çalışmasının bu bölümünde deneysel çalışmalarımızdaki en başarılı girdilerin olduğu modellerin başarısını daha doğru ve duyarlı bir şekilde ölçmek amacıyla ilgili deneyler 3 kere daha tekrarlanmış ve 4 deneyin ortalaması alınmıştır. Bu şekilde istatistiksel olarak daha anlamlı değerler elde edilmiştir. YawDD veri setinde tekrarlanan deneylerin eğitim doğrulama verileri üzerindeki kayıp ve doğruluk başarı değerleri Tablo 5.12’de, test verileri üzerindeki başarımlar Tablo 5.13’te gösterilmektedir. Ortalaması alınan ondalık değerlerde yukarı ve aşağı yuvarlamalar yapılmıştır.

Tablo 5.12. YawDD veri setinin eğitim ve doğrulama verileri üzerindeki ortalama kayıp ve doğruluk başarı değerleri.

Veri Seti	Görüntü boyutu (piksel)	Eğitim		Doğrulama	
YawDD Ağız	48	Kayıp	0,0333	Kayıp	0,0591
		Doğruluk (%)	99,02	Doğruluk (%)	97,95
YawDD Yüz	100	Kayıp	0,0238	Kayıp	0,1150
		Doğruluk (%)	99,20	Doğruluk (%)	97,13
YawDD Ağız, Yüz	48 100	Kayıp	0,0142	Kayıp	0,1064
		Doğruluk (%)	99,60	Doğruluk (%)	96,72
YawDD Sol göz, Sağ göz, Ağız, Yüz	24 24 48 100	Kayıp	0,0110	Kayıp	0,1387
		Doğruluk (%)	99,71	Doğruluk (%)	95,90

Tablo 5.13. YawDD veri setinin test verileri üzerinde elde edilen ortalama başarımların değerleri.

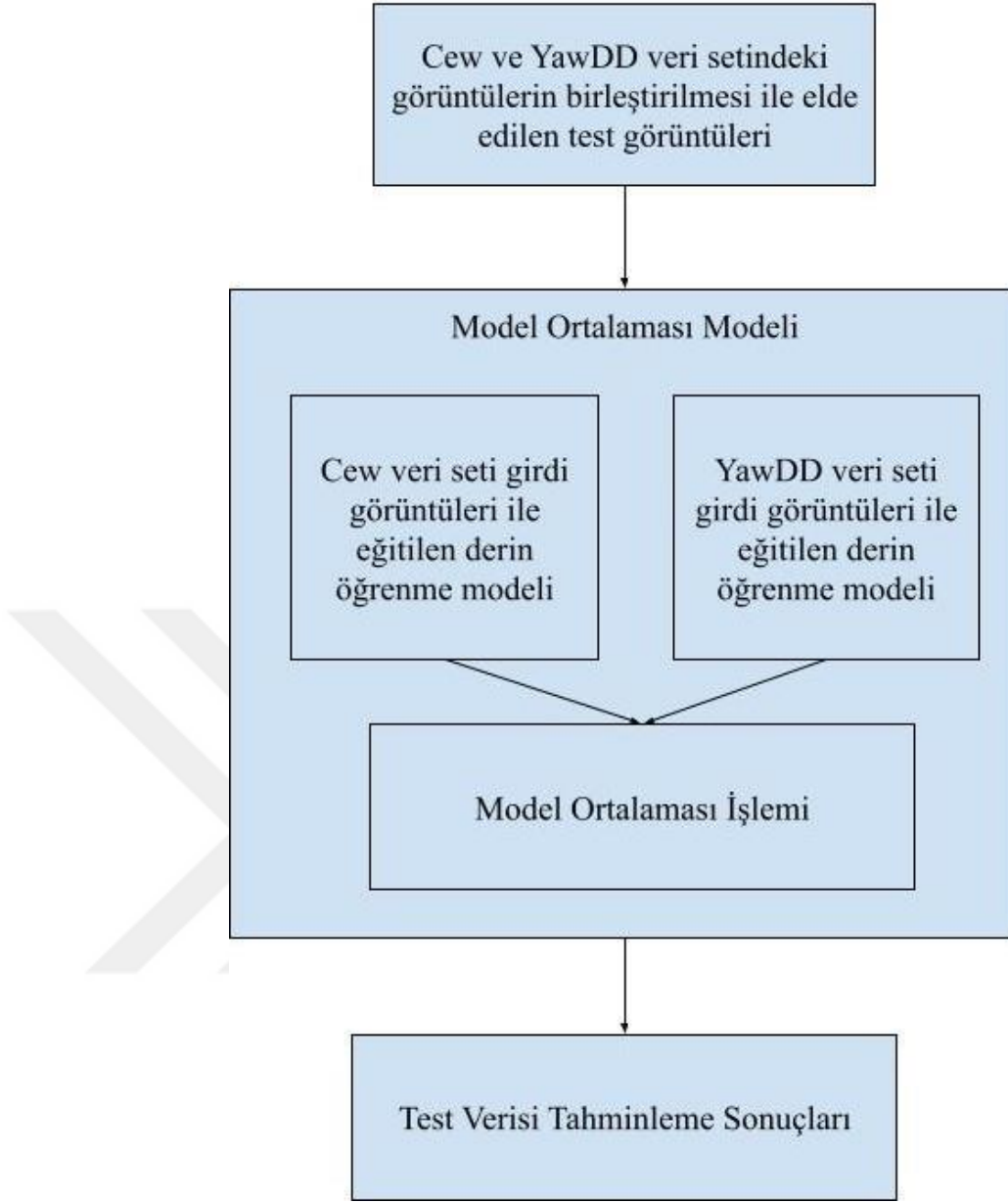
Veri Seti	Görüntü Boyutu (piksel)	Sınıf	Kesinlik	Duyarlılık	F-Ölçütü	Doğruluk (%)
YawDD Yüz	100	Esneyen	0,98	0,95	0,97	97,25
		Esname-yen	0,96	0,98	0,97	
YawDD Ağız	48	Esneyen	1,00	0,96	0,96	98,03
		Esname-yen	0,97	1,00	0,98	
YawDD Ağız, Yüz	48 100	Esneyen	0,98	0,98	0,98	97,64
		Esname-yen	0,98	0,98	0,98	
YawDD Sol Göz, Sağ Göz, Ağız, Yüz	24 24 48 100	Esneyen	1,00	0,96	0,98	98,03
		Esname-yen	0,97	1,00	0,98	

Tablo 5.12’de yer alan sonuçlara göre eğitim verileri üzerindeki en yüksek başarı değeri sol göz, sağ göz, ağız ve yüz özniteliklerinin birleşiminden oluşan modelde, doğrulama verileri üzerindeki en yüksek başarı değeri ise ağız girdi görüntü parçasının girdi olarak kullanıldığı tekil modelde elde edilmiştir. Tablo

5.13'te yer alan test verileri üzerinde de bu iki model %98,03 doğruluk değeri ile en yüksek başarı değerini paylaşmaktadır.

5.10. Cew ve YawDD Veri Setlerinin Birleştirilmesi ile Elde Edilen Deneysel Sonuçlar

Tez çalışmasının bu bölümünde, önerilen yorgunluk tanıma yöntemlerinin Cew ve YawDD veri seti üzerinde esneyen ve kapalı göz verileri, esneyen ve gözleri kapalı olan yüz görüntüleri, esnemeyen ve açık göz görüntüleri, esnemeyen ve gözleri açık yüz görüntüleri birleştirilerek test verileri oluşturulmuştur. Test veri seti, 240 adet esneyen ve gözleri kapalı olan yüz görüntüsü, 251 adet esnemeyen ve gözleri açık yüz görüntüsü, 480 adet esneyen ve kapalı göz görüntüsü, 502 adet esnemeyen ve açık göz görüntüsü olmak üzere toplamda 1473 adet görüntü içermektedir. Önerilen yöntem, benzer özelliklerde eğitilen modellerin ilgili özelliğe ait veriler üzerinde model ortalaması topluluk yöntemi kullanılarak Cew ve YawDD test verileri üzerinde test edilmesiyle elde edilen sınıflandırma sonuçları bu bölümde verilmiştir. Her iki veri setinde benzer özelliklerin (24x24, 48x48 boyutlarındaki göz; 100x100 boyutlarındaki yüz; sırasıyla 24x24, 24x24, 100x100 boyutlarındaki sol göz, sağ göz ve yüz; sırasıyla 48x48, 48x48, 100x100 boyutlarındaki sol göz, sağ göz ve yüz için) birleştirilmesi için gerçekleştirilen model ortalaması modeli Şekil 5.14'te yer almaktadır.



Şekil 5.14. Cew ve YawDD veri setlerinde benzer özelliklerin birleştirilmesi için geliştirilen model ortalaması modeli.

Model ortalama topluluk öğrenmesi yöntemi ile test verileri üzerinde yapılan testlerde elde edilen doğruluk değerleri ile standart sapma değerleri Tablo 5.14’te sunulmuştur.

Tablo 5.14. Cew ve YawDD veri setlerinin test görüntüleri üzerindeki model ortalama yöntemi sınıflandırma sonuçları.

Model Özellikleri				Test Sonuçları	
Veri Seti	Model Türü	Girdi Özelliği	Görüntü Boyutu (piksel)	Doğruluk Değeri (%)	Standart Sapma Değeri
Cew	Sıralı	Göz	24	86,4	0,005
YawDD					
Cew	Sıralı	Göz	48	84,4	0,001
YawDD					
Cew	Sıralı	Yüz	100	83,7	0,004
YawDD					
Cew	Fonksiyonel	Sol Göz, Sağ Göz, Yüz	24 24 100	81,0	0,027
YawDD					
Cew	Fonksiyonel	Sol Göz, Sağ Göz, Yüz	48 48 100	70,0	0,052
YawDD					

Veri setlerindeki yüz girdi görüntüleri ile eğitilen sıralı derin öğrenme modellerinin model ortalama yöntemi kullanılarak test verileri üzerinde yapılan sınıflandırma sonucu doğruluk değeri %83,7 olarak elde edilmiştir.

