



**DİYABETİK RETİNOPATİ EVRELERİNİN
DERİN ÖĞRENME TABANLI TESPİTİ**

YÜKSEK LİSANS TEZİ

Havva ÖZYILDIZ

Danışman

Dr. Öğr. Üyesi Fatih BAYRAM

MEKATRONİK MÜHENDİSLİĞİ ANABİLİM DALI

Mayıs 2025

AFYON KOCATEPE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

**DİYABETİK RETİNOPATİ EVRELERİNİN DERİN ÖĞRENME
TABANLI TESPİTİ**

Havva ÖZYILDIZ

Danışman

Dr. Öğr. Üyesi Fatih BAYRAM

MEKATRONİK MÜHENDİSLİĞİ ANABİLİM DALI

Mayıs 2025

TEZ ONAY SAYFASI

Havva ÖZYILDIZ tarafından hazırlanan “Diyabetik Retinopati Evrelerinin Derin Öğrenme Tabanlı Tespiti” adlı tez çalışması lisansüstü eğitim ve öğretim yönetmeliğinin ilgili maddeleri uyarınca 16/05/2025 tarihinde aşağıdaki jüri tarafından **oy birliği** ile Afyon Kocatepe Üniversitesi Fen Bilimleri Enstitüsü **Mekatronik Mühendisliği Anabilim Dalı’nda YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Danışman : Dr. Öğr. Üyesi Fatih BAYRAM

Başkan : Doç. Dr. Ahmet YÖNETKEN
Afyon Kocatepe Üniversitesi, Teknoloji Fakültesi

Üye : Dr. Öğr. Üyesi Fatih BAYRAM
Afyon Kocatepe Üniversitesi, Teknoloji Fakültesi

Üye : Doç. Dr. Ahmet BİÇER
Burdur Mehmet Akif Ersoy Üniversitesi, Fen Edebiyat
Fakültesi

Afyon Kocatepe Üniversitesi
Fen Bilimleri Enstitüsü Yönetim Kurulu’nun
..... /..... /..... tarih ve
..... sayılı kararıyla onaylanmıştır.

.....
Prof. Dr. Bekir YALÇIN
Enstitü Müdürü

BİLİMSEL ETİK BİLDİRİM SAYFASI

Afyon Kocatepe Üniversitesi

Fen Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- Görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- Başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- Atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- Kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- Ve bu tezin herhangi bir bölümünü bu üniversite veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

30 / 05 / 2025

Havva ÖZYILDIZ

ÖZET

Yüksek Lisans Tezi

DİYABETİK RETİNOPATİ EVRELERİNİN DERİN ÖĞRENME TABANLI TESPİTİ

Havva ÖZYILDIZ

Afyon Kocatepe Üniversitesi

Fen Bilimleri Enstitüsü

Mekatronik Mühendisliği Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Fatih BAYRAM

Bu yüksek lisans tezinde, diyabetik retinopatinin (DR) beş evresinin otomatik olarak sınıflandırılmasına yönelik derin öğrenme tabanlı bir model geliştirilmiştir. Çalışmada, APTOS 2019 veri setinde bulunan retina görüntüleri kullanılarak MobileNetV2, DenseNet169, NASNet, VGG19, ConvNeXt Large, EfficientNetB5, EfficientNetB6 ve EfficientNetV2-S olmak üzere sekiz farklı derin öğrenme mimarisi, transfer öğrenme yöntemiyle eğitilmiştir. Görüntü işleme ve veri artırma teknikleri ile sınıf dengesizliği problemleri azaltılmış, model performansı artırılmıştır. Elde edilen sonuçlara göre, EfficientNetV2-S modeli en yüksek başarıyı göstermiş ve ince ayar (fine-tuning) yöntemiyle performansı daha da iyileştirilmiştir. Test aşamasında, geliştirilen model test zamanı artırma (test-time augmentation) yöntemi ile %92,24 doğruluk, %96,77 özgüllük, %92,28 duyarlılık, %92,07 F1 skoru ve %98,99 AUC değerlerine ulaşmıştır. Elde edilen bulgular, geliştirilen sistemin diyabetik retinopatinin erken teşhisinde etkili bir yardımcı tanı aracı olabileceğini göstermektedir.

Anahtar Kelimeler: Derin öğrenme, Görüntü işleme, Transfer öğrenme, Diyabetik retinopati.

ABSTRACT

M.Sc. Thesis

DEEP LEARNING-BASED DETECTION OF DIABETIC RETINOPATHY STAGES

Havva ÖZYILDIZ

Afyon Kocatepe University

Graduate School of Natural and Applied Sciences

Department of Mechatronic Engineering

Supervisor: Asst. Prof. Fatih BAYRAM

In this master's thesis, a deep learning-based model was developed for the automatic classification of the five stages of diabetic retinopathy (DR). The study utilized retinal images from the APTOS 2019 dataset and employed eight different deep learning architectures—MobileNetV2, DenseNet169, NASNet, VGG19, ConvNeXt Large, EfficientNetB5, EfficientNetB6, and EfficientNetV2-S—trained using the transfer learning approach. Image processing and data augmentation techniques were applied to mitigate class imbalance issues and enhance model performance. Among the evaluated models, EfficientNetV2-S achieved the highest performance, which was further improved through fine-tuning. During the testing phase, the proposed model achieved 92.24% accuracy, 96.77% specificity, 92.28% sensitivity, 92.07% F1-score and 98.99% AUC by employing test-time augmentation. The findings indicate that the developed system can serve as an effective computer-aided diagnostic tool for the early detection of diabetic retinopathy.

Keywords: Deep learning, Image processing, Transfer learning, Diabetic retinopathy.

TEŞEKKÜR

Bu araştırmanın konusu, deneysel çalışmaların yönlendirilmesi, sonuçların değerlendirilmesi ve yazımı aşamasında yapmış olduğu büyük katkılarından dolayı tez danışmanım Sayın Dr. Öğr. Üyesi Fatih BAYRAM'a, her konuda öneri ve eleştirileriyle yardımlarını gördüğüm hocalarıma ve arkadaşlarıma teşekkür ederim.

Bu araştırma boyunca maddi ve manevi desteklerinden dolayı aileme teşekkür ederim.

Havva ÖZYILDIZ
Afyonkarahisar 2025

İÇİNDEKİLER DİZİNİ

	Sayfa
ÖZET	1
ABSTRACT	2
TEŞEKKÜR	3
İÇİNDEKİLER DİZİNİ.....	4
KISALTMALAR DİZİNİ	7
ŞEKİLLER DİZİNİ	8
ÇİZELGELER DİZİNİ.....	9
1. GİRİŞ.....	10
2. LİTERATÜR BİLGİLERİ	13
2.1 Makine Öğrenmesi Algoritmaları ile Yapılan Çalışmalar	13
2.2 Derin Öğrenme Mimarileri ile Yapılan Çalışmalar	15
2.3 Makine Öğrenmesi Algoritmaları ile Yapılan Diyabetik Retinopati Tespiti Çalışmaları	16
2.4 Derin Öğrenme Mimarileri ile Yapılan Diyabetik Retinopati Tespiti Çalışmaları	17
3. MATERYAL ve METOT	19
3.1 Veri Kümesi ve Ön İşleme.....	19
3.1.1 Diyabetik Retinopati Veri Seti	19
3.1.2 Ön İşleme	20
3.1.2.1 Gri Seviye Görüntü	20
3.1.2.2 Gauss Bulanıklığı	21
3.1.2.3 CLAHE (Kontrast Sınırlı Uyarlanabilir Histogram Eşitleme)	21
3.1.2.4 Medyan Bulanıklık Filtresi.....	22
3.1.2.5 Veri Artırma (Data Augmentation)	23
3.2 Python	24
3.2.1 TensorFlow Kütüphanesi	25
3.3 Yapay Zekâ (Artificial Intelligence)	25
3.4. Makine Öğrenmesi.....	26
3.5 Derin Öğrenme	26
3.5.1 Evrişimli Sinir Ağları (CNN).....	27
3.5.1.1 Giriş Katmanı	27

3.5.1.2 Evrişim Katmanı.....	28
3.5.1.3 Havuzlama Katmanı.....	28
3.5.1.4 Tam Bağlantı Katmanı	29
3.5.2 Transfer Öğrenme	30
3.5.3 Derin Öğrenme Modelleri	30
3.5.3.1 MobileNetV2.....	31
3.5.3.2 DenseNet169	32
3.5.3.3 NASNet	33
3.5.3.4 VGG19	33
3.5.3.5 ConvNeXt Large	34
3.5.3.6 EfficientNetB5.....	35
3.5.3.7 EfficientNetB6.....	36
3.5.3.8 EfficientNetV2-S.....	36
3.5.4 Ek Katmanlar	38
3.5.4.1 Dense Katmanı (2048 nöron)	38
3.5.4.2 Dropout Katmanı.....	39
3.5.4.3 Dense Katmanı (5 nöron)	39
3.5.4.4 Softmax Aktivasyon Fonksiyonu	39
3.5.5 Eğitim.....	40
3.5.5.1 Batch Size.....	41
3.5.5.2 Sınıf Ağırlıkları	41
3.5.5.3 Ağırlıklı Kategorik Çapraz Entropi Kayıp Fonksiyonu	42
3.5.5.4 Adam Optimizasyon Algoritması.....	42
3.5.5.5 Geri Çağırma Fonksiyonları (Callbacks) ve Aşırı Uyumun Önlenmesi	42
3.5.6 İnce Ayar (Fine Tuning).....	43
3.5.7 Test Zamanı Artırma (TTA).....	44
4. BULGULAR	46
4.1 Eğitim Sonuçları	46
4.2 İnce Ayar Sonuçları	48
4.3 Test Sonuçları	49
4.3.1 Karışıklık Matrisi (Confusion Matrix).....	50

4.3.2 Hassasiyet (Precision)	52
4.3.3 Duyarlılık (Recall)	52
4.3.4 Özgüllük (Specificity)	53
4.3.5 Doğruluk (Accuracy)	53
4.3.6 F1-Skoru (F1-Score)	53
4.3.7 ROC Eğrisi ve AUC Değeri	55
4.4 İşlenmiş ve İşlenmemiş Görüntüler Üzerinde Model Performans Analizi	58
4.5 Literatürdeki Yöntemlerle Modelin Karşılaştırmalı Performans Analizi	58
5. TARTIŞMA ve SONUÇ	61
6. KAYNAKLAR	63
ÖZGEÇMİŞ	69
EKLER	70

KISALTMALAR DİZİNİ

Kısaltmalar

AHE	Uyarlanabilir Histogram Eşitleme
AUC	Eğri Altındaki Alan
CLAHE	Kontrast Sınırlı Uyarlanabilir Histogram Eşitleme
CNN	Evrişimli Sinir Ağları
CPU	Merkezi İşlem Birimi
DT	Karar Ağaçları
ET	Ekstra Ağaçlar
FN	Yanlış Negatif
FP	Yanlış Pozitif
FPR	Yanlış Pozitif Oranı
GAN	Üretici Çekişmeli Ağ
GPU	Grafik İşlem Birimi
K-NN	K-En Yakın Komşu Algoritması
LSTM	Uzun Kısa Süreli Bellek
MLP	Çok Katmanlı Algılayıcı
MSA-Net	Çok Ölçekli Dikkat Ağı
R-CNN	Bölge Tabanlı Evrişimli Sinir Ağı
RF	Rastgele Orman Algoritması
RFE	Özyinelemeli Özellik Eleme
RGB	Kırmızı-Yeşil-Mavi (Renk Modeli)
RNN	Tekrarlayan Sinir Ağları
ROC	Alıcı İşletim Karakteristiği
SGD	Stokastik Gradyan İnişi
SVM	Destek Vektör Makineleri
TN	Doğru Negatif
TP	Doğru Pozitif
TTA	Test Zamanı Artırma
TPU	Tensör İşlem Birimi
VGG	Görüntü Tanıma Grubu

ŞEKİLLER DİZİNİ

	Sayfa
Şekil 1.1 Diyabetik retinopatili fundus görüntüsü (Dao vd. 2023)	11
Şekil 1.2 (a), (b), (c), (d), (e) – diyabetik retinopati aşamaları.....	11
Şekil 3.1 Ön İşleme Adımları.....	24
Şekil 3.2 Evrişimli Sinir Ağı Yapısı (İnt.Kyn.2).....	27
Şekil 3.3 Evrişim Katmanı Örneği (İnt.Kyn.3)	28
Şekil 3.4 Ortalama Havuzlama Katmanı Örneği (İnt.Kyn.3).....	29
Şekil 3.5 Tam Bağlantı Katmanı Örneği (İnt.Kyn.4)	30
Şekil 3.6 MobileNetV2 Mimarisi	31
Şekil 3.7 DenseNet169 Mimarisi	32
Şekil 3.8 NASNet Mimarisi	33
Şekil 3.9 VGG19 Mimarisi	34
Şekil 3.10 ConvNeXt Large Mimarisi	35
Şekil 3.11 EfficientNetB5 Mimarisi.....	36
Şekil 3.12 EfficientNetB5 Mimarisi.....	37
Şekil 3.13 EfficientNetV2-S Mimarisi.....	37
Şekil 3.14 Akış Şeması.....	45
Şekil 4.1 EfficientNetV2-S modelinin eğitim grafikleri	48
Şekil 4.2 EfficientNetV2-S modelinin ince ayar eğitim grafikleri.....	49
Şekil 4.3 ROC Eğrisi	57

ÇİZELGELER DİZİNİ

	Sayfa
Çizelge 3.1 APTOS veri setinde bulunan diyabetik retinopati sınıflarının dağılımı.....	19
Çizelge 4.1 Derin öğrenme modellerinin eğitim sonuçları	46
Çizelge 4.2 Karışıklık Matrisi	51
Çizelge 4.3 Test edilen CNN modelinin karışıklık matrisi	51
Çizelge 4.4 Test edilen CNN modelinin sınıflandırma performansı sonuçları	54
Çizelge 4.5 İşlenmiş ve işlenmemiş görüntülerin test sonuçları	58
Çizelge 4.6 APTOS veri setinde kullanılan farklı yöntemlerin karşılaştırmalı sonuçları	59



1. GİRİŞ

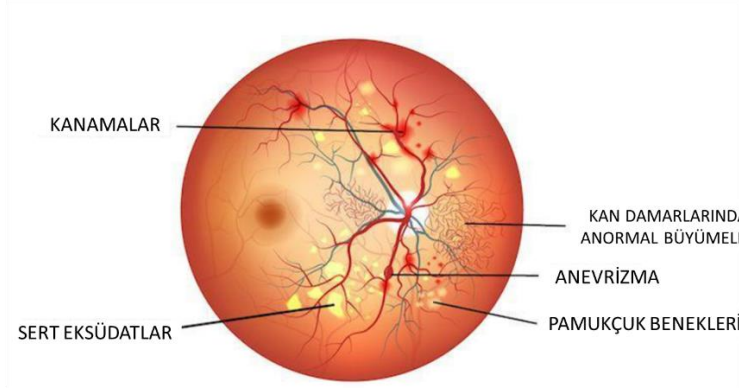
Diabetes mellitus, yaygın olarak bilinen adıyla diyabet, vücudun kan şekeri (glikoz) seviyelerini düzenleme yeteneğini olumsuz etkileyen bir metabolik bozukluktur (Association 2004). Genellikle yetersiz insülin üretimi, insülinin doğru şekilde kullanılmaması veya her iki durumun bir kombinasyonu nedeniyle ortaya çıkar. Diyabet kronik bir rahatsızlıktır.

Diyabet hastalarında yüksek kan şekeri seviyeleri zamanla çeşitli sağlık komplikasyonlarına yol açabilir. Retina ile ilişkili olduğunda, bu durum Diyabetik Retinopati olarak adlandırılır (Fong vd. 2004). Diyabetik retinopati, retinadaki kan damarlarını etkiler ve kontrol altına alınmadığında görme kaybına neden olabilir.

Diyabetik retinopati, diyabetin retinadaki kan damarları üzerindeki olumsuz etkilerinden kaynaklanır. Şekil 1.1’de diyabetik retinopati hastalığının bir retina üzerindeki etkilerini göstermektedir. Diyabetik retinopatide, uzun süreli yüksek kan şekeri seviyeleri nedeniyle retina damarlarında çeşitli yapısal bozukluklar meydana gelir. Görselde işaretlenen başlıca bulgular şunlardır:

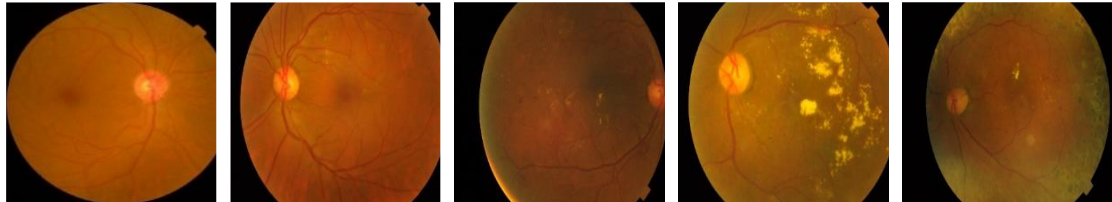
- Kanamalar:** Retinadaki küçük damarların hasar görmesi sonucu kan sızmasıdır. Görme kaybına yol açabilir.
- Kan damarlarında anormal büyümeler:** Proliferatif diyabetik retinopatide görülen, yeni ama kırılgan ve işlevsiz damar oluşumlarıdır.
- Anevrizma:** Kan damarlarında oluşan küçük balonlaşmalardır; duvarları zayıfladığı için sızıntı yapabilirler.
- Pamukçuk benekleri:** Sinir lif tabakasında lokal iskemi (kan akışı eksikliği) gösteren beyaz renkli lezyonlardır.
- Sert eksüdatlar:** Sızan lipid ve proteinlerin birikimi sonucu oluşan parlak sarımsı birikintilerdir.

Bu bulgular, retinanın oksijenlenmesinin ve beslenmesinin bozulduğunu gösterir ve zamanında müdahale edilmezse kalıcı görme kaybına neden olabilir. Bu etkiler arasında kan damarlarında şişme ve sızıntı, damarların daralması ve tıkanması bulunur. Durum ilerledikçe, yeni ve anormal kan damarlarının olduğu bir aşamaya ulaşır ve bu da sonunda ciddi görme kaybına neden olur (Tang ve Kern 2011).



Şekil 1.1 Diyabetik retinopatili fundus görüntüsü (Dao vd. 2023)

Bu hastalığın ilerleyişi dört ana evreye ayrılır. Şekil 1.2(a)'da normal olarak sınıflandırılan sağlıklı bir retina gösterilmektedir. Hastalığın birinci evresi olan hafif non-proliferatif diyabetik retinopati evresinde, retinadaki küçük kan damarlarında mikroanevrizma adı verilen balon benzeri küçük şişlikler bulunur (Şekil 1.2(b)). Bu mikroanevrizmalar retinaya sıvı sızdırabilir ve pamuk benzeri beyaz lezyonlar da ortaya çıkabilir. Orta düzey non-proliferatif diyabetik retinopati evresinde, bu anevrizmalardan sızan sıvının miktarında artış görülür (Şekil 1.2(c)). Şekil 1.2(d)'de, ileri düzey non-proliferatif diyabetik retinopatiye sahip bir hastanın retinasını göstermektedir. Bu evrede daha fazla kan damarı tıkanır ve bazı bölgelerde kan dolaşımı olmaz. Bu bölgeler, yeni kan damarlarının büyümesini teşvik eden büyüme faktörleri salgılar. Hastalığın en ileri evresi olan proliferatif diyabetik retinopati, Şekil 1.2(e)'de gösterilmektedir. Bu evrede, göz içinde anormal kan damarları gelişir. Bu yeni damarlar kırılgan yapıları nedeniyle yırtılabilir. Bu durum vitreus sıvısında kan birikimine yol açar. Bu tür kanamalar, görme alanında "kara lekeler" veya yüzen nesneler olarak adlandırılır. Vitreus kanaması ani görme kaybına neden olabilir (Nawaz vd. 2019).



Şekil 1.2 (a), (b), (c), (d), (e) – diyabetik retinopati aşamaları.

Diyabetik retinopatinin etkilerini kontrol altına almanın en önemli yollarından biri erken

teşhistir. Hastalık erken evrelerde tespit edildiğinde, uygun tedavi yöntemleriyle yönetilerek görme kaybı önlenir. Diyabetik retinopati teşhisi için sıkça kullanılan yöntem, retina görüntülerinin manuel değerlendirilmesidir. Ancak, bu yöntemin zaman alıcı olması ve öznel yorumlara dayanması nedeniyle, bilgisayar destekli tanı sistemlerinin kullanımı son yıllarda artış göstermiştir. Bu sistemler, retina görüntüleri üzerinde otomatik analizler gerçekleştirerek tanı sürecini hızlandırır.

Bu tez çalışmasında, veri yetersizliği ve model genelleme kapasitesi gibi önceki araştırmalarda tespit edilen temel sorunlar ele alınarak daha verimli ve tutarlı bir derin öğrenme tabanlı model geliştirilmiştir. Önerilen model, hafif, orta, şiddetli, proliferatif ve normal retinalar dahil olmak üzere retina görüntülerinden hastalığın beş farklı aşamasını tanımlamak için derin öğrenme tekniklerini kullanmaktadır. Retinal görüntüleri ilk olarak yeniden boyutlandırma, arka plan ayırma, kontrast geliştirme ve filtreleme gibi ön işleme adımlarına tabi tutulmuştur. Daha sonra eğitim görüntülerine veri artırma yöntemleri uygulanmıştır. İşlenmiş görüntüler kullanılarak, EfficientNetB5-B6, EfficientNetV2-S, VGG19, DenseNet169, ConvNeXt Large, NasNet Large ve MobileNetV2 gibi modeller, transfer öğrenme metodu kullanılarak eğitilmiştir. Eğitim sonuçlarına göre en yüksek performansı elde eden EfficientNetV2-S modeline ince ayar yöntemi uygulanmıştır. Son aşamada ise EfficientNetV2-S modelinin test verileri üzerindeki performansı analiz edilmiştir. Test verileri ayrıca Test Zamanı Artırma (TTA) yöntemi kullanılarak değerlendirilmiş ve sonuçlar karşılaştırılmıştır.

2. LİTERATÜR BİLGİLERİ

Literatürde, görüntü işleme, makine öğrenimi ve derin öğrenme yöntemleri kullanılarak farklı hastalık teşhisine yönelik bir dizi çalışma gerçekleştirilmiştir.

2.1 Makine Öğrenmesi Algoritmaları ile Yapılan Çalışmalar

Makine öğrenmesi, bilgisayarların açıkça programlanmaksızın veri üzerinden öğrenmesini ve bu veriler doğrultusunda tahminlerde bulunmasını sağlayan bir yapay zekâ alt alanıdır. Bu alanda kullanılan algoritmalar genellikle denetimli, denetimsiz ve pekiştirmeli öğrenme olarak üç ana gruba ayrılır. Denetimli öğrenme algoritmaları, etiketlenmiş verilerle eğitilerek sınıflandırma veya regresyon problemlerinde kullanılır. Bu algoritmalar, sağlık alanında hastalık teşhisi, risk analizi ve klinik karar destek sistemleri gibi birçok uygulamada kullanılmaktadır.

So vd. (2017) demans destek merkezlerinde kullanılan yöntemlerden esinlenerek, makine öğrenmesi teknikleri kullanarak demansın erken teşhisi için iki aşamalı bir yöntem geliştirmişlerdir. Naive Bayes, Bayes Ağı, Bagging, Lojistik Regresyon, Rastgele Orman, Destek Vektör Makineleri (SVM) ve Çok Katmanlı Algılayıcı (MLP) gibi makine öğrenmesi algoritmaları kullanmışlardır. Güney Kore'deki Gangbuk-Gu demans merkezinde 2008-2013 yılları arasında tarama yapılan hastalardan toplanan veriler üzerinde testler gerçekleştirilmiştir. İlk aşamada hastalar normal ve anormal olarak ikiye ayrılmıştır. İkinci aşamada ise anormal gruplar hafif bilişsel bozukluk veya demans olmak üzere yeniden iki grupta sınıflandırılmıştır. İlk aşamada en yüksek sınıflandırma başarısını MLP göstererek 0.97 F1-Skoruna ulaşmıştır. İkinci aşamada ise SVM algoritması en yüksek sonucu vererek 0.739 F1-Skoruna sahiptir.

Asri vd. (2016) meme kanseri teşhisi için makine öğrenmesi algoritmaları kullanmışlardır. Wisconsin Meme Kanseri veri seti üzerinde Destek Vektör Makineleri (SVM), Naive Bayes, Karar Ağacı ve K-En Yakın Komşu (K-NN) algoritmaları test edilmiştir. Bu veri seti 699 örnekten oluşmakta ve bu örneklerin 458'i iyi huylu (benign), 241'i kötü huylu (malignant) olarak etiketlenmiştir. Çalışmanın sonuçlarına göre SVM

algoritması en yüksek doğruluk oranını (0.9713) elde etmiştir.

Naji vd. (2021) Meme kanserinin tahmini ve teşhisi için beş farklı makine öğrenmesi algoritması kullanmıştır. Bunlar Destek Vektör Makineleri (SVM), Rastgele Orman, Lojistik Regresyon, Karar Ağacı ve k-En Yakın Komşu (KNN) algoritmalarıdır. Çalışmada, Wisconsin Meme Kanseri Teşhis Veri Seti kullanılmıştır. Bu veri seti, 569 örnekten oluşmakta ve bu örneklerin 357'si iyi huylu (benign), 212'si kötü huylu (malignant) olarak ayrılmaktadır. Çalışmanın sonuçlarına göre, Destek Vektör Makineleri (SVM) algoritması 0.972 ile en yüksek doğruluk oranını elde etmiştir.

Karapınar Senturk (2020) yaptığı çalışmada parkinson hastalığının erken teşhisi için makine öğrenmesi tabanlı bir yöntem önermiştir. Destek Vektör Makineleri (SVM), Yapay Sinir Ağları ve Sınıflandırma ve Regresyon Ağaçları gibi algoritmalar kullanarak Parkinson hastalığının sınıflandırılması gerçekleştirilmiştir. Çalışmada, UCI Makine Öğrenmesi veri setinden alınan parkinson hastalarına ve sağlıklı bireylere ait ses sinyalleri kullanılmıştır. Veri seti, 31 kişiden (23 Parkinson hastası ve 8 sağlıklı birey) alınan 195 ses ölçümünü içermektedir. Özellik seçimi için Özellik Önemi ve Özyinelemeli Özellik Eleme (RFE) yöntemleri kullanılmıştır. Çalışmanın sonuçlarına göre, SVM algoritması, RFE yöntemi ile birlikte kullanıldığında en yüksek doğruluk oranını (0.9384) elde etmiştir. Regresyon Ağaçları doğruluk oranı 0.9076 iken YSA doğruluk oranı 0.9154 olarak ölçülmüştür.

Derin ven trombozu (DVT), kan pıhtılarının damarlarda oluşması ve kan akışını engellemesi sonucu ortaya çıkan bir sağlık sorunudur. Contreras-Luján vd. (2022) bu hastalığın erken teşhisi için bazı makine öğrenmesi algoritmaları kullanmışlardır. Algoritmaların performansı Wells kriterleri, hasta yaşı ve cinsiyeti gibi veriler kullanılarak ölçülmüştür. Çalışmada Karar Ağaçları (DT), Ekstra Ağaçlar (ET), k-En Yakın Komşu (K-NN), Çok Katmanlı Algılayıcı Sinir Ağı, Rastgele Orman (RF) ve Destek Vektör Makineleri (SVM) algoritmaları kullanılmıştır. Ayrıca, modellerin Raspberry Pi 4 gibi düşük maliyetli ve taşınabilir bir cihazda çalıştırılabilirliği test edilmiştir. Çalışmanın sonuçlarına göre K-NN algoritması en yüksek doğruluk oranını (0.904) elde etmiştir.

2.2 Derin Öğrenme Mimarileri ile Yapılan Çalışmalar

Derin öğrenme, makine öğrenmesinin bir alt dalı olup çok katmanlı yapay sinir ağları aracılığıyla veriden özellik çıkarımı ve öğrenme gerçekleştiren bir yöntemdir. Özellikle büyük veri ve yüksek işlem gücü gerektiren problemler için kullanılmaktadır. Görüntü işleme alanında yaygın olarak kullanılan derin öğrenme mimarileri arasında Evrişimli Sinir Ağları (Convolutional Neural Networks - CNN) en çok kullanılan yöntemler arasındadır. CNN tabanlı mimariler, görüntülerdeki mekansal ve yapısal özellikleri otomatik olarak öğrenerek sınıflandırma, nesne tanıma ve segmentasyon gibi çalışmalarda yüksek doğruluk sağlamaktadır.

Wang (2022) çalışmasında derin öğrenme tekniklerini kullanarak akciğer kanseri ve pulmoner nodül tespiti yapmıştır. CNN mimarileri ve U-Net aracılığıyla nodül segmentasyonu, tespiti ve sınıflandırmasında, LIDC-IDRI (1018 BT görüntüsü), LUNA16 (888 BT), ve NLST (10.058 hasta) gibi kamuya açık veri setleri kullanmıştır. Önerdiği transfer öğrenme tabanlı model 0.997 gibi yüksek bir doğruluk oranına ulaşmıştır.