24x24 piksel boyutundaki göz girdi görüntüleri ile eğitilen sıralı derin öğrenme modellerinin model ortalama yöntemi kullanılarak test verileri üzerinde yapılan sınıflandırma sonucu doğruluk değeri %86,4 olarak elde edilmiştir.

48x48 piksel boyutundaki göz girdi görüntüleri ile eğitilen sıralı derin öğrenme modellerinin model ortalama yöntemi kullanılarak test verileri üzerinde yapılan sınıflandırma sonucu doğruluk değeri %84,4 olarak elde edilmiştir.

Yüz, 24x24 piksel boyutundaki sol göz ve sağ göz görüntüleri ile eğitilen fonksiyonel derin öğrenme modellerinin model ortalama yöntemi kullanılarak test

verileri üzerinde yapılan sınıflandırma sonucu doğruluk değeri %81,0 olarak elde edilmiştir.

Yüz, 48x48 piksel boyutundaki sol göz ve sağ göz görüntüleri ile eğitilen fonksiyonel derin öğrenme modellerinin model ortalaması yöntemi kullanılarak test verileri üzerinde yapılan sınıflandırma sonucu doğruluk değeri %70,0 olarak elde edilmiştir.

Tablo 5.14'te yer alan Cew ve YawDD veri setlerinin birleştirilmesi ile elde edilen test verilerindeki en iyi doğruluk başarısı, sıralı modellerde 24x24 piksel boyutlarındaki göz görüntülerinde %86,4 olarak, fonksiyonel modellerde ise sırasıyla 24x24, 24x24, 100x100 piksel boyutlarındaki sol göz, sağ göz ve yüz görüntülerinde %81,0 olarak elde edilmiştir.

Tablo 5.14 ile Tablo 5.11'daki sıralı modellerde göz test görüntüleri üzerindeki doğruluk başarı değeri karşılaştırıldığında, model ortalaması yönteminin kullanılmasıyla YawDD veri setinde 24x24 piksel boyutlarındaki göz görüntülerindeki test doğruluk değerinin %83,07'den %86,4'e yükseldiği, 48x48 piksel boyutlarındaki göz görüntülerindeki test doğruluk değerinin ise %79,53'ten %84,4'e yükseldiği görülmektedir. Model ortalaması yönteminin kullanılmasıyla YawDD veri setindeki göz görüntüleri üzerinde gerçekleştirilen sıralı modellerin test verileri üzerindeki doğruluk başarı değerinde daha iyi sonuçlar elde edildiği görülmektedir. Ancak elde edilen başarı değerleri sadece Cew veri setindeki göz görüntüleri ile elde edilenlerden daha düşüktür.

5.11. Önerilen Yöntemin Literatürdeki Çalışmalar ile Karşılaştırılması

Tez süresi boyunca geliştirilen yöntemler arasında YawDD veri seti üzerinde yüz, gözler ve ağız görüntülerinin birlikte değerlendirildiği fonksiyonel model ile ağız görüntülerinin değerlendirildiği sıralı model, en iyi sonucu veren modeller olarak görülmektedir. Literatürdeki çalışmalar ile yapılan karşılaştırma, kullandıkları veri seti, sınıflandırma modeli ve doğruluk başarı ölçütüne göre Tablo 5.15 ve 5.16'da sunulmuştur. YawDD veri seti için yapılan karşılaştırmalarda

verilen sonuçlar (YawDD göz girdi görüntü parçaları için olan sonuç hariç), tekrarlı deneysel çalışmalarımızla elde ettiğimiz sonuçlardır.

Tablo 5.15. Önerilen sıralı modellerin literatürdeki çalışmalar ile karşılaştırılması.

Çalışma	Yöntem	Model	Veri seti	Sonuç (Doğruluk) (%)
Bu çalışmada önerilen modeller	Derin Öğrenme	CNN	Cew Göz	98,35
			Cew Yüz	95,33
			YawDD Göz	83,07
			YawDD Ağız	98,03
Tümen vd. (2018)	Derin Öğrenme	CNN	Cew Göz	92,99–96,5
Alparslan et al. (2020)	Derin Öğrenme	CNN	Cew Yüz	92
Ji et al. (2020)	Derin Öğrenme	CNN	Cew Göz	97,87
			YawDD Göz	98,2
			YawDD Ağız	94,7
Xie et al. (2018)	Derin Öğrenme	CNN (Inception-v3), LSTM	YawDD, NTHU- DDD (Ağız)	98,4

Tablo 5.16. Önerilen YawDD veri seti için en başarılı fonksiyonel modelin literatürdeki çalışmalar ile karşılaştırılması.

Çalışma	Yöntem	Model	Veri Seti	Sonuç (Doğruluk) (%)
Bu çalışmada önerilen yöntem	Derin Öğrenme	CNN	YawDD (Yüz, Sol Göz, Sağ Göz, Ağız)	98,03
Shen et al. (2020)	Derin Öğrenme	Optical Flow, CLAHE, CNN	NTHU- DDD (Yüz, Gözler, Ağız)	94,46

Cew veri setinde eğitim ve test olarak bir ayırım söz konusu olmadığından eğitim, doğrulama ve test için verilerin ayrıştırılması rastgele seçim ile belirlenmiştir. YawDD veri setinde ise veriler eğitim ve test olarak ayrılmış durumdadır. Eğitim verilerinde doğrulama verileri için Scikit-learn kütüphanesinin `train_test_split()` fonksiyonu kullanılarak %15 oranında eğitim verilerinden doğrulama verileri oluşturulmuştur.

Videolardan durağan görüntüler elde edilirken yapılan seçim işlemi; görüntüleri yüz, göz, ağız gibi parçalarına ayırırken kullanılan yöntem vb. aşamalarda veri setlerinde oluşan farklılıklardan dolayı hassas bir başarımla karşılaştırması yapmak mümkün olamamaktadır. Literatürdeki farklı çalışmalarda da farklı yaklaşımlar bulunması da karşılaştırmayı zorlaştırmaktadır. Ancak durağan görüntülerde, yorgunluk ve esneme veri setlerindeki yüz, göz ve ağız görüntü parçalarının birleştirildiği modeller oluşturularak deneysel çalışmalar yapılmıştır. Her iki veri seti için de görüntü veri setleri ve görüntü parçalarının melezleştirilmesi ile, durağan görüntüler kullanıldığında literatürdeki çalışmalarla yarışabilir düzeyde sonuçlar elde edilmiştir. Ayrıca, hazır veri setlerini kullanarak oluşturduğumuz veri setlerimiz sabit tutulduğunda, bazı durumlarda, görüntü parçaları özneliklerini birleştirme yaklaşımımız ile, tekil görüntülerden elde edilen en iyi başarı değerinden daha yüksek başarılar elde edilebildiği de gözlemlenmiştir.

6. SONUÇ VE ÖNERİLER

Bu tez çalışmasında maddi ve manevi kayıpların yaşandığı trafik kazalarını önlemek; çalışan veya öğrencilerin durumlarını belirleyip verimlerini artırmak amacıyla görüntüler üzerinden yorgunluk/uykusuzluk tanıma konusunda derin öğrenme ve topluluk öğrenmesi yöntemlerinin birlikte kullanıldığı modeller önerilerek geliştirilmiştir.

Kişilerde, yorgunluk/uykusuzluk belirtileri olarak en belirgin özellikler gözlerin kapanması ve esnemedir. Bu tez çalışmasında görüntüler üzerinde yüz, göz ve ağız bölgelerine odaklanılarak gri tonlamalı görüntüler üzerinde CNN mimarisine dayalı derin öğrenme modelleri geliştirilmiştir. Yüz, gözler ve ağız özelliklerinin birlikte değerlendirilerek yorgunluk tanıma alanında başarıyı artırmak amaçlı yüz, sol göz ve sağ göz görüntülerinin birlikte eğitildiği, yüz ve ağız görüntülerinin birlikte eğitildiği; yüz, sol göz, sağ göz ve ağız görüntülerinin birlikte eğitildiği derin öğrenme modelleri geliştirilmiştir. Geliştirilen modellerin durağan görüntüler üzerinde işletilebilir ve birden fazla özellik kümesinin birlikte kullanılabilir olması açısından yorgunluk tanıma alanındaki ilk çalışma olması sebebiyle literatüre katkısı bulunmaktadır.

Bu çalışmada önerilen yöntemler yorgunluk/uykusuzluk tanıma alanında sıkça kullanılan, farklı etnik kökenden ve farklı cinsiyetten kişilere ait farklı aydınlatma koşulları ve bulanıklıkta görüntüler içeren iki farklı veri seti üzerinde test edilmiştir. Cew veri setinde ve YawDD veri setinde model eğitimleri birer kez gerçekleştirilmiştir. YawDD veri setinde daha başarılı sonuçlar elde edildiğinden test verileri üzerinde en yüksek doğruluk değerine sahip modellerin eğitimleri için üçer tekrar daha gerçekleştirilip dört deneyin sonuçlarının ortalaması alınmıştır. Tekrarlanan deneylerden sonra önerilen modelin YawDD veri setinde yapılan eğitim verileri üzerinde %99,71; doğrulama verileri üzerinde %95,90; test verileri üzerinde %98,03 doğruluk değeri elde edilmiştir. Cew veri setinde ise orijinal görüntü boyutlarındaki (sırasıyla 100x100, 24x24, 24x24 piksel) yüz, sol göz ve sağ göz görüntülerinin birlikte değerlendirildiği derin öğrenme modelinde eğitim verileri üzerinde %99,71; doğrulama verileri üzerinde %94,78; test verileri üzerinde %96,98 doğruluk değeri elde edilmiştir. Model ortalaması yönteminin

kullanılmasıyla göz görüntüleri üzerindeki tekil modelin test verilerinde elde edilen doğruluk başarı değerlerinde, Cew veri setindeki değerden daha düşük sonuçlar elde edilmiş fakat YawDD veri setindeki değerden daha başarılı sonuçlar alınmıştır. Yapılan deneysel çalışmalardaki sonuçlar Cew veri setinde göz özniteliklerinin, YawDD veri setinde ise ağız öznitelikliğin sınıflandırma başarısında daha etkili olduğunu göstermektedir. Ayrıca, önerilen yöntemler değerlendirildiğinde bazı durumlarda yüz, gözler ve ağız özniteliklerinin birleştirilmesinin yorgunluk tanımada sınıflandırma başarısını artırabildiği görülmektedir. Her ne kadar büyük artışlar söz konusu olmasa da umut verici sonuçlar alınmıştır. Yaklaşımlarımıza / önerdiğimiz sisteme daha farklı yöntemler de eklenerek ve farklı hiperparametre eniyilemesi metotlarının uygulanması ile performansın daha da yükseltilebileceği düşünülmektedir.