Houssein vd. (2022) çalışmalarında, meme kanseri teşhisi için derin öğrenme tabanlı bir model önermişlerdir. Çalışmada, deniz avcılığı algoritması tabanlı bir optimizasyon yöntemi geliştirilmiş ve bu yöntem, ResNet50 mimarisi ile birleştirilerek IMPA-ResNet50 adlı bir model oluşturulmuştur. Çalışmada MIAS (322 mamografi görüntüsü) ve CBIS-DDSM (5283 mamografi görüntüsü) veri setleri kullanılmıştır. IMPA-ResNet50 modeli, CBIS-DDSM veri setinde %98.32 doğruluk oranı elde etmiştir.

Jojoa Acosta vd. (2021) yaptıkları çalışmada melanom hastalığının erken teşhisi için lezyonların veya cilt pigmentasyonunun dermatoskopik görüntüsü aracılığıyla derin öğrenme tabanlı bir yöntem geliştirmişlerdir. Veri seti olarak 2017 Uluslararası Biyomedikal Görüntüleme Sempozyumu'nda sunulan yarışmaya ait veri tabanı kullanılmıştır. İki aşamadan oluşan bu çalışmanın ilk aşamasında Mask R-CNN tekniği kullanılarak lezyonların olduğu bölge kırpılmıştır. İkinci aşamada ise ResNet152 mimarisi kullanılarak lezyonlar “iyi huylu” veya “kötü huylu” olarak sınıflandırılmıştır. Modelin doğruluk oranı 0.904 olarak ölçülmüştür.

Liu Y vd. (2020) cilt hastalıklarının tespiti için Inception-v4 tabanlı bir model önermiştir. 16.114 anonimleştirilmiş vakanın (fotoğraflar ve klinik veriler) bulunduğu bir teledermatoloji veri setini kullanarak, 26 yaygın cilt hastalığını ve ek olarak 419 farklı cilt hastalığını tespit etmeyi amaçlamışlardır. 0.66 oranında doğruluk elde etmişlerdir.

Li Y vd. (2019) mide kanserinin erken teşhisi için ResNet34 tabanlı yeni bir yöntem geliştirmişlerdir. Flüoresans hiperspektral görüntüleme yöntemi kullanılarak mide mukozasından alınan 120 taze doku örneği kullanılmıştır. Bu görüntüler Çin’de bulunan 2 hastaneden alınmıştır. Görüntüleri 3 farklı kategoride sınıflayan bu model 0.965 doğruluk oranına ulaşmıştır.

Liu S vd. (2014) çalışmalarında Alzheimer hastalığı ve demansın erken teşhisi için derin öğrenme tabanlı DEMNET modelini önermiştir. MRI görüntüleri kullanılarak Alzheimer’ın dört farklı aşaması sınıflandırılmıştır. Veri seti olarak Kaggle Alzheimer’s MRI veri seti ve ADNI veri seti kullanılmıştır. Modelin Kaggle veri setinde doğruluk oranı 0.9523 iken ADNI veri setinde 0.8483’tür.

2.3 Makine Öğrenmesi Algoritmaları ile Yapılan Diyabetik Retinopati Tespiti Çalışmaları

Reddy vd. (2020), Rastgele Orman Algoritması, Karar Ağacı Algoritması, Adaboost Algoritması, K-En Yakın Komşu (K-NN) Algoritması ve Lojistik Regresyon Algoritması gibi makine öğrenimi algoritmalarını kullanarak diyabetik retinopati tespiti için karşılaştırmalı bir çalışma yapmıştır. Test sonuçlarına göre doğruluk oranları sırasıyla 0.68, 0.55, 0.67, 0.65 ve 0.77 olarak rapor edilmiştir.

Sharma A vd. (2021), DIARETDB0 ve DIARETDB1 veri setlerini kullanarak üç farklı makine öğrenimi algoritmasıyla diyabetik retinopati tespiti yapmıştır. Görüntüler, akıllı kenar algılama, gri tonlama ve görüntü genişletme gibi morfolojik işlemlerle işlenmiştir. Kullanılan algoritmalar Ağırlıklı K-NN, Kübik SVM ve Basit Ağaç olup, doğruluk oranları sırasıyla 0.858, 0.872 ve 0.886 olarak elde edilmiştir.

Sumathy vd. (2022), 900 hasta kaydından oluşan bir veri setinde Logistic Regression, K-NN, SVM, Bagged Tree ve Boosted Tree makine öğrenimi algoritmalarıyla diyabetik retinopati sınıflandırması yapmıştır. Aşırı öğrenmeyi önlemek ve en uygun özellikleri seçmek için 2 katlı çapraz doğrulama uygulanmış, en iyi doğruluk oranı (0.901) K-NN algoritmasıyla elde edilmiştir. Ancak kullanılan veri setinin sınırlı olması nedeniyle modelin genelleme performansı sorgulanabilir.

2.4 Derin Öğrenme Mimarileri ile Yapılan Diyabetik Retinopati Tespiti Çalışmaları

Menaouer vd. (2022), APTOS 2019 veri setindeki görüntü dengesizliğini gerçek zamanlı veri artırma teknikleriyle ele almış ve görüntüleri 128x128 piksel boyutuna yeniden boyutlandırmıştır. VGG16 ve VGG19 modellerinden oluşan hibrit bir ağ eğitmiş ve 0.906 doğruluk oranına ulaşmıştır.

Al-Antary ve Arafa (2021), Eyepacs ve APTOS veri setlerini kullanarak çok ölçekli bir dikkat ağı (MSA-Net) modelini test etmiştir. Ön işleme adımı bir Gaussian Bulanıklık filtresi uygulanmıştır. Model, APTOS veri setinde 0.846, EyePacs veri setinde ise 0.878 doğruluk oranına ulaşmıştır.

Mondal vd. (2023), CLAHE yöntemiyle görüntü işleme yapmış ve APTOS 2019 ile DIARETDB1 veri setlerini birleştirmiştir. Yeni görüntüler oluşturmak için GAN tabanlı veri artırma teknikleri uygulamış ve EDLDR adlı bir model geliştirmiştir. Model, beş sınıflı görüntü sınıflandırmasında 0.8608 doğruluk oranı elde etmiştir. Diyabetik retinopati sınıflandırması için, görüntüleri 819x614 çözünürlüğe yeniden boyutlandırmış, [-1, 1] aralığında normalleştirmiş ve bir medyan filtre kullanılarak filtrelemiştir.

Kassani vd. (2019), Xception modelini diyabetik retinopati tespiti için modifiye etmiş ve sınıflandırma performansını minimum pooling katmanı ekleyerek iyileştirmiştir. Modelleri test aşamasında 0.8309 doğruluk oranına ulaşmıştır.

Kallel ve Echtioui (2024), APTOS 2019 veri setini kullanarak dört farklı transfer öğrenme modeliyle diyabetik retinopati tespiti yapmıştır. Görüntüleri yeniden boyutlandırma ve

dairesel kırpma yöntemleriyle ön işlemiş, veri artırma için görüntüleri döndürme uygulamıştır. En başarılı model InceptionV3 olmuş, test aşamasında 0.866 F1 skoru ve 0.9688 doğruluk oranı elde etmiştir. Doğruluk oranı yüksek olmasına rağmen, F1 puanı nispeten daha geride kalmıştır.

Khan vd. (2024), halka açık yedi aşamalı diyabetik retinopati veri setini kullanarak dört farklı derin öğrenme modelini eğitmiş ve değerlendirmiştir. Veri artırma tekniklerini kullanarak gerçekleştirdikleri test sonuçlarına göre VGG16 (%96.95), VGG19 (%96.75), AlexNet (%96.09) ve SqueezeNet (%92.96) yüksek doğruluk elde etmiştir.



3. MATERYAL ve METOT

Bu bölümde çalışmada kullanılan veri kümesi, görüntü ön işleme adımları, yazılım araçları ve derin öğrenme yöntemleri açıklanmıştır. Diyabetik retinopati sınıflandırması için kullanılan modeller, eğitim süreci ve değerlendirme aşamaları detaylı olarak sunulmuştur. Yöntemlerin anlaşılır ve tekrarlanabilir olması amaçlanmıştır.

3.1 Veri Kümesi ve Ön İşleme

Derin öğrenme tabanlı modellerin başarısı, büyük ve dengeli veri kümeleri ve ön işleme adımlarına bağlıdır. Bu bölümde, çalışmada kullanılan veri kümesine dair genel bilgiler verilmekte ve model eğitimi öncesinde uygulanan ön işleme teknikleri açıklanmaktadır. Ön işleme adımları, veri kalitesini artırmak, gürültüyü azaltmak ve modelin öğrenme kapasitesini geliştirmek amacıyla gerçekleştirilmiştir.

3.1.1 Diyabetik Retinopati Veri Seti

Bu çalışmada 2019 yılında yayınlanan APTOS (Asya Pasifik Tele-Oftalmoloji Derneği) veri seti kullanılmıştır. Kaggle platformunda Diyabetik Retinopati Sınıflandırma Yarışması için oluşturulan APTOS veri seti beş farklı sınıftan oluşmaktadır. Bu sınıflar, normal retina görüntülerinden proliferatif aşamalara kadar hastalığın beş farklı aşamasını içermektedir. Her görüntü, optik sinir başı, makula ve vasküler yapılar gibi gözün retina bölgelerini içermektedir. Veri seti toplam 3662 retina görüntüsünden oluşmaktadır (İnt.Kyn.1). Hastalığın evreleri ve sınıf başına düşen görüntü sayısı Çizelge 3.1'de sunulmuştur.

Çizelge 3.1 APTOS veri setinde bulunan diyabetik retinopati sınıflarının dağılımı

Sınıf	Hastalık Evresi	Görüntü Sayısı
0	Normal	1805
1	Hafif	370
2	Orta	999
3	Şiddetli	193
4	Proliferatif	295

Bu dağılım, sınıflar arasında önemli bir dengesizliği göstermektedir. Örneğin, hastalığın yokluğunu temsil eden 0 sınıfına ait görüntüler, veri setinin neredeyse yarısını oluşturmaktadır. Buna karşılık, şiddetli ve proliferatif diyabetik retinopati sınıflarında daha az görüntü bulunmaktadır. Bu tür dengesizlikler model eğitimi önemli ölçüde etkileyebilmektedir. Bu nedenle bu tez çalışmasında veri artırma teknikleri kullanılmış ve sınıf ağırlıkları ayarlanmıştır.

3.1.2 Ön İşleme

Görüntülerdeki paraziti gidermek ve sınıflandırma performansını artırmak için çeşitli ön işleme adımları uygulanmıştır. Başlangıçta, görüntüler üç bölüme ayrılmıştır: eğitim için %70, doğrulama için %15 ve test için %15. Eğitim ve doğrulama görüntüleri ön işlemeye tabi tutulurken, test görüntülerine herhangi bir ön işleme uygulanmamıştır. Eğitim ve doğrulama görüntüleri, tüm görüntülerde tek tip boyutlar sağlamak için önce 400x400 piksel olarak yeniden boyutlandırılmıştır. Bu büyüklük, modelin hem hesaplama verimliliğini koruduğu hem de retina detaylarını yeterli düzeyde içerdiği göz önünde bulundurularak belirlenmiştir. Alternatif olarak, daha büyük veya daha küçük boyutlar test edilmiş, ancak 400x400 piksel hem performans hem de işlem süresi açısından en uygun dengeyi sağlamıştır.

3.1.2.1 Gri Seviye Görüntü

Görüntüler 400x400 piksel olarak yeniden boyutlandırılmasının ardından gri seviyeye dönüştürülmüştür. Görüntü işlemede yaygın olarak kullanılan gri seviye görüntüler, yalnızca parlaklık bilgisi içeren tek kanallı görüntülerdir. Renkli görüntüler kırmızı, yeşil ve mavi (RGB) bileşenlerden oluşurken, gri seviye görüntülerde her piksel, 0 ile 255 arasında değişen bir yoğunluk değeri ile oluşmaktadır. Gri seviye görüntüler, görüntü işlemenin erken dönemlerinden itibaren bilgisayarla görme, tıbbi görüntüleme ve uzaktan algılama gibi alanlarda kullanılmıştır.

Gri seviye dönüşümünün amacı, görüntüdeki gereksiz renk bilgisini kaldırarak işlemeyi basitleştirmek ve görüntü özelliklerinin daha iyi analiz edilmesini sağlamaktır. Tıbbi

görüntüleme alanında, retina görüntüleri gibi biyomedikal verilerin işlenmesinde gri seviye görüntüler ile renk bilgisinin azaltılması, görüntüdeki gürültünün etkisini minimize etmektedir.

3.1.2.2 Gauss Bulanıklığı

Gauss bulanıklığı, görüntü işleme alanında yaygın olarak kullanılan bir filtreleme tekniğidir. Bu yöntem, gauss fonksiyonuna dayalı bir ağırlıklı ortalama hesaplaması yaparak görüntüleri yumuşatır (Gedraite ve Hadad 2011). Bu tez çalışmasında gürültüyü azaltmak için Gauss bulanıklık filtresi uygulanmış ardından arka plandan ayırmak için, retina görüntülerinin ana çerçevesi tanımlanarak daha net görüntüler elde edilmiştir. Bu adımdan sonra, tüm görüntülerin odak bölgeleri başarıyla çıkarılmıştır.

3.1.2.3 CLAHE (Kontrast Sınırlı Uyarlanabilir Histogram Eşitleme)

Histogram eşitleme, uzun zamandır görüntü işleme alanında yaygın bir teknik olarak kullanılmıştır. Ancak, geleneksel histogram eşitleme yöntemleri, tüm görüntüye tek bir dönüşüm uyguladığı için yerel kontrast farklılıklarını yeterince iyileştiremez ve aşırı parlak veya aşırı koyu bölgeler oluşturur. Bunu önlemek amacıyla Uyarlanabilir Histogram Eşitleme (AHE - Adaptive Histogram Equalization) geliştirilmiştir. AHE, görüntüyü küçük bölgelere (genellikle 8×8 veya 16×16 piksellik bloklar) ayırarak her bölgenin kontrastını bağımsız olarak artırır. Ancak, AHE'nin en büyük dezavantajlarından biri, gürültüyü artırmasıdır. Bu sorunu gidermek için CLAHE yöntemi geliştirilmiştir. CLAHE, AHE'nin geliştirilmiş bir versiyonudur ve kontrastı sınırlandırarak aşırı parlak bölgeler oluşmasını engeller (Sharma R ve Kamra 2023).