Tez çalışması kapsamında sıralı ve fonksiyonel modellerdeki katman sayıları değiştirilerek her iki veri seti üzerinde deneysel çalışmalar yapılmış ve en iyi sonucun hem sıralı hem fonksiyonel modeller için 4 adet evrişim katmanı, 2 adet tam bağlantı katmanı, 1 adet çıktı katmanından oluşan modellerden elde edildiği gözlemlenmiştir.

Gelecek çalışmalarda, kendinden eğitim ve test olarak ayrılmamış veri setinde farklı veri ayırma yöntemlerinin denenmesi, derin öğrenme modellerinin gerçekleştiriminde tekrar sayılarının artırılması ve hiperparametre eniyilemesi yöntemlerinin denenmesiyle başarımın ve ölçüm hassasiyetinin artırılması hedeflenmektedir. Ayrıca, önerilen yöntemlerin farklı uygulama alanlarındaki farklı veri setleri üzerinde de uygulanması planlanmıştır.

KAYNAKLAR DİZİNİ

- Abtahi, S., Omidyeganeh, M., Shirmohammadi, S. and Hariri, B.,** 2014, March, YawDD: A yawning detection dataset. In *Proceedings of the 5th ACM multimedia systems conference* (pp. 24-28).
- Albawi, S., Mohammed, T. A. and Al-Zawi, S.,** 2017, Understanding of a convolutional neural network. In *2017 International Conference on Engineering and Technology (ICET)* (pp. 1-6). IEEE.
- Alparslan, K., Alparslan, Y. and Burlick, M.,** 2020, Towards Evaluating Driver Fatigue with Robust Deep Learning Models. *arXiv preprint arXiv:2007.08453*.
- Alpaydin, E.,** 2010, Introduction to machine learning, The MIT Press, London, 539p.
- Amini, S. M. and Parmeter, C. F.,** 2012, Comparison of model averaging techniques: Assessing growth determinants. *Journal of Applied Econometrics*, 27(5), 870-876.
- Baş, F.,** 2015, Geçici Fiziksel Özelliklerden Yorgunluk ve Uykusuzluğun Sürücü Davranışları Üzerine Etkisi, Yüksek Lisans Tezi, Atatürk Üniversitesi Fen Bilimleri Enstitüsü, 71s.
- Beauxis-Aussalet, E. and Hardman, L.,** 2014, Simplifying the visualization of confusion matrix. In *26th Benelux Conference on Artificial Intelligence (BNAIC)*.
- Bhukya, R.,** 2021, Exploring Machine Learning: A Beginners Perspective, Horizon Books, 254p.
- Brownlee, J.,** 2018, *Better deep learning: train faster, reduce overfitting, and make better predictions*. Machine Learning Mastery.
- Celona, L., Mammana, L., Bianco, S. and Schettini, R.,** 2018, September, A Multi-Task CNN Framework for Driver Face Monitoring. In *2018 IEEE 8th International Conference on Consumer Electronics-Berlin (ICCE-Berlin)* (pp. 1-4). IEEE.

KAYNAKLAR DİZİNİ (devam)

- Choi, I. H., Hong, S. K. and Kim, Y. G.,** 2016, January, Real-time categorization of driver's gaze zone using the deep learning techniques. In *2016 International Conference on Big Data and Smart Computing (BigComp)* (pp. 143-148). IEEE.
- Coşkun, C. ve Baykal, A.,** 2011, Veri madenciliğinde sınıflandırma algoritmalarının bir örnek üzerinde karşılaştırılması. *Akademik Bilişim, 2011*, 1-8.
- Dalianis, H.,** 2018, *Clinical text mining: Secondary use of electronic patient records*. Springer Nature, 181p.
- Dietterich, T. G.,** 2002, Ensemble learning. *The handbook of brain theory and neural networks*, 2(1), 110-125.
- Doğan, F. ve Türkoğlu, İ.,** 2019, Derin öğrenme modelleri ve uygulama alanlarına ilişkin bir derleme. *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, 10(2), 409-445.
- Dwivedi, K., Biswaranjan, K. and Sethi, A.,** 2014, February, Drowsy driver detection using representation learning. In *2014 IEEE International Advance Computing Conference (IACC)* (pp. 995-999). IEEE.
- Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S. and Lew, M. S.,** 2016, Deep learning for visual understanding: A review. *Neurocomputing*, 187, 27-48.
- Güzel, K.** 2018., “Kolektif Öğrenme (Ensemble Learning) ve Bagging Algoritması. “, <https://medium.com/@billmuhh/kolektif-%C3%B6%C4%9Frenme-ve-bagging-algoritmas%C4%B1-e8ea3d932b72> (Erişim Tarihi: 19 Ağustos 2019)
- Hossin, M. and Sulaiman, M. N.,** 2015, A review on evaluation metrics for data classification evaluations. *International journal of data mining & knowledge management process*, 5(2), 1.
- Iansiti, M. ve Lakhani, K. R.,** 2021, Optimist Yayın Grubu, Yapay Zekâ Çağında Rekabet, 296s.

KAYNAKLAR DİZİNİ (devam)

- İnik, Ö. ve Ülker, E.,** 2017, Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri. *Gaziosmanpaşa Bilimsel Araştırma Dergisi*, 6(3), 85-104.
- Jabbar, R., Al-Khalifa, K., Kharbeche, M., Alhajyaseen, W., Jafari, M. and Jiang, S.,** 2018, Real-time driver drowsiness detection for android application using deep neural networks techniques. *Procedia computer science*, 130, 400-407.
- Ji, Y., Wang, S., Zhao, Y., Wei, J. and Lu, Y.,** 2019, Fatigue state detection based on multi-index fusion and state recognition network. *IEEE Access*, 7, 64136-64147.
- Kayaalp, K. ve Süzen, A. A.,** 2018, Derin Öğrenme ve Türkiye'deki Uygulamaları, İksad Yayınevi, Ankara, 89s.
- Khabaralak, K. and Koriashkina, L.,** 2021, Fast Facial Landmark Detection and Applications: A Survey. *arXiv preprint arXiv:2101.10808*.
- Labach, A., Salehinejad, H. and Valaee, S.,** 2019, Survey of dropout methods for deep neural networks. *arXiv preprint arXiv:1904.13310*.
- Laouz, Hamed,** 2019, Driver's Drowsiness and fatigue détection system, 65p.
- Li, X., Yang, Z. and Wu, H.,** 2020, Face detection based on receptive field enhanced multi-task cascaded convolutional neural networks. *IEEE Access*, 8, 174922-174930.
- Liu, F., Li, X., Lv, T. and Xu, F.,** 2019, January, A Review of Driver Fatigue Detection: Progress and Prospect. In *2019 IEEE International Conference on Consumer Electronics (ICCE)* (pp. 1-6). IEEE.
- Liu, W., Qian, J., Yao, Z., Jiao, X. and Pan, J.,** 2019, Convolutional Two-Stream Network Using Multi-Facial Feature Fusion for Driver Fatigue Detection. *Future Internet*, 11(5), 115.
- Mahesh, B.,** 2020, Machine Learning Algorithms-A Review. *International Journal of Science and Research (IJSR)*, 381-386, 9p.

KAYNAKLAR DİZİNİ (devam)

- Mahpod, S., Das, R., Maiorana, E., Keller, Y. and Campisi, P.,** 2021, Facial landmarks localization using cascaded neural networks. *Computer Vision and Image Understanding*, 205, 103171.
- Metlek, S. ve Çetiner, H.,** 2021, Matlab Ortamında Derin Öğrenme Uygulamaları, İksad Yayınevi, Ankara, 80s.
- Mitchell, T.,** 2006, The discipline of machine learning, Technical Report, School of Computer Science Carnegie Mellon University, 7p.
- Mu, R. and Zeng, X.,** 2019, A review of deep learning research. *KSII Transactions on Internet and Information Systems (TIIS)*, 13(4), 1738-1764.
- Müller, A. C. and Guido, S.,** 2016, *Introduction to machine learning with Python: a guide for data scientists*. "O'Reilly Media, Inc."
- Patel, A.** “Ensemble Learning- The heart of Machine learning. “, <https://medium.com/ml-research-lab/ensemble-learning-the-heart-of-machine-learning-b4f59a5f9777> (Erişim Tarihi: 19 Ağustos 2019)
- Raju, S.,** 2020, “yawn_eye_dataset_new”, https://www.kaggle.com/serenaraju/yawn-eye-dataset-new?select=dataset_new (Erişim Tarihi: 9 Ocak 2021)
- Sagi, O. and Rokach, L.,** 2018, Ensemble learning: A survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4), e1249.
- Shahar, M., Das, R., Maiorana, E. and Campisi, P.,** 2018, Facial Landmark Point Localization using Coarse-to-Fine Deep Recurrent Neural Network, 10p.
- Shakeel, M. F., Bajwa, N. A., Anwaar, A. M., Sohail, A. and Khan, A.,** 2019, June, Detecting Driver Drowsiness in Real Time Through Deep Learning Based Object Detection. In *International Work-Conference on Artificial Neural Networks* (pp. 283-296). Springer, Cham.
- Shams, A., Nasim, B. and Faysal, I.,** 2019, *Analysis on face recognition based on five different viewpoint of face images using MTCNN and FaceNet* (Doctoral dissertation, Brac University).