Görüntü işleme tekniklerinden biri olan Kontrast Sınırlı Uyarlanabilir Histogram Eşitleme (CLAHE), düşük kontrastlı görüntülerin iyileştirilmesi için kullanılan gelişmiş bir histogram eşitleme yöntemidir. Özellikle tıbbi görüntüleme, uydu görüntüleri ve düşük ışık koşullarındaki görüntü iyileştirme süreçlerinde yaygın olarak kullanılmaktadır.

CLAHE yönteminde öncelikle görüntüler küçük bölümlere (örneğin 8×8 veya 16×16

piksel) ayrılır. Her küçük bölgedeki parlaklık değerleri için ayrı bir histogram oluşturulur. Histogramdaki maksimum yoğunluk belirli bir eşik değeri ile sınırlandırılır. Eğer belirli bir parlaklık seviyesinin frekansı bu eşik değerini aşarsa, fazla pikseller diğer parlaklık seviyelerine dağıtılır. Bu işlem aşırı kontrast artışını engeller. Ardından blokların birleşim bölgelerinde ani geçişleri önlemek için komşu bölgeler arasında doğrusal interpolasyon uygulanarak pürüzsüz geçişler sağlanır.

Bu tez çalışmasında, CLAHE yöntemi, retina görüntülerindeki düşük kontrastlı alanları iyileştirmek amacıyla kullanılmıştır. Bu işlem sayesinde, retinadaki önemli hastalık yapıları daha belirgin hale gelmiştir.

3.1.2.4 Medyan Bulanıklık Filtresi

Medyan filtresi, geleneksel doğrusal (linear) filtrelerden farklı olarak doğrusal olmayan (non-linear) bir filtreleme tekniğidir. Matematiksel olarak medyan filtresi, belirli bir çap içinde kalan piksellerin parlaklık değerlerini alır, bunları küçükten büyüğe sıralar ve ortadaki (medyan) değeri kullanarak merkezi pikseli günceller. Böylece aşırı parlak veya aşırı koyu piksellerin etkisi azalır ve görüntüdeki keskin geçişler yumuşatılır.

Medyan bulanıklık filtresi, görüntü işleme alanında yaygın olarak kullanılan bir gürültü azaltma tekniğidir. Özellikle tuz ve biber gürültüsü (salt and pepper noise) olarak bilinen rastgele siyah-beyaz noktaların bulunduğu görüntülerde oldukça etkilidir. Bu yöntem, her pikselin parlaklık değerini komşu piksellerin medyan değeriyle değiştirerek görüntüyü yumuşatır ve istenmeyen gürültüyü giderir (Li X vd. 2021).

Bu tez çalışmasında, son olarak görüntülere bir medyan bulanıklık filtresi uygulanarak kaydedilmiştir. Medyan bulanıklık filtresi retina görüntülerindeki gürültüyü azaltmak için kullanılmıştır. Özellikle düşük kaliteli ve parazit içeren görüntülerde, önemli detayları koruyarak netliği artırmak için tercih edilmiştir. Bu işlem sayesinde, retina görüntüleri daha temiz hale getirilmiştir.

3.1.2.5 Veri Artırma (Data Augmentation)

Derin öğrenme modellerinin başarısı, kullanılan veri setinin çeşitliliği ve kalitesi ile doğrudan ilişkilidir. Ancak, tıbbi görüntüleme gibi alanlarda yeterli miktarda etiketlenmiş veri elde etmek oldukça zordur. Özellikle retina görüntüleri gibi spesifik ve medikal uzmanlık gerektiren veri kümeleri, büyük miktarlarda toplanması zor olan veri türleridir. Veri artırma teknikleri kullanılarak bu zorluğun önüne geçmek mümkündür (Shorten ve Khoshgoftaar 2019).

APTOS 2019 veri setinde sınıflar arası dengesizlik oldukça fazla olup bazı sınıflarda yeteri kadar veri yoktur. Bu nedenle, veri artırma (data augmentation) teknikleri kullanılarak bu veri setinin çeşitliliği artırılmıştır. Aynı zamanda modelin aşırı öğrenmesini (overfitting) engelleme amacıyla da bu tekniğe başvurulmuştur. Veri artırma yöntemi olarak birden fazla teknik kullanılmıştır.

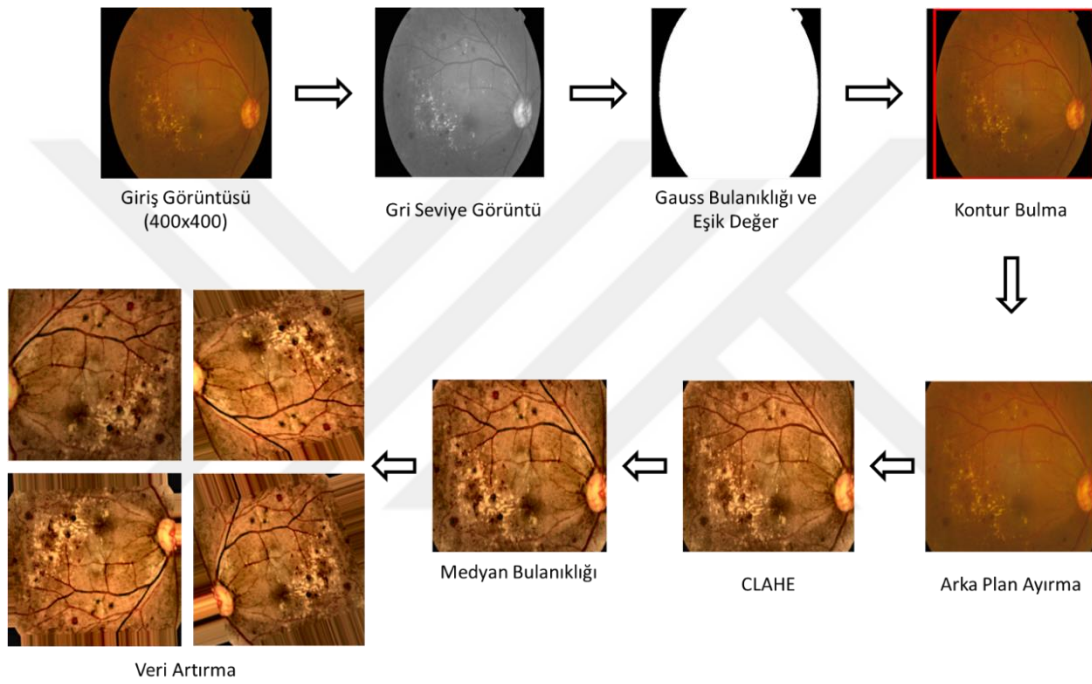
Dönme işlemi, görüntünün belirli bir açıyla çevrilerek yeni bir perspektif kazanmasını sağlayan bir veri artırma tekniğidir. Bu çalışmanın ön testlerinde 30 derece ve üzerindeki dönmelerin retinal vasküler yapıyı bozduğu ve modelin doğruluğunu olumsuz etkilediği gözlemlenmiştir. Retina görüntülerinde damar yapılarının belirli bir yönde düzenlendiği göz önüne alındığında, aşırı döndürme işlemleri modelin yanlış öğrenmesine neden olabilir. Bu nedenle, 0-20 derece aralığında dönme açısı seçilerek bir denge sağlanmıştır. Görüntüler rastgele 0 ila 20 derece arasında döndürülerek hem yatay hem de dikey eksenler boyunca rastgele çeviriler yapılmıştır.

Retina görüntülerinde simetri özellikleri dikkate alınarak hem yatay hem de dikey ekseninde rastgele aynalama işlemleri uygulanmıştır. Bunun yanı sıra rastgele çevirme işlemi uygulanarak görüntülerin bazı piksel değerleri sağa, sola, yukarı veya aşağı kaydırılmıştır. Bu işlem, modelin hafif konum değişikliğine karşı daha sağlam hale gelmesini sağlamıştır.

Tıbbi görüntülerde ışık koşulları değişkenlik gösterebileceğinden, modelin farklı aydınlatma koşullarına adapte olmasını sağlamak için parlaklık seviyeleri rastgele

değiştirilerek veri artırımı yapılmıştır. Aynı zamanda retina görüntülerindeki farklı ölçeklerdeki özellikleri modelin öğrenmesini sağlamak amacıyla rastgele yakınlaştırma ve uzaklaştırma işlemleri uygulanmıştır.

Tüm bu görüntü işleme ve veri artırma teknikleri, eğitim ve doğrulama veri setinin çeşitliliğini artırarak modelin genelleme yeteneğini güçlendirmiştir. Ön işleme adımlarını gösteren bir diyagram Şekil 3.1'de sunulmuştur.



Şekil 3.1 Ön İşleme Adımları

Görüntü ön işleme adımlarında kullanılan teknikler, modelin düşük kaliteli görüntülerden daha sağlam özellikler öğrenmesini sağlamıştır. Özellikle, kontrast geliştirme adımı, diyabetik retinopati hastalığının ayrıntılarını daha belirgin hale getirerek modelin sınıflandırma doğruluğunu artırmıştır. Görüntülerin yeniden boyutlandırılması ve işlenmesi, eğitim sürecindeki hesaplama yükünü de azaltarak daha verimli model eğitimi sağlamaktadır.

3.2 Python

Python, 1991 yılında geliştirilmiş olan açık kaynaklı, yüksek seviyeli ve nesne yönelimli

bir programlama dilidir. Geliştirilme amacı, okunaklı ve öğrenmesi kolay bir dil sunarak programcılarının daha hızlı ve verimli yazılım geliştirmesini sağlamaktır. Basit sözdizimi ve esnek yapısı sayesinde bilimsel hesaplamalar, yapay zeka, web geliştirme, veri analizi gibi pek çok alanda kullanılmaktadır. Günümüzde akademide ve endüstride yaygın olarak kullanılan Python, büyük ölçekli projelerde de fazlaca tercih edilmektedir (Raschka ve Mirjalili 2019).

Python, çok yönlü yapısı sayesinde farklı alanlarda geniş bir kullanıma sahiptir. Yapay zekâ ve makine öğrenmesi uygulamalarında, Python'ın TensorFlow ve PyTorch kütüphanelerinden faydalanılmaktadır. Büyük veri analizlerinde ise Pandas ve NumPy gibi kütüphaneler sayesinde karmaşık veri kümeleri üzerinde analizler gerçekleştirilmektedir. Bu tez çalışmasında TensorFlow kütüphanesi kullanılmıştır.

3.2.1 TensorFlow Kütüphanesi

TensorFlow, Google Brain ekibi tarafından geliştirilen, açık kaynaklı ve genişletilebilir bir makine öğrenmesi kütüphanesidir. İlk olarak 2015 yılında piyasaya sürülen TensorFlow, özellikle derin öğrenme modellerinin oluşturulması ve eğitilmesi için tasarlanmıştır. Esneklik, ölçeklenebilirlik ve yüksek performanslı hesaplama yetenekleri sayesinde birçok uygulamada yaygın olarak kullanılmaktadır.

TensorFlow, CPU, GPU ve TPU gibi donanım hızlandırıcılarıyla uyumlu çalışarak hesaplama sürecini hızlandırır. Bu özellik, özellikle derin öğrenme modellerinin eğitimi sırasında büyük veri kümeleri üzerinde işlem yaparken oldukça faydalıdır (Raschka ve Mirjalili 2019).

3.3 Yapay Zekâ (Artificial Intelligence)

Yapay zekâ, insan zekasını taklit eden, öğrenme, problem çözme, karar verme ve dil anlama gibi bilişsel yetenekleri gerçekleştirebilen bilgisayar sistemleri geliştirmeyi amaçlayan bir bilim dalıdır. Yapay zekâ, makine öğrenmesi, derin öğrenme, doğal dil işleme, bilgisayarla görme ve robotik gibi birçok alt disiplini içeren geniş bir alandır.

Günümüzde yapay zekâ; sağlık, finans, otomotiv, perakende, güvenlik ve eğitim gibi pek çok sektörde kullanılmaktadır. Otonom araçlar, yüz tanıma sistemleri, sesli asistanlar, öneri sistemleri ve tıbbi teşhis sistemleri gibi birçok uygulama yapay zekâ sayesinde geliştirilmektedir. Yapay zekâ, birçok farklı teknik kullanılarak oluşturulmaktadır.

3.4. Makine Öğrenmesi

Makine öğrenmesi, bilgisayar sistemlerinin açıkça programlanmadan, verilerden öğrenerek belirli görevleri yerine getirmesini sağlayan bir yapay zekâ alt dalıdır. Bu alan, sistemlerin büyük miktarda veriyi analiz edip öğrenmesine ve daha önce görmediği veriler üzerinde tahmin yapmasına dayanır. Makine öğrenmesi, bilgisayarların insan müdahalesi olmadan öğrenmesini sağlayan bir model geliştirme sürecidir. Derin öğrenmenin temelini oluşturur.

3.5 Derin Öğrenme

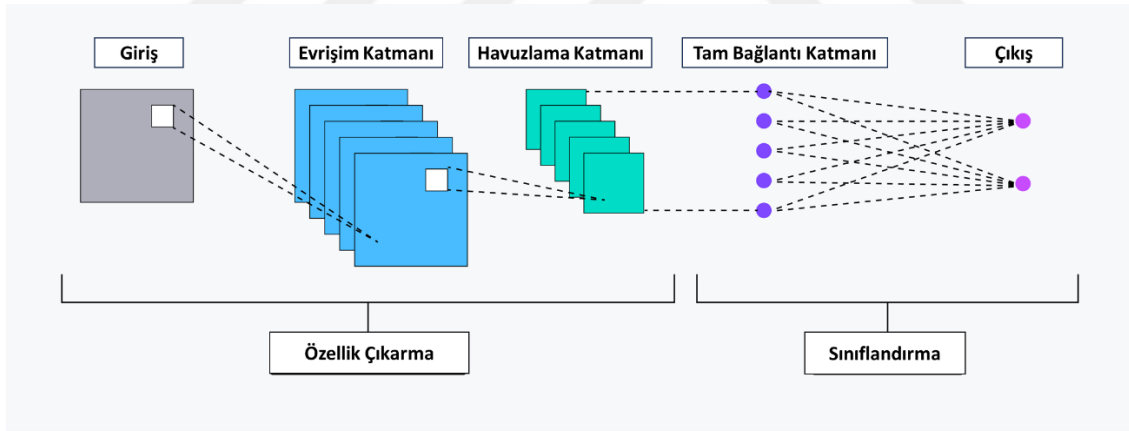
Makine öğreniminin bir alt kümesi olan derin öğrenme, büyük ve yüksek boyutlu veri kümelerini işleyerek kalıpları ve ilişkileri otomatik olarak öğrenebilen yapay sinir ağlarını kullanır. Bu yapay sinir ağları, insan beyninin çalışmasından ilham alan matematiksel modellerdir (Emmert-Streib vd. 2020). Derin öğrenme modelleri özellikle görüntü, metin, video ve ses gibi verileri işlemek için kullanılır ve bu tür verileri içeren birçok uygulamada genellikle geleneksel makine öğrenimi yöntemlerinden daha iyi performans gösterir (Jia vd. 2024, Manalı vd. 2024). Geleneksel makine öğrenmesi modellerinden farklı olarak, öznitelik çıkarma gibi manuel işlemler gerektirmez. Bunun yerine, model kendiliğinden uygun özellikleri öğrenmektedir.

Derin öğrenme modelleri makine öğrenmesinden farklı olarak çok katmanlı sinir ağlarını içerir. Ancak hesaplama maliyetleri daha yüksektir ve büyük miktarda veriye ihtiyaç duyar. Derin öğrenme birçok farklı mimariye sahiptir. Bu mimariler Çok Katmanlı Algılayıcı (MLP), Evrişimli Sinir Ağları (CNN), Tekrarlayan Sinir Ağları (RNN), Uzun Kısa Süreli Bellek (LSTM) ve Transformer Modelleridir. Bu tez çalışmasında Evrişimli

Sinir Ağları (CNN) kullanılmıştır.

3.5.1 Evrişimli Sinir Ağları (CNN)

Evrişimli Sinir Ağları, derin öğrenme yöntemlerinde en güçlü ve yaygın olarak kullanılan mimariler arasındadır. Evrişimli Sinir Ağları, biyolojik sinir sisteminin yapısından ilham alan çok katmanlı bir mimari kullanarak verilerdeki özellikleri öğrenir (Lecun vd. 1998). Bu ağlar, görüntü işleme, konuşma tanıma ve metin analizi gibi alanlarda geleneksel makine öğrenimi yöntemlerine kıyasla üstün performans göstermektedir (Bayram ve Eleyan 2022). CNN'lerin temel avantajlarından biri, filtreler ve evrişim işlemleri yoluyla görüntülerdeki düşük seviyeli özellikleri (kenarlar ve dokular gibi) öğrenmeleri ve ardından daha derin katmanlarda yüksek seviyeli özelliklere (nesne tanıma gibi) ulaşmalarıdır (Emmert-Streib vd. 2020). Özellikle, tıbbi görüntüleme alanında, CNN tabanlı modeller, insan gözü tarafından tespit edilemeyen ayrıntıları yakalar. Şekil 3.2'de basit bir evrişimli sinir ağı yapısı verilmiştir.



Şekil 3.2 Evrişimli Sinir Ağı Yapısı (İnt.Kyn.2)

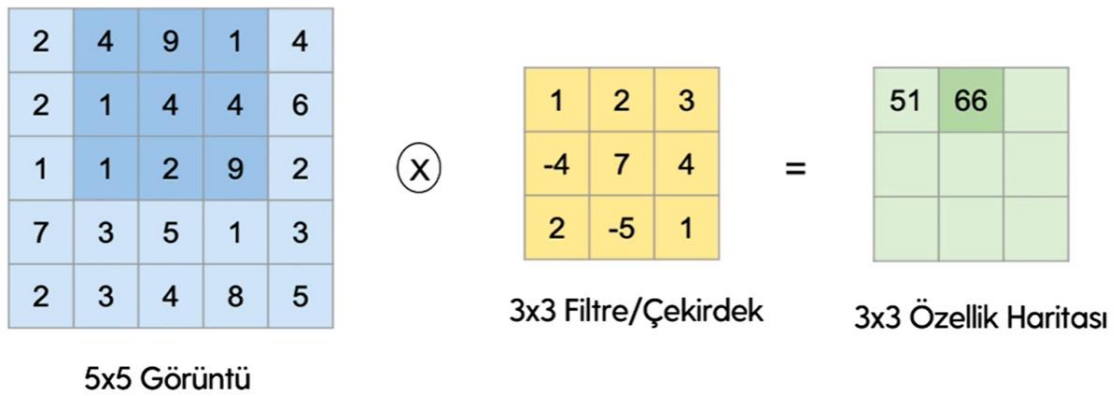
3.5.1.1 Giriş Katmanı

Evrişimli sinir ağlarında giriş katmanı, modelin aldığı veriyi işleyerek ağı derin katmanlarına uygun hale getiren ilk aşamadır. Görüntü tabanlı uygulamalarda giriş katmanı, genellikle belirli bir çözünürlüğe sahip RGB veya gri tonlamalı görüntüleri giriş olarak alır. Bu görüntüler, piksel değerlerinden oluşan çok boyutlu diziler şeklinde

modele verilir. Giriş katmanında ayrıca görüntülerin boyutları standart hale getirilir. Örneğin, farklı boyutlardaki görüntüler bir CNN modeline verilmeden önce sabit bir boyuta ölçeklendirilir. Sonuç olarak, verinin model için uygun bir formata getirilmesini sağlayarak CNN'nin öğrenme sürecinin temelini oluşturur.

3.5.1.2 Evrişim Katmanı

Evrişimli sinir ağlarının (CNN) en temel ve en önemli bileşenlerinden biri evrişim katmanıdır. Bu katman, giriş görüntüsünden anlamlı öznitelikler çıkarır. Evrişim işlemi, çekirdek (kernel) veya filtre adı verilen küçük boyutlu matrislerin giriş verisi üzerinde kaydırılmasıyla gerçekleştirilir. Çekirdek, genellikle 3×3 , 5×5 gibi küçük boyutlu bir matristir ve giriş görüntüsü üzerinde belirli bir adım (stride) ile kaydırılır. Her konumda, çekirdeğin elemanları ile giriş görüntüsünün ilgili bölgesi çarpılıp toplanarak yeni bir değer elde edilir. Bu işlem, tüm görüntü boyunca tekrarlandığında özellik haritası (feature map) oluşturulur. Çekirdek sayesinde görüntüdeki kenarlar, dokular, desenler ve nesne şekilleri tespit edilir. Bu sayede manuel öznitelik çıkarma işlemine gerek duyulmaz. Evrişim işlemi sonucunda, giriş görüntüsüne göre küçültülmüş ancak belirli özellikleri içeren yeni bir çıktı matrisi oluşur. Şekil 3.3'te evrişim katmanının bir örneği verilmiştir.

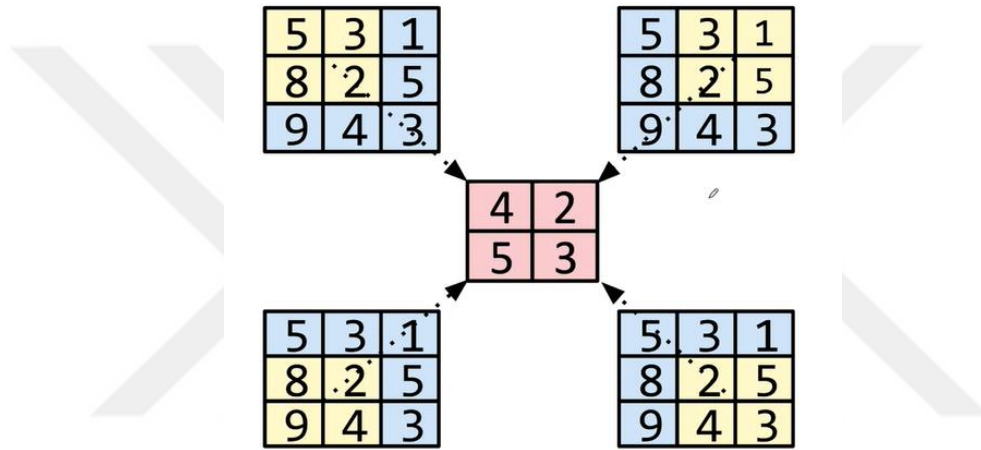


Şekil 3.3 Evrişim Katmanı Örneği (İnt.Kyn.3)

3.5.1.3 Havuzlama Katmanı

Havuzlama katmanı, evrişim katmanından elde edilen özellik haritalarının boyutunu küçültmek, hesaplama maliyetini azaltmak ve en belirgin bilgileri koruyarak gereksiz

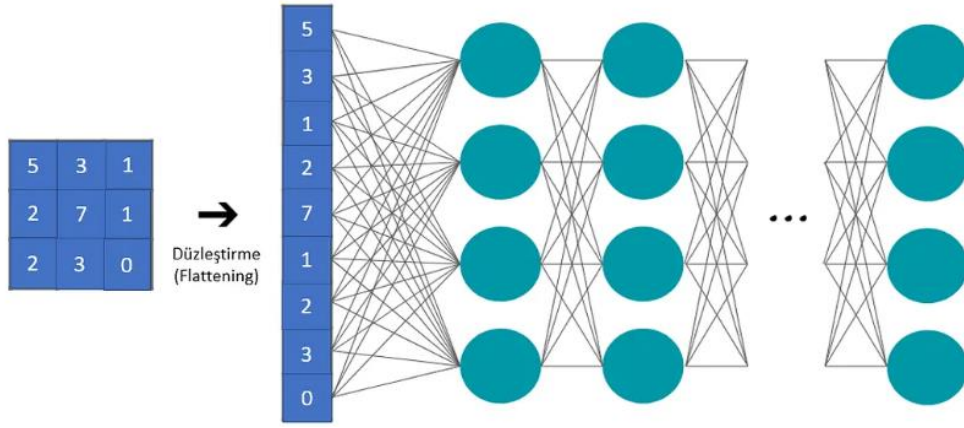
detayları eksiltmek için kullanılan bir katmandır. Genellikle evrişim katmanından hemen sonra ve tam bağlantı katmanından önce parametre sayısını azaltmak için kullanılır. Havuzlama, belirli bir pencere (örneğin, 2×2 veya 3×3 boyutunda bir filtre) kullanarak özet bir bilgi çıkarır. Bu pencere giriş görüntüsü üzerinde belirli bir adımla (stride) hareket eder ve her adımda havuzlama türüne göre değeri seçerek yeni bir çıktı matrisi oluşturur. Maksimum havuzlama ve ortalama havuzlama olmak üzere iki farklı havuzlama türü vardır. Bu tez çalışmasında ortalama havuzlama kullanılmıştır. Ortalama havuzlama örneği Şekil 3.4'te verilmiştir.



Şekil 3.4 Ortalama Havuzlama Katmanı Örneği (İnt.Kyn.3)

3.5.1.4 Tam Bağlantı Katmanı

Tam bağlantı katmanı, Evrişimli Sinir Ağlarının son aşamasında yer alan ve sınıflandırma işlemini gerçekleştiren katmandır. Bu katmanda, önceki katmanlardan gelen özellik haritaları, düzleştirilerek (flattening) bir vektöre dönüştürülür ve tamamen bağlı bir yapıya geçirilir. Yani evrişim ve havuzlama katmanından gelen matrisler bir vektöre dönüştürülerek tam bağlı hale getirilir. Tam bağlantılı katmanda, her nöron önceki katmandaki tüm nöronlarla bağlıdır. Bu sayede, veriler arasındaki tüm ilişkiler model tarafından öğrenilir. Her nöron, giriş değerleriyle bir ağırlık çarpımı alır ve bir aktivasyon fonksiyonuna tabi tutulur. Bu tez çalışmasında tam bağlantı katmanlar, gelen veriyi analiz ederek verinin hangi sınıfa ait olduğunu belirleyip nihai tahminleri üretmiştir. Şekil 3.5'te tam bağlantı katmanı örneği verilmiştir.



Şekil 3.5 Tam Bağlantı Katmanı Örneği (İnt.Kyn.4)

3.5.2 Transfer Öğrenme

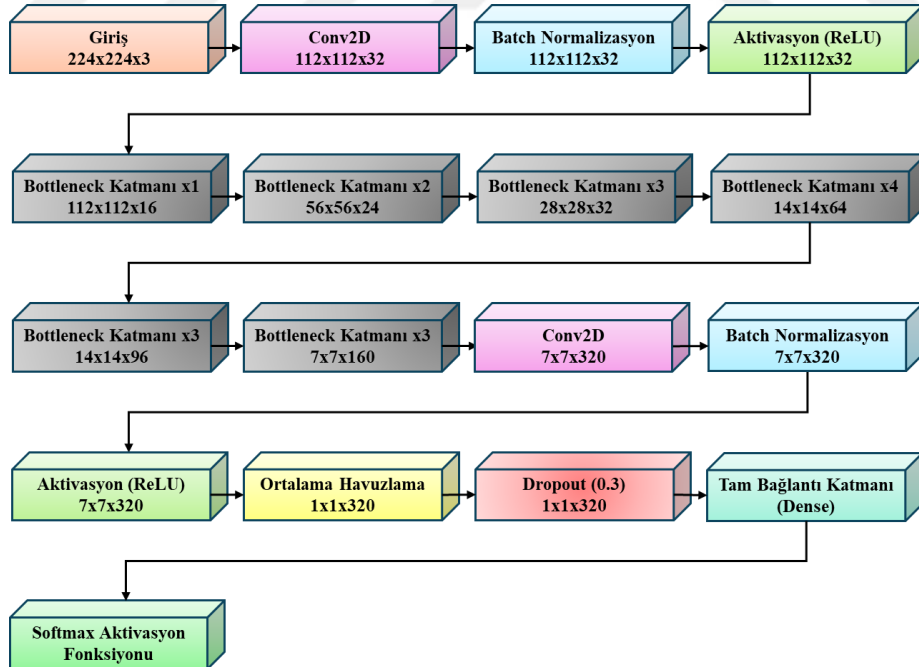
Derin öğrenme modelleri, özellikle evrişimli sinir ağları (CNN) gibi derin ağlar, genellikle büyük ölçekli veri kümeleri üzerinde eğitilir. Örneğin, ImageNet gibi milyonlarca görüntü içeren veri kümelerinde eğitilmiş bir model, çeşitli görsel özellikleri (kenarlar, dokular, nesneler vb.) öğrenmiştir. Transfer öğrenme sayesinde önceden eğitilmiş bir model alınır ve yeni bir problem için bu modelin bazı katmanları dondurulur bu şekilde önceden öğrenilmiş özellikler korunur (Weiss vd. 2016). Son katmanlar, yeni veri kümesine uygun olacak şekilde değiştirilir ve yeniden eğitilir. Bu süreç, daha az veri ve daha kısa eğitim süresiyle yüksek doğruluk elde edilmesini sağlar. Yeni bir modeli sıfırdan eğitmek için milyonlarca veri gerekebilirken, transfer öğrenme sayesinde daha küçük veri kümelerinde de başarılı sonuçlar elde edilir (Zhuang vd. 2021). Bu tez çalışmasında modellerin başlangıç katmanları dondurularak modellere ek katmanlar eklenmiş ve modeller transfer öğrenme yöntemi kullanılarak yeniden yapılandırılmıştır.