KAYNAKLAR DİZİNİ (devam)

- Shen, Q., Zhao, S., Zhang, R. and Zhang, B.,** 2020, Robust Two-Stream Multi-Features Network for Driver Drowsiness Detection. In *Proceedings of the 2020 2nd International Conference on Robotics, Intelligent Control and Artificial Intelligence* (pp. 271-277).
- Shinde, P. P. and Shah, S.,** 2018, A Review of Machine Learning and Deep Learning Applications. In *2018 Fourth international conference on computing communication control and automation (ICCUBE)* (pp. 1-6). IEEE.
- Song, F., Tan, X., Liu, X. and Chen, S.,** 2014, Eyes closeness detection from still images with multi-scale histograms of principal oriented gradients. *Pattern Recognition*, 47(9), 2825-2838.
- Şeker, A., Diri, B. ve Balık, H. H.,** 2017, Derin Öğrenme Yöntemleri ve Uygulamaları Hakkında Bir İnceleme. *Gazi Mühendislik Bilimleri Dergisi (GMBD)*, 3(3), 47-64.
- Tümen, V., Yıldırım, Ö. and Ergen, B.,** 2018, April, Detection of driver drowsiness in driving environment using deep learning methods. In *2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT)* (pp. 1-5). IEEE.
- Ursavaş, A. ve Ege, E.,** 2004, Uyku apne sendromu ve trafik kazaları. *Uludağ Üniversitesi Tıp Fakültesi Dergisi*, 30(1), 37-41.
- Valliani, A. A. A., Ranti, D. and Oermann, E. K.,** 2019, Deep learning and neurology: a systematic review. *Neurology and therapy*, 8(2), 351-365.
- Vargas, R., Mosavi, A. and Ruiz, R.,** 2017, Deep learning: a review, *Advances in Intelligent Systems and Computing*, 10p.
- Voulodimos, A., Doulamis, N., Doulamis, A. and Protopapadakis, E.,** 2018, Deep learning for computer vision: A brief review. *Computational intelligence and neuroscience*, 13p.
- Wu, Y. and Ji, Q.,** 2019, Facial landmark detection: A literature survey. *International Journal of Computer Vision*, 127(2), 115-142.

KAYNAKLAR DİZİNİ (devam)

- Xie, Y., Chen, K. and Murphey, Y. L.**, 2018, November, Real-time and Robust Driver Yawning Detection with Deep Neural Networks. In *2018 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 532-538). IEEE.
- Yadawadkar, S., Mayer, B., Lokegaonkar, S., Islam, M. R., Ramakrishnan, N., Song, M. and Mollenhauer, M.**, 2018, December, Identifying Distracted and Drowsy Drivers Using Naturalistic Driving Data. In *2018 IEEE International Conference on Big Data (Big Data)* (pp. 2019-2026). IEEE.
- Zhang, K., Zhang, Z., Li, Z. and Qiao, Y.**, 2016, Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 23(10), 1499-1503.

TEŞEKKÜR

Tez çalışmam boyunca bilgi, yardım, görüş ve fikirlerini esirgemeyen, her fırsatta kıymetli vaktini ayırarak çalışmamla yakından ilgilenip öneride bulunan ve yönlendirmeleriyle ilerlememi, gelişmemi sağlayan saygıdeğer hocam Prof. Dr. Aybars UĞUR'a içten teşekkürlerimi sunarım.

Tezin biçimlenmesinde katkıda bulunan, değerli vaktini ayırıp yardımını esirgemeyen Bursa Teknik Üniversitesi Ar. Gör. Esma İBİŞ'e teşekkür ederim.

Çalışma hayatım ve bu tez çalışmam süresince desteğini hissettiğim başta Yazılım Müdürü'm Sayın Dr. Fatih DAĞDEVİREN ve Proje Yönetici'm Cihan BİRCAN olmak üzere tüm UNİPA A.Ş. ailesine teşekkür ederim.

Tüm eğitim öğretim hayatım süresince her zaman destek olan ve tez çalışma sürecimde de gösterdikleri fedakarlıklardan dolayı başta annem Esin KÖK olmak üzere tüm aileme teşekkür ederim.

01 / 09 / 2021

Rabia KÖK

ÖZGEÇMİŞ

Kişisel Bilgiler

Soyadı: KÖK

Adı: Rabia

Eğitim Bilgileri

Yüksek Lisans: Ege Üniversitesi, Bilgisayar Mühendisliği Bölümü, 2018 - Devam ediyor

Lisans: Ege Üniversitesi, Bilgisayar Mühendisliği Bölümü, 2012 - 2017

İş Deneyimi

Stajyer: Veriyaz Yazılım Ltd. Şti., 2016

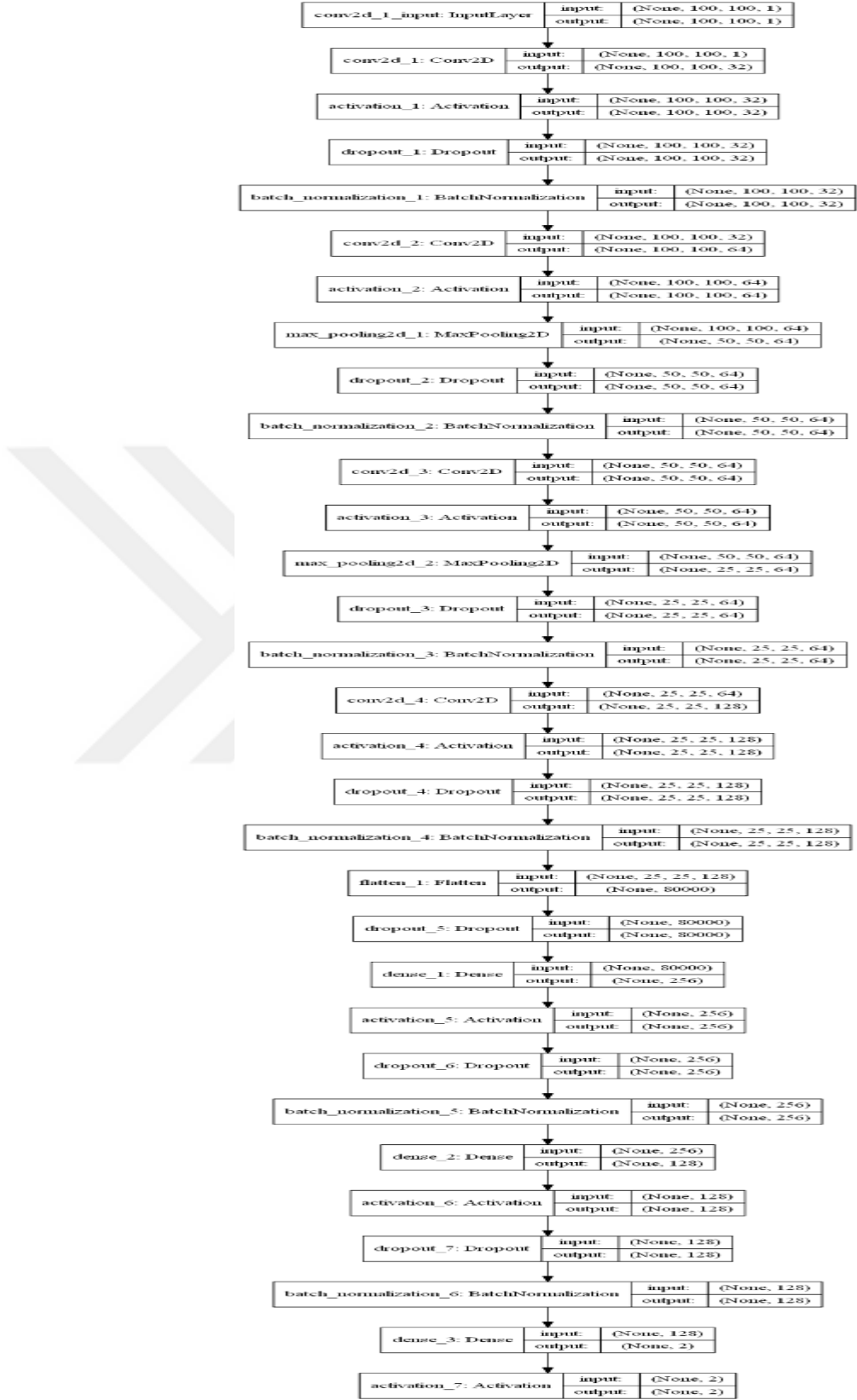
Yazılım Stajyeri: UNİPA A.Ş., Aralık 2019 - Haziran 2020