3.5.3 Derin Öğrenme Modelleri

Diyabetik retinopati aşamalarının sınıflandırılması için kullanılan derin öğrenme modelleri, görüntü işleme çalışmalarında yüksek performans gösteren çeşitli mimarilerden oluşmaktadır. Eğitim için sekiz farklı derin öğrenme modeli kullanılmıştır. Model seçiminde hem klasik hem de modern derin öğrenme mimarilerini kapsayarak geniş bir karşılaştırma yapılması amaçlanmıştır. Bu modeller, farklı yapıları ve avantajları sayesinde kapsamlı bir performans değerlendirmesi yapmayı mümkün kılmıştır.

3.5.3.1 MobileNetV2

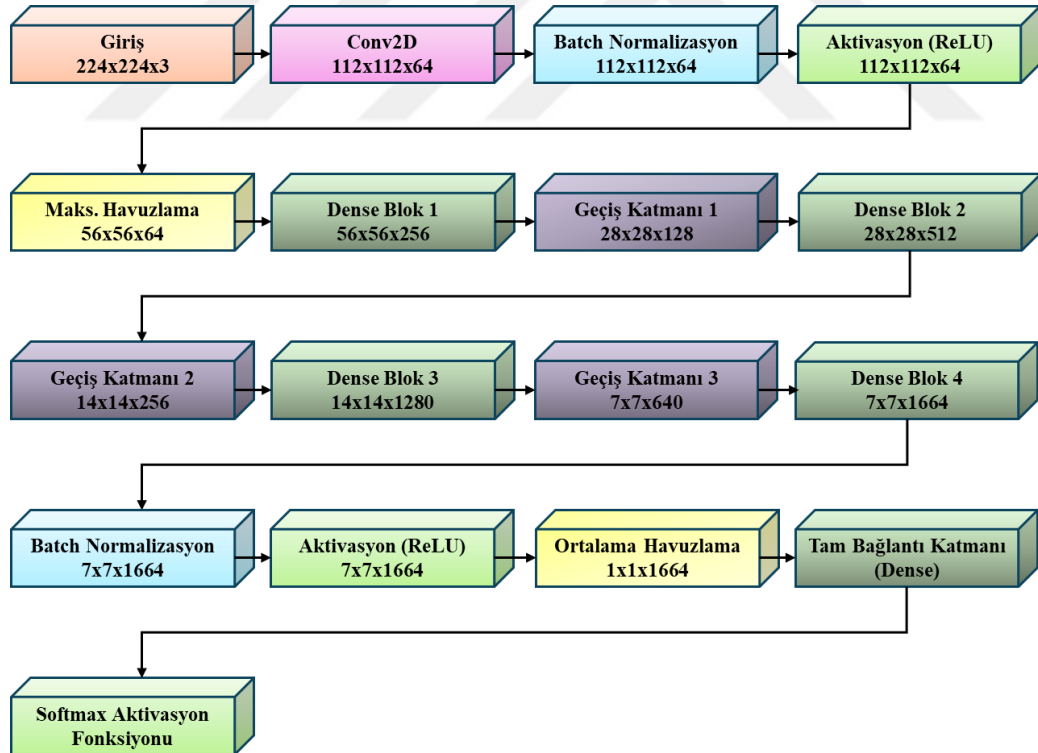
MobileNetV2 modeli, hafif mimarisi ve düşük hesaplama maliyeti ile bilinmektedir. Google tarafından 2018 yılında yayınlanan, hafif ve verimli bir derin öğrenme modelidir. Mobil cihazlar ve gömülü sistemler için optimize edilmiş olup, özellikle sınırlı hesaplama gücüne sahip donanımlarda yüksek doğruluk elde etmek amacıyla tasarlanmıştır.(Sandler vd. 2018). Bu model, MobileNetV1'in geliştirilmiş bir versiyonudur ve derin öğrenme tabanlı görüntü işleme görevlerinde hız, doğruluk ve model boyutu açısından önemli iyileştirmeler sunmaktadır. MobileNetV2'nin mimarisi, geleneksel konvolüsyonel sinir ağlarından farklı olarak derinlik ayırık (depthwise separable) evrişimlere ve inverted residual bloklara sahiptir. Bu yapı, hesaplama yükünü azaltırken modelin performansını artırır. Bu özelliğinden dolayı, karmaşık sınıflandırma görevleri için retina görüntülerindeki ince taneli doku özelliklerini diğer modellere kıyasla daha başarılı bir şekilde öğreneceği düşünülmüştür. Eğitim sürecinde, düşük bellek kullanımı hızlı sonuçlar alınmasını sağlamıştır. Şekil 3.6'da MobileNetV2 modeline ait mimari özet olarak verilmiştir.



Şekil 3.6 MobileNetV2 Mimarisi

3.5.3.2 DenseNet169

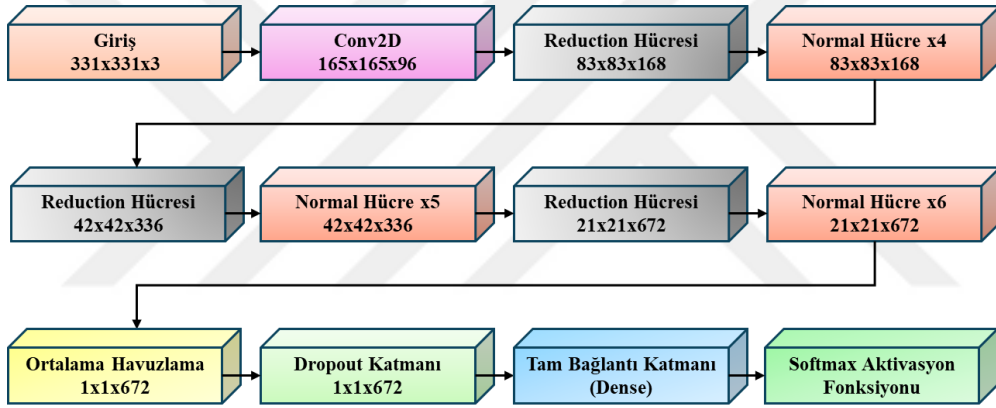
DenseNet (Densely Connected Convolutional Network), 2017 yılında Huang ve arkadaşları tarafından önerilen derin öğrenme modelidir. Bu model, geleneksel CNN'lerden farklı olarak her katmandaki çıktıyı, sonraki tüm katmanlara girdi olarak veren yoğun bağlantı yöntemini kullanır. DenseNet, diğer derin sinir ağıları gibi konvolüsyonel ve tam bağlantı katmanlarını kullanır ancak en büyük farkı katmanlar arasındaki bağlantı yapısıdır. DenseNet169, DenseNet mimarisinin orta büyüklükteki versiyonudur ve 169 katmandan oluşur. Daha az parametre ile daha yüksek doğruluk sunması, ağda bilgi kaybını azaltması ve daha verimli gradyan yayılımı sağlaması nedeniyle tercih edilmiştir (Krešo vd. 2018). DenseNet169, yoğun bağlantı katmanları aracılığıyla bilgileri verimli bir şekilde yayarak yeniden kullanır. Bu mimari, modelin sınırlı verilerle bile yüksek bir genelleme kapasitesi elde etmesini sağlamıştır. Şekil 3.7'de DenseNet169 mimarisinin özet yapısı bulunmaktadır.



Şekil 3.7 DenseNet169 Mimarisi

3.5.3.3 NASNet

NASNet, Google Brain ekibi tarafından geliştirilmiş ve evrimsel algoritmalar kullanılarak otomatik olarak tasarlanmış bir sinir ağı mimarisidir. Geleneksel derin öğrenme modellerinin aksine, NASNet'in mimarisi yapay zeka tarafından tasarlanarak insan tarafından manuel olarak tasarlanmamıştır (Şekil 3.8). NASNet, sınırlı veri kümelerinde güçlü performansa sahip olup, görüntü sınıflandırma ve bilgisayarla görme gibi çalışmalarda son derece yüksek doğruluk sağlayan bir CNN tabanlı derin öğrenme modelidir (Zoph vd. 2018). Bu mimari, otomatik olarak tasarlanmış hücre yapıları ve sağlam özellik çıkarma yetenekleri nedeniyle seçilmiştir. Bu tez çalışmasında NASNetLarge versiyonu eğitilmiştir.



Şekil 3.8 NASNet Mimarisi

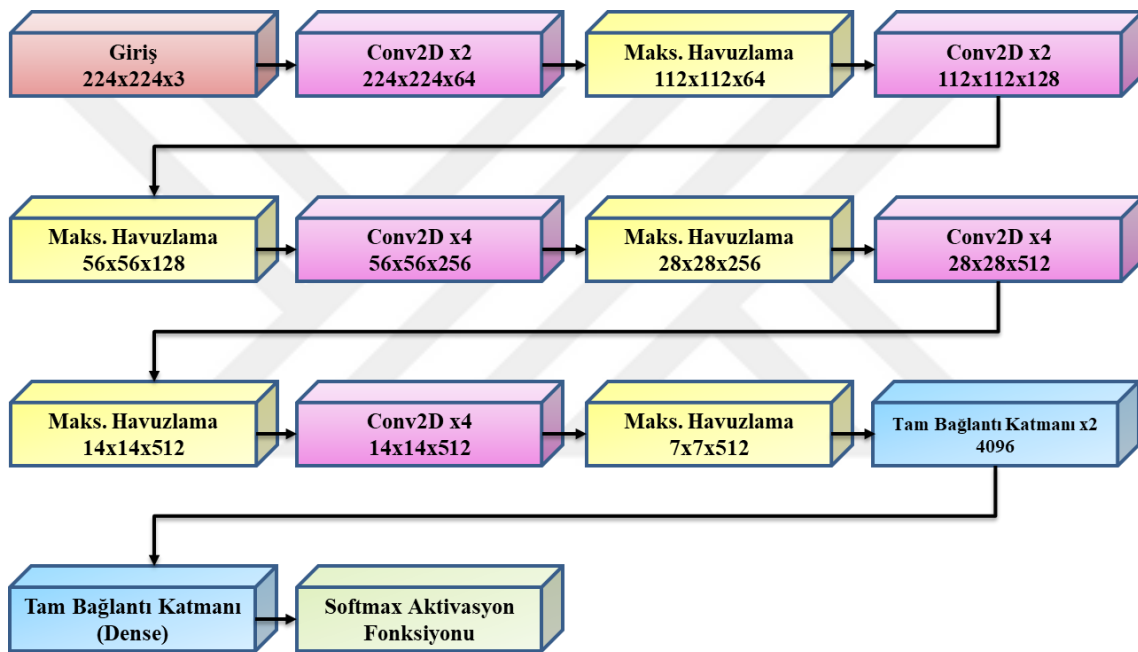
3.5.3.4 VGG19

VGG19, Oxford Üniversitesi'ndeki Görüntü Tanıma Grubu (Visual Geometry Group - VGG) tarafından geliştirilen derin bir evrişimli sinir ağı modelidir. 2014 yılında ImageNet Büyük Ölçekli Görüntü Tanıma Yarışması (ILSVRC-2014)'te en iyi modellerden biri olarak tanıtılmıştır. VGG grubu, VGG16 ve VGG19 olmak üzere iki temel modelden oluşur. VGG19, 19 katmanlı derin bir CNN mimarisidir ve VGG16'ya kıyasla daha fazla konvolüsyon katmanına sahiptir, bu da daha karmaşık özellikleri öğrenebilmesini sağlar.

VGG19, derin, basit ve düzenli bir yapıya sahip evrişimli sinir ağıdır. Model, küçük

filtreler (3x3 konvolüsyon) ve maksimum havuzlama katmanları (2x2) kullanarak görüntülerin özelliklerini derinlemesine çıkarır. Model, sabit boyutlu filtreler ve derin katmanlar kullanarak ayrıntılı özellik haritalarını ayıklar (Mascarenhas ve Agarwal 2021).

Diyabetik retinopatinin farklı aşamalarını temsil eden retina görüntülerinde, VGG19'un sabit boyutlu filtreleri tarafından elde edilen tutarlı özellik çıkarımı, modelin sınıflandırma performansını artırmıştır. Şekil 3.9'da VGG19 mimarisi verilmiştir.



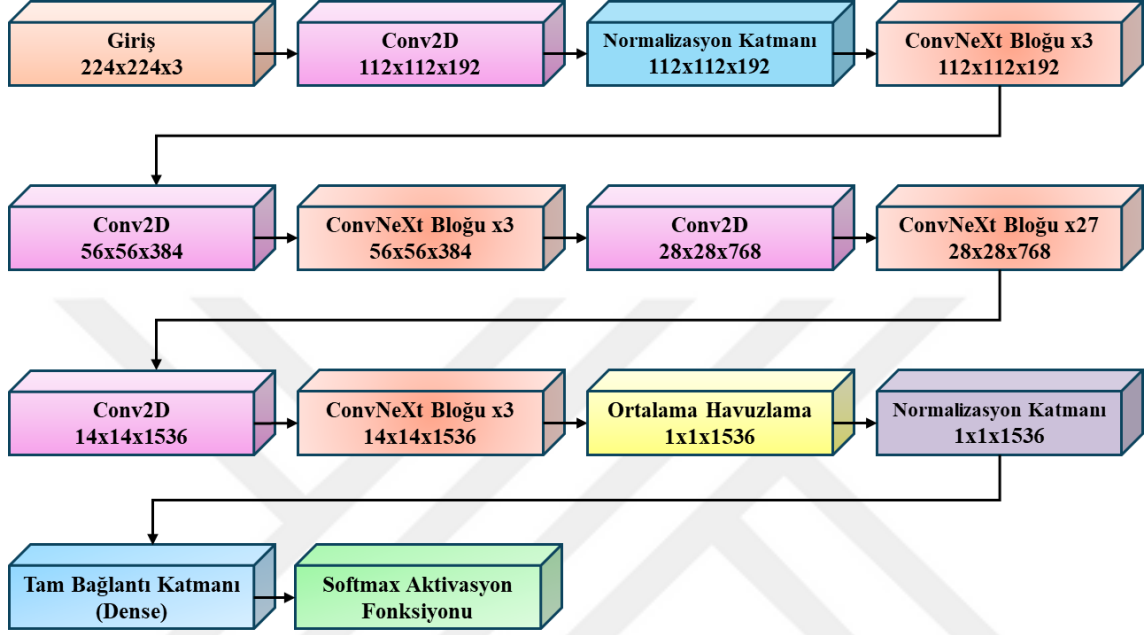
Şekil 3.9 VGG19 Mimarisi

3.5.3.5 ConvNeXt Large

ConvNeXt Large, Facebook AI Research (FAIR) ekibi tarafından 2022 yılında geliştirilmiş, evrişim tabanlı modern bir sinir ağı modelidir. ConvNeXt grubu (Tiny, Small, Base, Large, X-Large), ResNet mimarisinden esinlenerek oluşturulmuş ancak birçok yapısal iyileştirme içeren gelişmiş bir konvolüsyon tabanlı modeldir. ConvNeXt Large, ConvNeXt modellerinin büyük boyutlu versiyonudur ve yüksek performans gerektiren görüntü sınıflandırma ve bilgisayarla görme gibi çalışmalarda kullanılmak için uygundur. Transformatör tabanlı yaklaşımlardan ilham alan evrişimli sinir ağlarının

modernize edilmiş bir versiyonudur (Liu Z vd. 2022).