Yazılım Uzmanı: UNİPA A.Ş., Temmuz 2020 - Devam ediyor

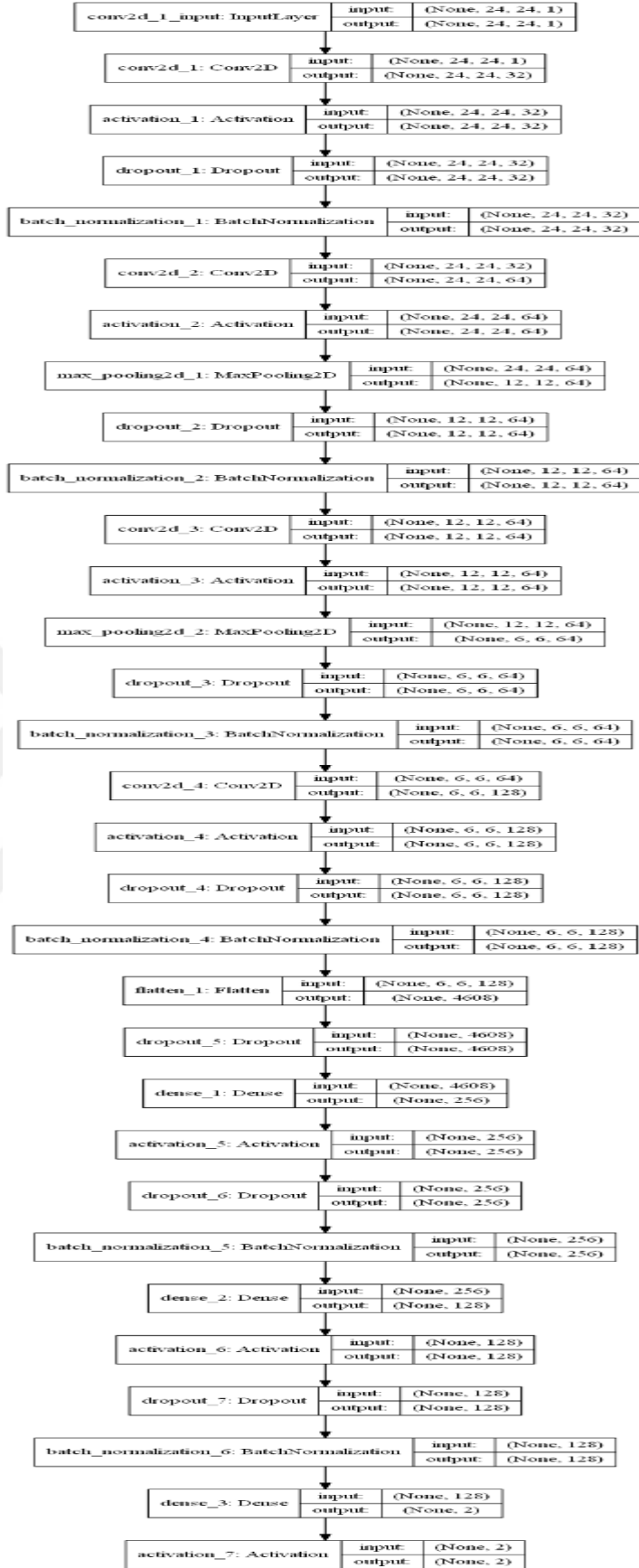
EKLER

- Şekil A.1.** Cew veri seti yüz girdi görüntüleri için geliştirilen sıralı model katmanları.
- Şekil A.2.** Cew veri seti 24x24 piksel boyutundaki göz girdi görüntüleri için geliştirilen sıralı model katmanları.
- Şekil A.3.** Cew veri seti 48x48 piksel boyutundaki göz girdi görüntüleri için geliştirilen sıralı model katmanları.
- Şekil A.4.** YawDD veri seti yüz girdi görüntüleri için geliştirilen sıralı model katmanları.
- Şekil A.5.** YawDD veri seti ağız girdi görüntüleri için geliştirilen sıralı model katmanları.
- Şekil A.6.** YawDD veri seti 24x24 piksel boyutundaki göz girdi görüntüleri için geliştirilen sıralı model katmanları.
- Şekil A.7.** YawDD veri seti 48x48 piksel boyutundaki göz girdi görüntüleri için geliştirilen sıralı model katmanları.
- Şekil A.8.** Cew veri seti yüz ve 24x24 piksel boyutundaki göz girdi görüntüleri için geliştirilen fonksiyonel model katmanları.
- Şekil A.9.** Cew veri seti yüz ve 48x48 piksel boyutundaki göz girdi görüntüleri için geliştirilen fonksiyonel model katmanları.
- Şekil A.10.** YawDD veri seti yüz ve 24x24 piksel boyutundaki göz girdi görüntüleri için geliştirilen fonksiyonel model katmanları.
- Şekil A.11.** YawDD veri seti yüz ve 48x48 piksel boyutundaki göz girdi görüntüleri için geliştirilen fonksiyonel model katmanları.
- Şekil A.12.** YawDD veri seti yüz, 24x24 piksel boyutundaki göz ,48x48 piksel boyutundaki ağız girdi görüntüleri için geliştirilen fonksiyonel model katmanları.
- Şekil A.13.** YawDD veri seti yüz, 48x48 piksel boyutundaki ağız girdi görüntüleri için geliştirilen fonksiyonel model katmanları.

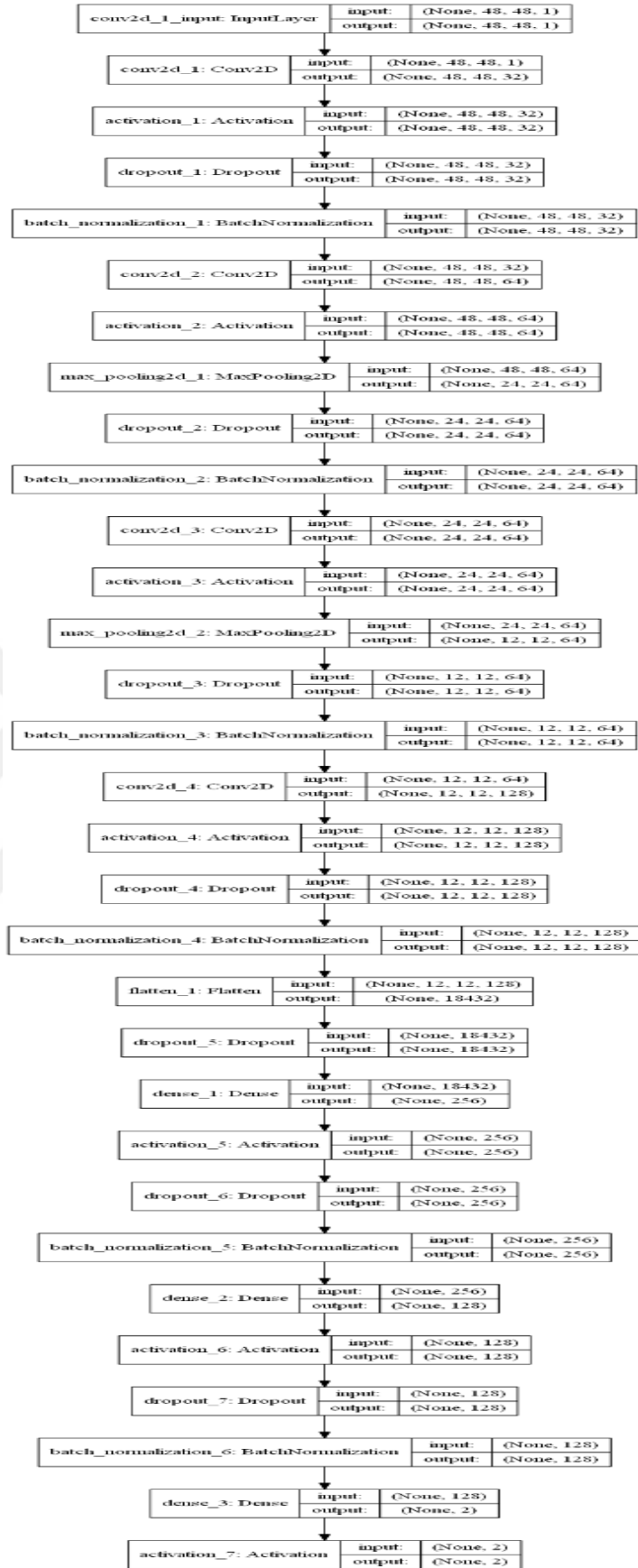
EK Açıklamalar - A



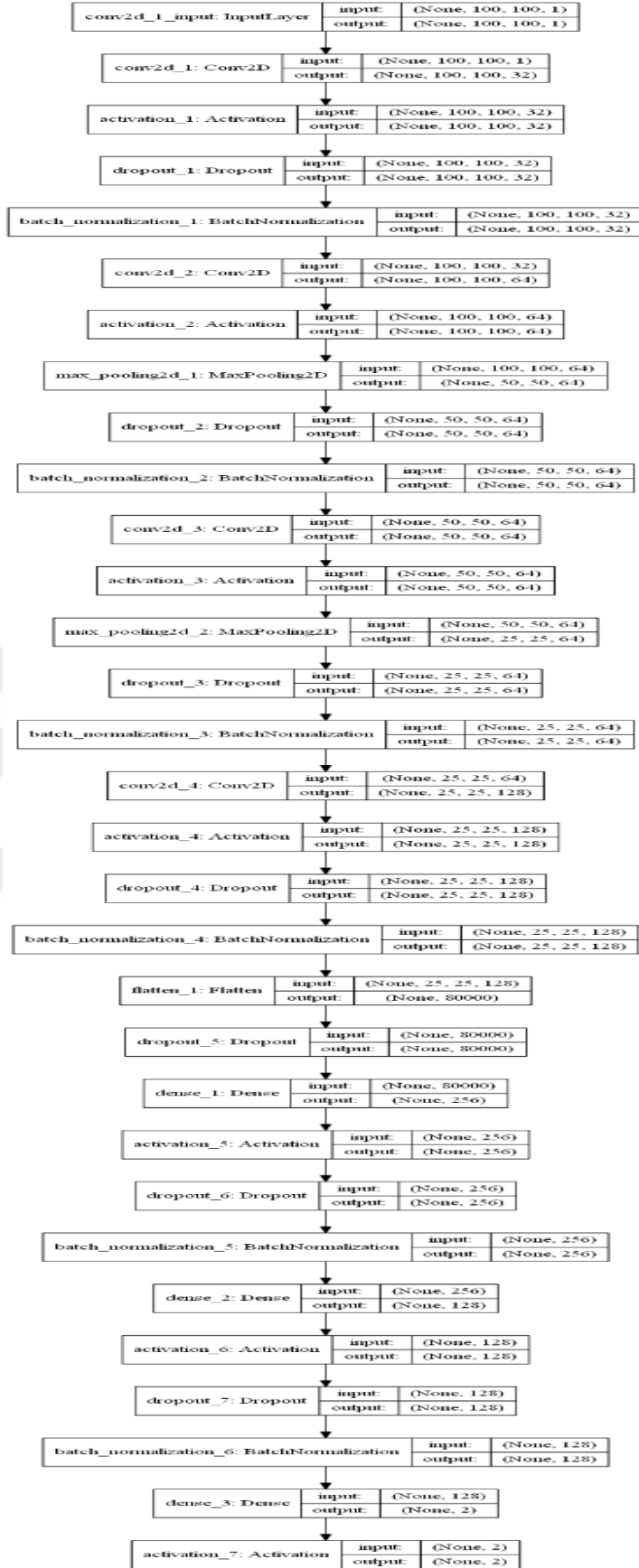
Şekil A.1. Cew veri seti yüz girdi görüntüleri için geliştirilen sıralı model katmanları.



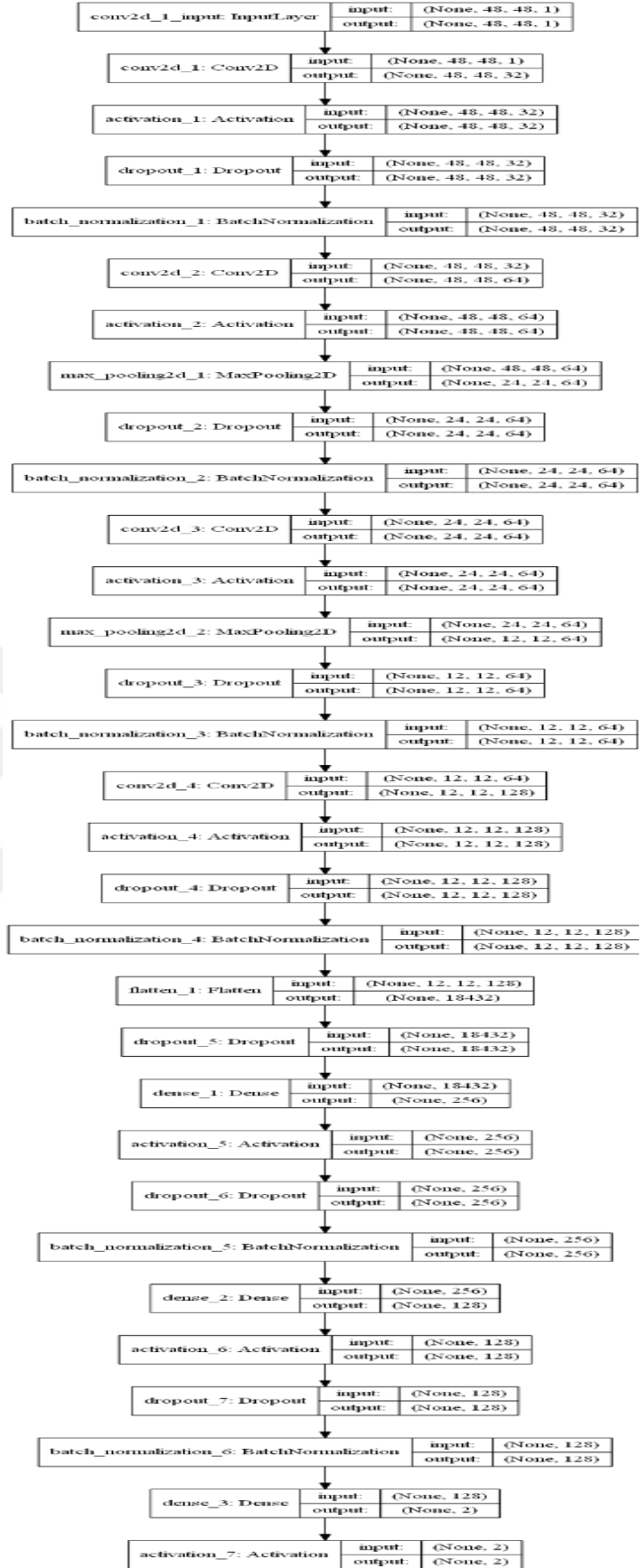
Şekil A.2. Cew veri seti 24x24 piksel boyutundaki göz girdi görüntüleri için geliştirilen sıralı model katmanları.



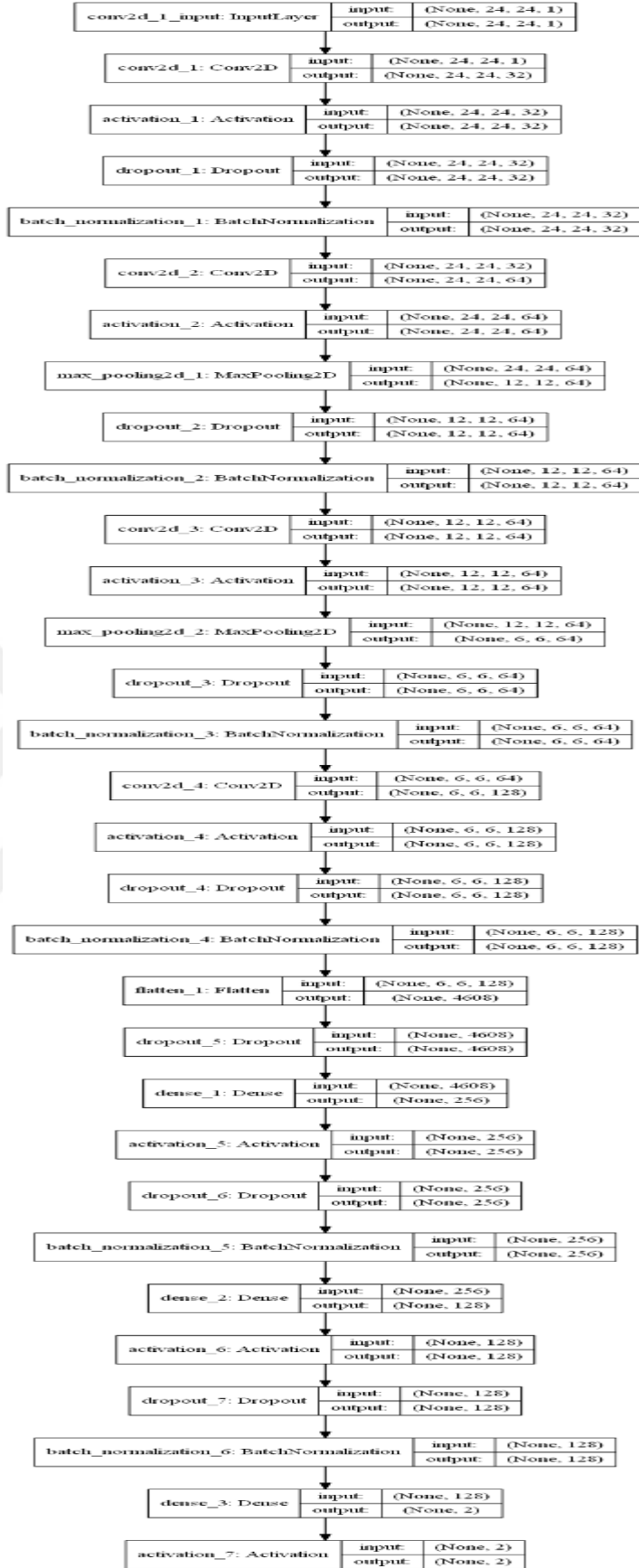
Şekil A.3. Cew veri seti 48x48 piksel boyutundaki göz girdi görüntüleri için geliştirilen sıralı model katmanları.



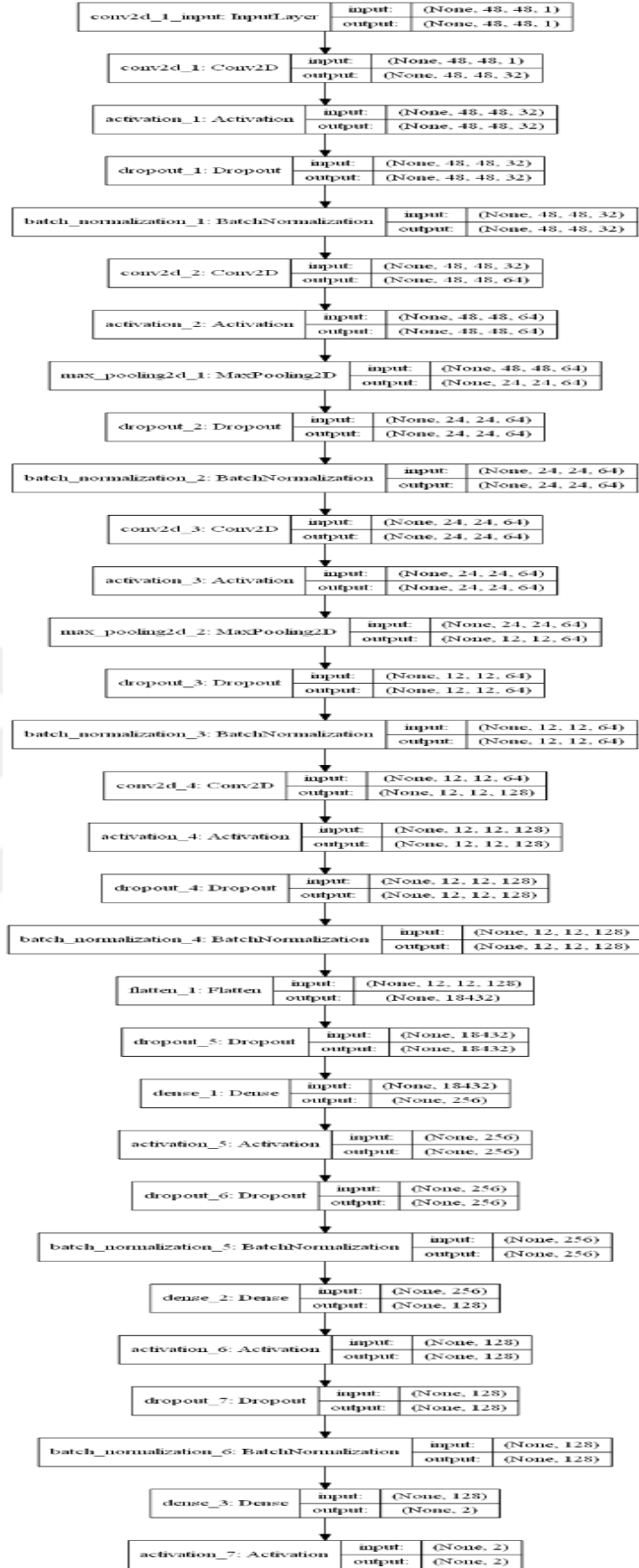
Şekil A.4. YawDD veri seti yüz girdi görüntüleri için geliştirilen sıralı model katmanları .