ConvNeXt Large modelinin geniş katmanları ve yüksek parametre sayısı, retina görüntülerinin özellikle karmaşık detaylarını öğrenmede başarılı olmuştur. Modelin büyük boyutu doğruluğu artırırken, eğitim süresini uzatmıştır (Şekil 3.10).

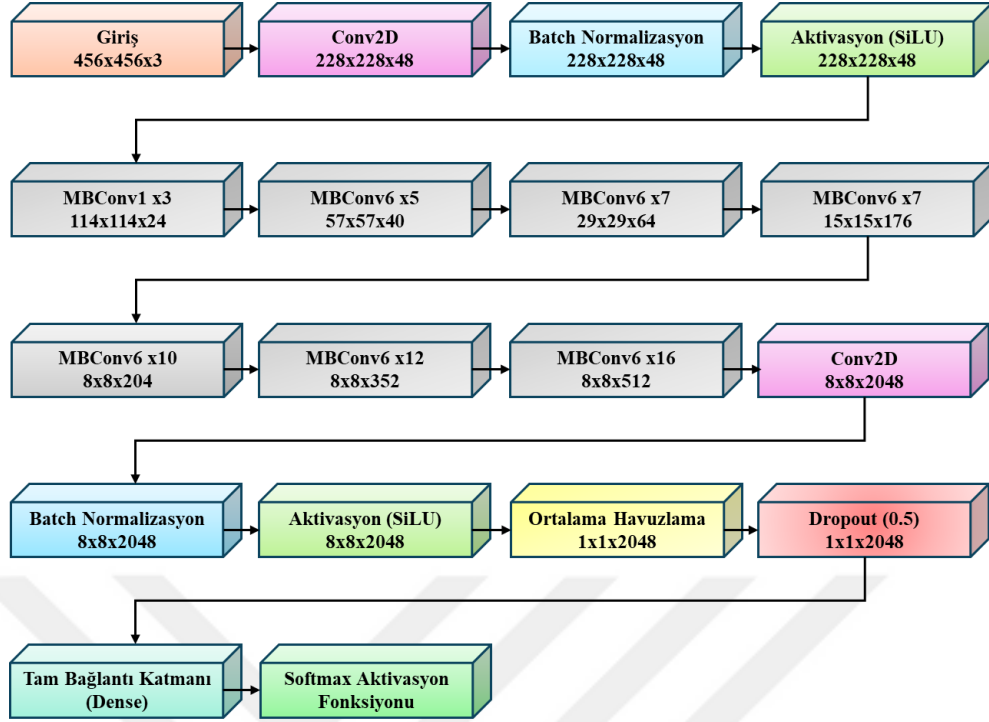


Şekil 3.10 ConvNeXt Large Mimarisi

3.5.3.6 EfficientNetB5

EfficientNetB5, Google Brain ekibi tarafından geliştirilen EfficientNet mimarisinin orta seviyedeki bir üyesidir. Bu model, "Compound Scaling" adı verilen özel bir ölçeklendirme yöntemi kullanarak model boyutunu ve hesaplama verimliliğini dengeler. EfficientNetB5, mobil cihazlardan büyük ölçekli sunuculara kadar geniş bir kullanım alanına sahiptir ve görüntü sınıflandırma, nesne algılama gibi çalışmalarda yüksek doğruluk oranları elde eder (Tan ve Le 2019).

Hem genişlik hem de derinlik açısından dengeli olan bu model, artırılmış görüntü çözünürlüğünü başarılı bir şekilde ele alarak retina görüntülerindeki mikroskobik ayrıntıları öğrenmiş ve yüksek sınıflandırma performansı elde etmiştir (Şekil 3.11).



Şekil 3.11 EfficientNetB5 Mimarisi

3.5.3.7 EfficientNetB6

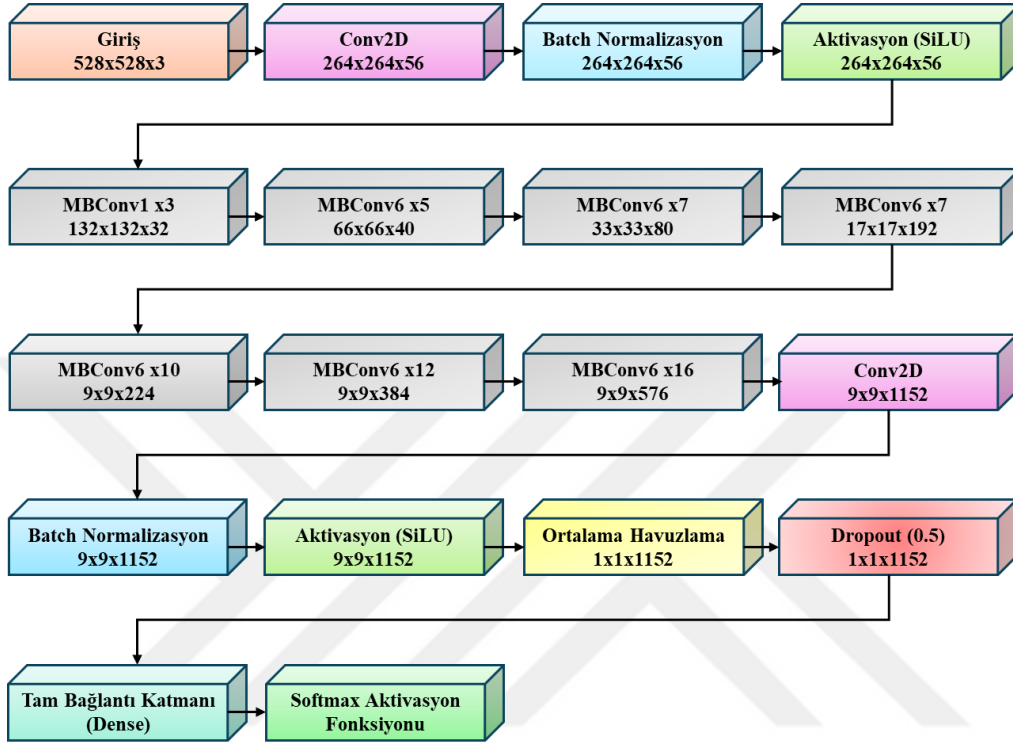
EfficientNetB6, EfficientNetB5 modeline kıyasla daha derin bir mimariye ve daha fazla sayıda parametreye sahiptir (Tan ve Le 2019). Bu özellikler, diyabetik retinopati sınıflandırmasında daha yüksek bir doğruluk elde etmesini sağlamıştır. EfficientNetB5 güçlü sonuçlar verdiği için, gelişmiş versiyonu olan EfficientNetB6 da deneyler için özellikle seçilmiştir. Şekil 3.12’de modele ait mimarinin özet versiyonu verilmiştir.

3.5.3.8 EfficientNetV2-S

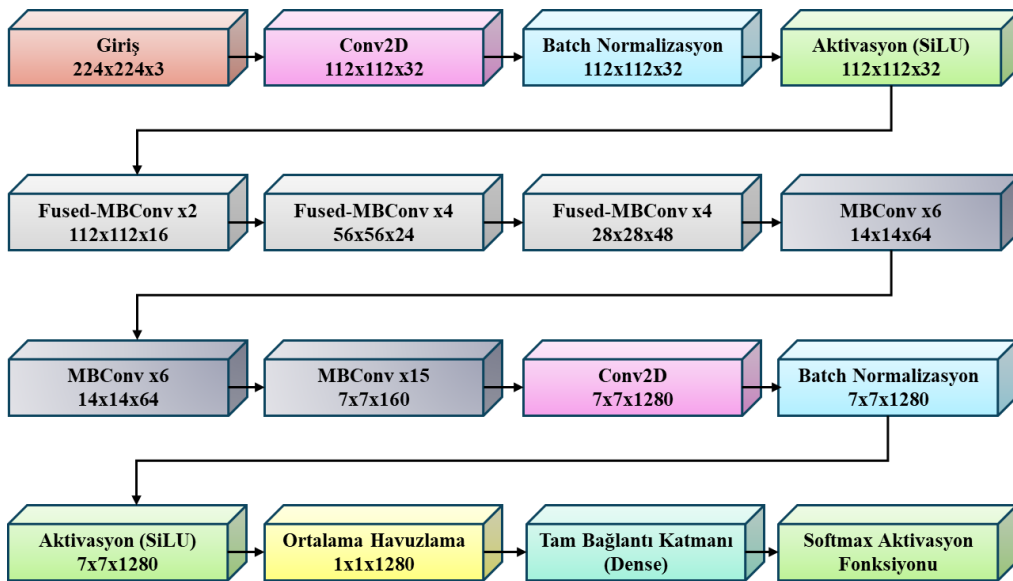
EfficientNetV2-S, Google Brain ekibi tarafından geliştirilmiştir. EfficientNetV2 mimarisi, önceki EfficientNet modellerinin eksikliklerini gidererek geliştirilmiştir. EfficientNetV2 modelleri EfficientNetV1’e göre daha hızlı eğitim süresi, daha yüksek doğruluk ve daha düşük hesaplama maliyeti sunmaktadır (Tan ve Le 2021).

Yapılan deneylerde EfficientNet modellerinin güçlü başarısı göz önüne alındığında, bu sürüm de değerlendirilmiştir. Model, görüntülerdeki ince ayrıntıları daha hızlı ve başarılı bir şekilde analiz etmek için tasarlanmış iç içe geçmiş evrimsel katmanları kullanarak

düşük çözünürlüklü ayrıntıların bile öğrenilmesini sağlamıştır (Şekil 3.13). Hızlı ve başarılı bir öğrenme süreci sağlarken karmaşık retina görüntülerinin analizini daha az parametreyle verimli bir şekilde gerçekleştirmiştir.



Şekil 3.12 EfficientNetB5 Mimarisi



Şekil 3.13 EfficientNetV2-S Mimarisi

Modeller, TensorFlow ortamında, karışık hassasiyetli hesaplamalar devre dışı bırakılarak, float32 hassasiyetinde çalışarak yürütülmüştür ve bu sayede daha kararlı sonuçlar elde edilmiştir.

3.5.4 Ek Katmanlar

Derin öğrenme modellerinde ek katmanlar, modelin son katmanlarının dondurulmasıyla birlikte eklenmektedir. Transfer öğrenme sürecinde, önceden eğitilmiş modellerin temel özellikleri çıkaran katmanları korunurken, yeni veri kümesine uyum sağlayacak şekilde tamamen bağlı katmanlar (Dense), düzenlileştirme katmanları (Dropout) ve çıkış katmanı (Softmax) eklenerek modelin performansı artırılmaktadır. Bu çalışmada, son katmanlar dondurulduktan sonra sırasıyla 2048 nöronlu bir Dense katmanı, 0.3 oranında Dropout katmanı, 5 nöronlu Dense katmanı ve Softmax aktivasyonu kullanılarak modellerin son katmanları oluşturulmuştur.

3.5.4.1 Dense Katmanı (2048 nöron)

Tam bağlı katmanlar (Dense katmanları), her bir nöronun bir önceki katmandaki tüm nöronlarla bağlantılı olduğu yapılar olup, derin sinir ağlarında genellikle sınıflandırıcı olarak kullanılmaktadır. Bu tez çalışmasında, transfer öğrenme modeli üzerine eklenen ilk ek katman 2048 nöronlu bir Dense katmanıdır. Bu yüksek nöron sayısı, modelin geniş bir özellik kümesini öğrenebilmesini sağlamaktadır. 2048 nöronlu bu katman, modelin önceki konvolüsyonel katmanlar tarafından çıkarılmış olan yüksek seviyeli özellikleri işlemektedir. Modelin doğrusal olmayan ilişkileri daha iyi öğrenebilmesi için bu katmanda ReLU (Rectified Linear Unit) aktivasyon fonksiyonu kullanılmıştır. ReLU, negatif girdileri sıfıra eşitleyerek ağırlıkların güncellenmesini kolaylaştırarak gradyan kaybolma (vanishing gradient) problemini önlemektedir (Javid vd. 2021).

Bu tam bağlantılı katmanın eklenmesi, modelin daha yüksek kapasiteli bir öğrenme mekanizmasına sahip olmasını sağlarken, aynı zamanda modelin son katmandaki karar mekanizmasını daha esnek hale getirerek farklı sınıfları daha doğru bir şekilde ayırt etmesine yardımcı olmaktadır.

3.5.4.2 Dropout Katmanı

Derin sinir ağıları, yüksek sayıda parametre içermesi nedeniyle aşırı öğrenme (overfitting) problemine yatkın hale gelebilmektedir. Bu problemi önlemek için düzenlileştirme (regularization) tekniklerinden biri olan Dropout katmanı kullanılmaktadır. Dropout, her eğitim iterasyonunda belirli oranlarda nöronların rastgele devre dışı bırakılmasıyla çalışarak modelin belirli nöronlara aşırı bağımlılık geliştirmesini engellemektedir.

Bu çalışmada, 0.3 oranında bir Dropout katmanı kullanılmıştır. Bu, her eğitim adımında nöronların %30'unun rastgele devre dışı bırakıldığı anlamına gelmektedir. Böylece model, eğitim sırasında belirli bir alt ağ üzerinden öğrenme yapmaktan kaçınmaktadır. Bu şekilde daha genel bir öğrenme sağlanarak aşırı uyum engellenmektedir. Dropout katmanının en büyük avantajı, modelin sadece belirli özelliklere aşırı bağımlı olmasını önleyerek, daha dengeli ve genelleştirilebilir bir öğrenme süreci sağlamasıdır. Bu sayede, eğitim verisine aşırı uyum sağlayarak doğruluğu düşürme riski azaltılmış olur (Park ve Kwak 2017).