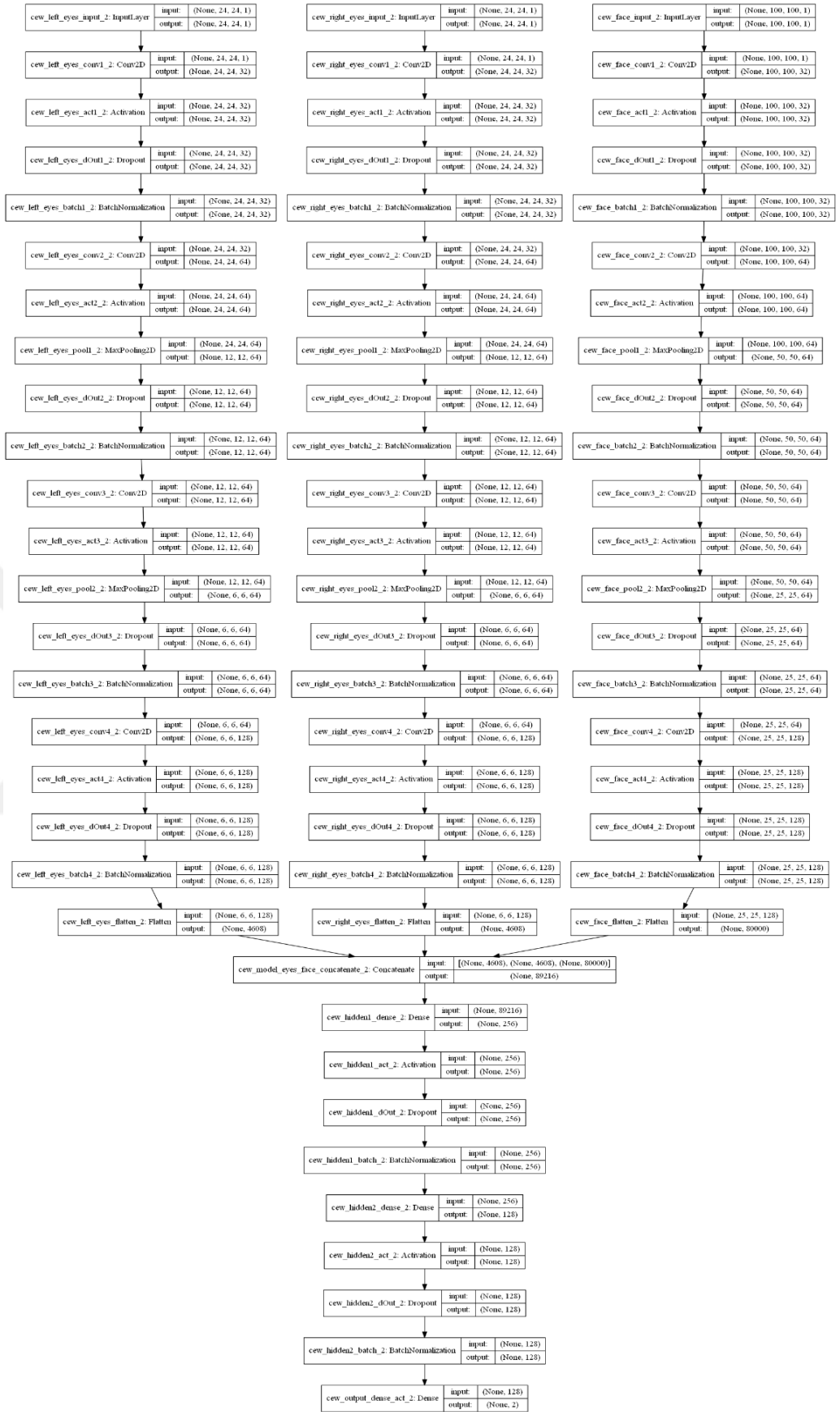
Şekil A.5. YawDD veri seti ağız girdi görüntüleri için geliştirilen sıralı model katmanları.



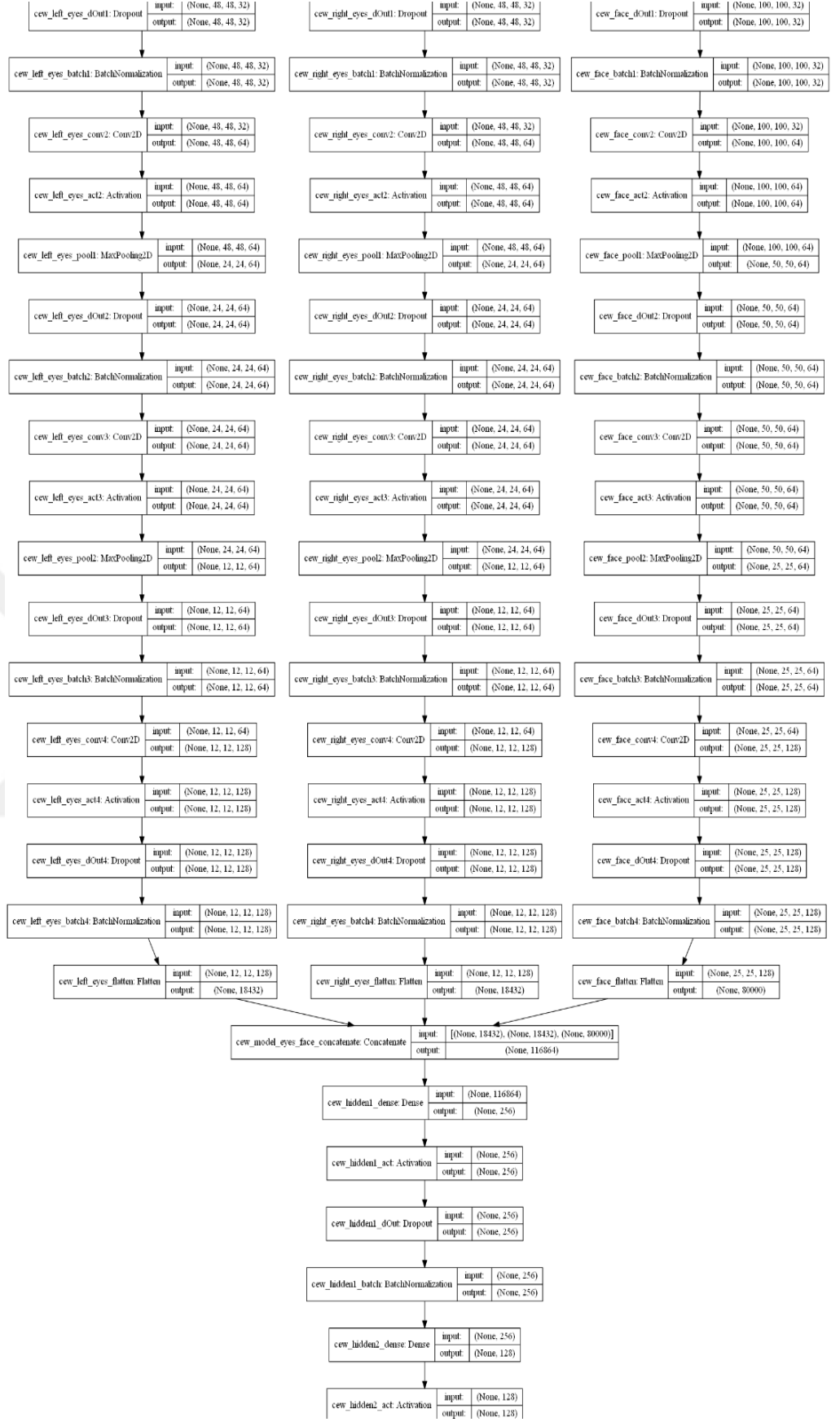
Şekil A.6. YawDD veri seti 24x24 piksel boyutundaki göz girdi görüntüleri için geliştirilen sıralı model katmanları.



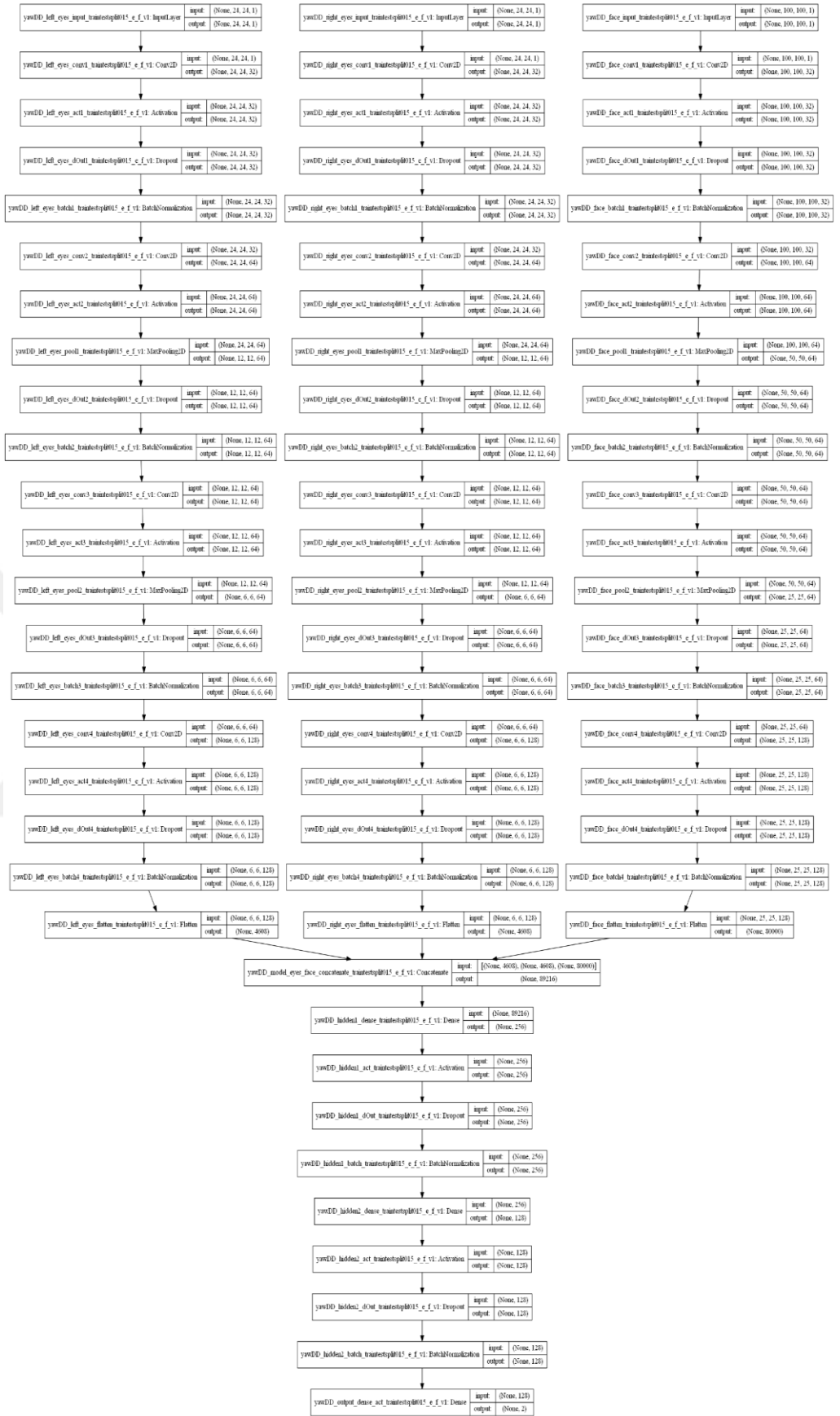
Şekil A.7. YawDD veri seti 48x48 piksel boyutundaki göz girdi görüntüleri için geliştirilen sıralı model katmanları.



Şekil A.8. Cew veri seti yüz ve 24x24 piksel boyutundaki göz girdi görüntüleri için geliştirilen fonksiyonel model katmanları.



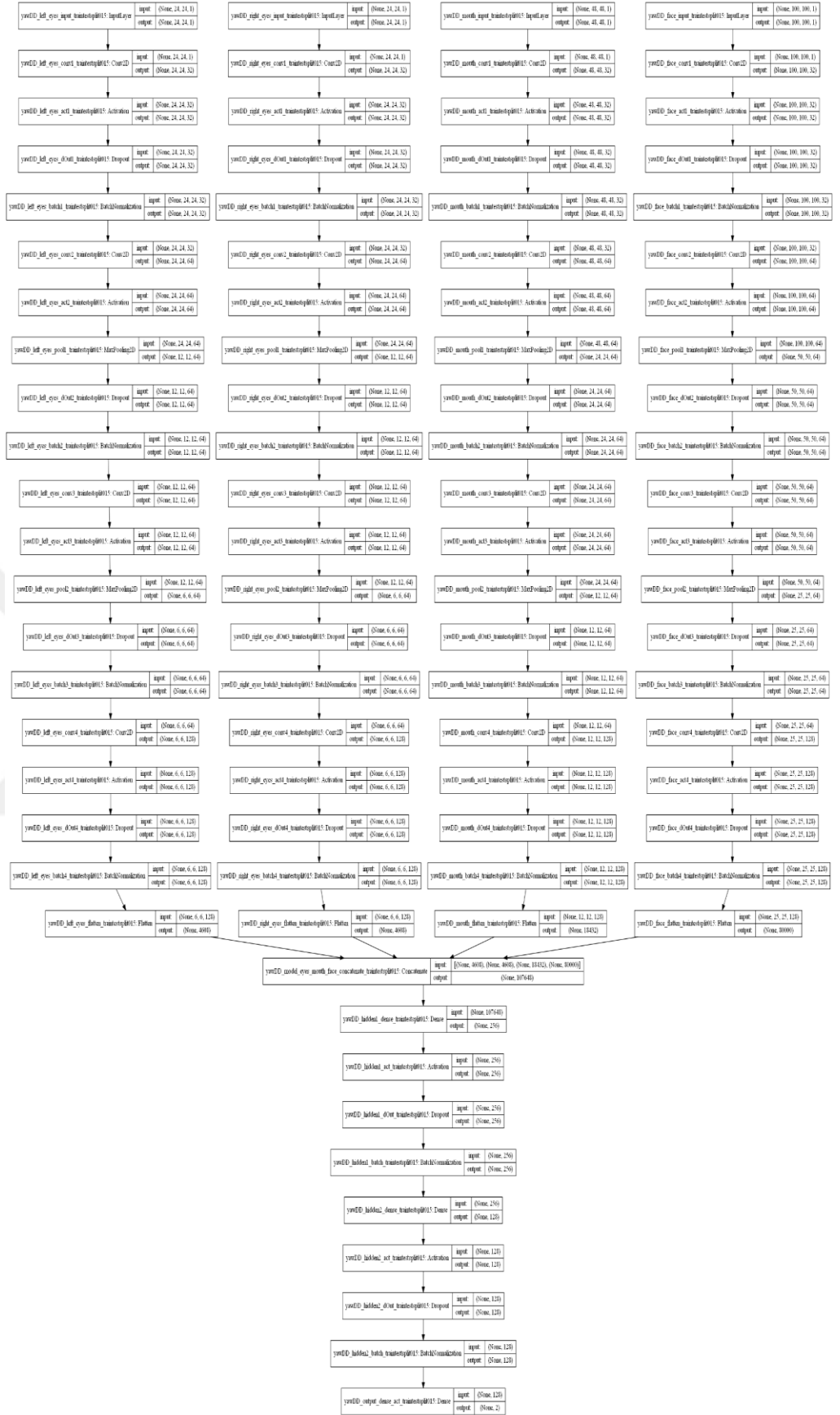
Şekil A.9. Cew veri seti yüz ve 48x48 piksel boyutundaki göz girdi görüntüleri için geliştirilen fonksiyonel model katmanları.



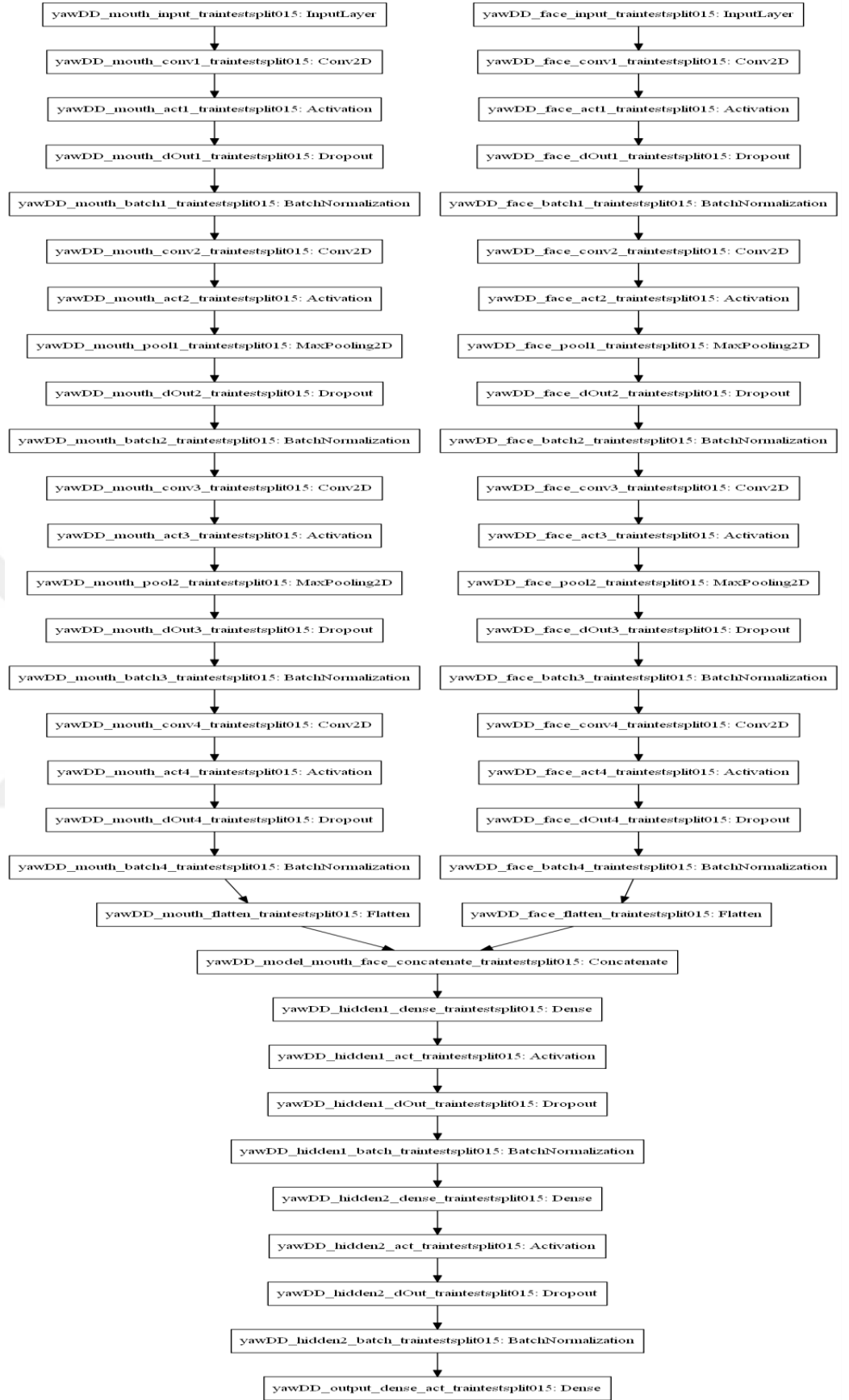
Şekil A.10. YawDD veri seti yüz ve 24x24 piksel boyutundaki göz girdi görüntüleri için geliştirilen fonksiyonel model katmanları.



Şekil A.11. YawDD veri seti yüz ve 48x48 piksel boyutundaki göz girdi görüntüleri için geliştirilen fonksiyonel model katmanları.



Şekil A.12. YawDD veri seti yüz, 24x24 piksel boyutundaki göz ,48x48 piksel boyutundaki ağır girdi görüntüleri için geliştirilen fonksiyonel model katmanları.



Şekil A.13. YawDD veri seti yüz, 48x48 piksel boyutundaki ağız girdi görüntüleri için geliştirilen fonksiyonel model katmanları.