3.5.4.3 Dense Katmanı (5 nöron)

Modelin başarılı bir sınıflandırma gerçekleştirebilmesi için son aşamada 5 nöronlu bir Dense katmanı eklenmiştir. Bu katman, modelin çıktısının kaç sınıf üzerinden değerlendirileceğini belirlemektedir. Bu çalışmada, 5 farklı sınıf bulunduğundan, çıktı katmanında 5 nöron bulunmaktadır ve her bir nöron, modelin belirli bir sınıfını temsil etmektedir. Bu aşamada kullanılan 5 nöron, modelin çok sınıflı (multi-class) bir sınıflandırma üzerinde çalıştığını göstermektedir. Eğer problem iki sınıflı (binary classification) olsaydı, çıktı katmanında yalnızca 1 nöron bulunur ve sigmoid aktivasyon fonksiyonu kullanılırdı. Ancak çok sınıflı problemler için Softmax, en yaygın tercih edilen aktivasyon fonksiyonudur.

3.5.4.4 Softmax Aktivasyon Fonksiyonu

Softmax fonksiyonu, birden çok sınıf arasında bir olasılık dağılımı sağladığı için çok sınıflı sınıflandırma problemleri için en uygun olan aktivasyon fonksiyonudur. Softmax,

her nöronun çıktı değerini bir olasılık dağılımına dönüştürerek, modelin hangi sınıfa daha yüksek güvenle karar verdiğini belirlemesini sağlamaktadır. Softmax fonksiyonu her sınıf için olasılığı hesaplamaktadır. Denklem 3.1’de , softmax aktivasyon fonksiyonunun denklemi verilmiştir (Sharma S vd. 2017).

$$\text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}} \quad (3.1)$$

Burada, z_i modelin çıktıları, e doğal logaritmanın tabanı ve n toplam sınıf sayısını göstermektedir. Bu formül sayesinde, modelin çıktıları bir olasılık dağılımı haline gelir ve toplamaları 1 olacak şekilde normalize edilir.

Softmax aktivasyon fonksiyonu modelin her bir sınıfı için bir olasılık tahmini sunar. Örneğin, bir görüntünün belirli bir hastalık sınıfına ait olup olmadığını tahmin eden bir model, çıktı olarak her sınıf için bir olasılık değeri döndürecektir. Bu sayede, model yalnızca en yüksek olasılığa sahip olan sınıfı belirlemekle kalmaz, aynı zamanda farklı sınıflar arasındaki belirsizlik seviyesini de gösterir.

Özet olarak bu tez çalışmasında, transfer öğrenme uygulanarak, önceden eğitilmiş modellerin son katmanları dondurulmuş ve yeni bir sınıflandırıcı baştan oluşturulmuştur. Modelin özelleştirilmesi için 2048 nöronlu bir Dense katmanı, aşırı öğrenmeyi önlemek için 0.3 Dropout oranına sahip bir düzenleme katmanı, 5 nöronlu bir Dense çıktı katmanı ve çok sınıflı sınıflandırma için Softmax aktivasyon fonksiyonu kullanılmıştır. Bu sayede model, geniş bir özellik kümesini işleyerek daha iyi genelleştirme yapabilmekte ve aşırı uyum riskini minimize etmektedir.

3.5.5 Eğitim

Makine öğrenimi ve derin öğrenme modellerinin başarısı büyük ölçüde eğitim sürecine bağlıdır. Eğitim, modelin belirli bir veri kümesi üzerinden öğrenmesini sağlamak amacıyla gerçekleştirilmektedir. Eğitim esnasında model, giriş verilerinden anlamlı kalıplar çıkarmayı ve bu kalıpları kullanarak yeni verilere yönelik doğru tahminlerde bulunmayı öğrenir.

Derin öğrenme modellerinde eğitim genellikle denetimli öğrenme (supervised learning) yaklaşımıyla gerçekleştirilir. Bu yöntemde model, giriş verileriyle birlikte bu verilere ait doğru etiketleri kullanarak eğitilir. Model, her iterasyonda tahminleri ile gerçek etiketler arasındaki farkı hesaplayan bir kayıp fonksiyonunu (loss function) en aza indirmeye çalışır. Kayıp fonksiyonunun çıktısına bağlı olarak modelin ağırlıkları bir optimizasyon algoritması aracılığıyla güncellenir ve modelin daha doğru tahminler yapması sağlanır.

3.5.5.1 Batch Size

Derin öğrenme modellerinin eğitim sürecinde batch size (mini-batch boyutu), her iterasyonda modele verilen verilerin sayısını belirleyen bir hiperparametredir. Eğitim sırasında model, veri kümesini doğrudan tüm örnekleri kullanarak değil, belirli boyutlardaki mini-batch'ler halinde işleyerek ağırlıklarını günceller.

Batch size seçimi, modelin bellek kullanımı, öğrenme sürecinin kararlılığı ve eğitim süresinin verimliliği açısından önemlidir. Küçük batch boyutları (örneğin 16 veya 32), modelin daha genel bir öğrenme gerçekleştirmesine yardımcı olurken, büyük batch boyutları (örneğin 128 veya 256) daha hızlı eğitim sağlamasının yanı sıra öğrenme kapasitesini düşürebilir. Bu çalışmada, modelin eğitimi için batch size 16 olarak belirlenmiştir. Bu değer yapılan deneyler sonucunda en başarılı performansı göstermesi nedeniyle tercih edilmiştir.

3.5.5.2 Sınıf Ağırlıkları

Derin öğrenme modellerinin eğitiminde en önemli sorunlardan biri sınıf dengesizliği problemidir. Veri setinde bazı sınıfların diğerlerine kıyasla çok daha fazla örneğe sahip olması, modelin bu baskın sınıfları öğrenme eğiliminde olmasına ve az bulunan sınıfların öğrenilmemesine neden olmaktadır. Bu durum, özellikle nadir görülen hastalıkların teşhis edilmesi gibi tıbbi görüntüleme alanındaki çalışmalarda oldukça önemli bir sorundur.

Sınıf dengesizliği problemini gidermek için sınıf ağırlıkları kullanılmıştır. Sınıf ağırlıkları, her bir sınıfın veri setindeki örnek sayısı ile ters orantılı olacak şekilde

hesaplanmıştır. Daha az bulunan sınıflar için daha büyük ağırlıklar eklenmiştir. Böylece modelin bu sınıflara daha fazla önem vermesi sağlanmıştır. Bu ağırlıklar, modelin kayıp fonksiyonuna entegre edilerek eğitim sürecinin dengesiz veri dağılımından etkilenmemesi amaçlanmıştır.

3.5.5.3 Ağırlıklı Kategorik Çapraz Entropi Kayıp Fonksiyonu

Modelin eğitimi sırasında kullanılan kayıp fonksiyonu, modelin hatalarını hesaplayarak doğru tahminler yapabilmesi için yönlendirilmesini sağlamaktadır. Bu tez çalışmasında ağırlıklı kategorik çapraz entropi kayıp fonksiyonu kullanılmıştır. Standart çapraz entropi kayıp fonksiyonu, tüm sınıfları eşit önemde değerlendirirken, ağırlıklı versiyonu veri setinde daha az bulunan sınıfların kaybını daha önemli hale getirerek modelin nadir görülen sınıfları daha iyi öğrenmesine yardımcı olur. Böylece, model yalnızca baskın sınıflara öğrenmez, aynı zamanda tüm sınıflar için dengeli bir öğrenme gerçekleştirir (He ve Garcia 2009).

3.5.5.4 Adam Optimizasyon Algoritması

Bu çalışmada modelin parametre güncellemeleri için Adam optimizasyon algoritması kullanılmıştır. Adam, Adaptif Moment Tahmini (Adaptive Moment Estimation) yöntemine dayalı olup, Stokastik Gradyan İnişi (SGD) algoritmasının bir uzantısıdır. Adam, öğrenme oranını her parametre için dinamik olarak ayarlayarak daha hızlı ve verimli bir optimizasyon süreci sağlar. İlk momentum (momentum 1) geçmiş gradyanların hareketli ortalamasını tutarken, ikinci momentum (momentum 2) ise gradyanların karelerinin hareketli ortalamasını tutar. Bu sayede öğrenme oranını otomatik olarak ayarlayarak eğitim sürecini hızlandırır (Kingma ve Ba 2014).

3.5.5.5 Geri Çağırma Fonksiyonları (Callbacks) ve Aşırı Uyumun Önlenmesi

Aşırı uyum, modelin eğitim verisine fazla uyum sağlaması ve bu nedenle yeni verilere karşı genelleme yapamaması durumudur. Bu durumda model, eğitim verisindeki gürültüyü ve rastgele örüntüleri öğrenerek, doğrulama veya test verisinde düşük

performans gösterir. Aşırı uyumun belirtileri arasında eğitim doğruluğunun hızla artarken doğrulama doğruluğunun aynı seviyede kalması veya düşmesi yer alır. Eğitim sürecinde modelin aşırı uyum (overfitting) göstermemesi için EarlyStopping ve ReduceLROnPlateau gibi geri çağırma fonksiyonları kullanılmıştır. Bunun yanı sıra aşırı uyumu önlemek amacıyla eğitim sürecinde dropout gibi düzenleme teknikleri de kullanılmıştır.

EarlyStopping fonksiyonu modelin doğrulama kaybında belirli bir süre boyunca iyileşme gözlemlenmediğinde eğitimi erken durdurarak, modelin gereğinden fazla eğitilmesini ve aşırı uyum göstermesini engellemektedir (Heaton 2018). Bu çalışmada, ardışık 10 epok boyunca doğrulama kaybında herhangi bir iyileşme görülmezse, bu fonksiyon sayesinde eğitim durdurulmuştur.

ReduceLROnPlateau modelin öğrenme oranını (learning rate) dinamik olarak ayarlayan bir fonksiyondur. Eğitim sürecinde doğrulama kaybı belirli bir süre boyunca değişmediğinde öğrenme oranı azaltılarak modelin daha küçük adımlarla öğrenmesi sağlanır. Bu sayede modelin ağırlıkları aniden büyük değişimler göstermeden daha kararlı bir şekilde değiştirilir (Heaton 2018). Bu çalışmada, öğrenme oranı başlangıçta 0,0005 olarak ayarlanmış ve doğrulama kaybında herhangi bir iyileşme görülmediğinde öğrenme oranı %20 oranında azaltılmıştır.

3.5.6 İnce Ayar (Fine Tuning)

İnce ayar, makine öğrenimi ve derin öğrenmede yaygın olarak kullanılan, belirli bir amaç için önceden eğitilmiş bir modeli yeniden yapılandırma işlemidir. Transfer öğrenme yöntemlerinin bir parçası olan bu teknik, model tarafından önceki eğitimden öğrenilen genel özellikleri korurken ince ayarlamalar yapılmasına izin verir. Bu süreçte, modelin mimarisi ve öğrenilen ağırlıklar büyük ölçüde korunur ve sıfırdan eğitim yerine, zaten var olan güçlü bir temel üzerine ince ayar yapılır. Özellikle daha önce eğitilen katmanlarda yapılan ayarlamalar modelin performansını artırır. İnce ayar, modelin önceden öğrenilmiş ağırlıklarında önemli değişiklikleri önlemek için düşük bir öğrenme oranıyla gerçekleştirilir.

Bu tez çalışmasında, modelin tüm başlangıç katmanları yeniden eğitime açılmış, ancak modelin önceki öğrenme sürecinde elde ettiği istatistikleri korumak amacıyla BatchNormalization katmanları dondurulmuştur (eğitimden çıkarılmıştır). BatchNormalization katmanlarının dondurulması, modelin önceki dağılım bilgilerini koruyarak modelin ağırlıklarının aşırı değişmesini önlemiştir (Ioffe ve Szegedy 2015). Bunun yanı sıra, diğer tüm katmanlar eğitilebilir olarak bırakılarak modelin daha iyi bir öğrenme süreci geçirmesi hedeflenmiştir. Bu aşamada, modelin önceden eğitilmiş ağırlıklarını büyük ölçüde değiştirmeden yeni veriye uyum sağlamasını sağlamak amacıyla, öğrenme oranı (learning rate) 0.00001 gibi son derece düşük bir değere ayarlanmıştır. Düşük öğrenme oranı, modelin küçük adımlarla güncellenmesine olanak tanıyarak aşırı uyum riskini azaltmış ve modelin daha kararlı bir şekilde öğrenmesini sağlamıştır.

İnce Ayar aşamasında batch size 16 olarak ayarlanmış ve iterasyon sayısı 15 ile sınırlandırılmıştır. EarlyStopping geri çağırma fonksiyonu, doğrulama kaybı ard arda beş iterasyon boyunca düzelmezse eğitimi sonlandıracak şekilde yapılandırılmıştır. İlk eğitim aşamasına benzer şekilde, ReduceLROnPlateau geri çağırma fonksiyonu, doğrulama kaybında herhangi bir iyileşme tespit edilmediğinde öğrenme oranını %20 oranında azaltmaya devam etmiştir.

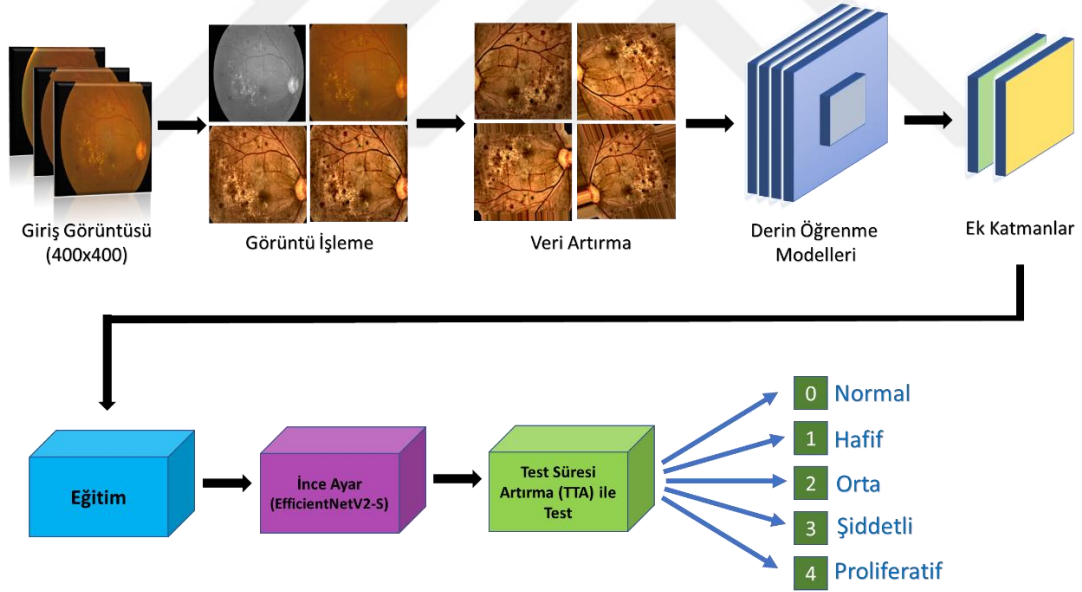
3.5.7 Test Zamanı Artırma (TTA)

Derin öğrenme modellerinin genelleme yeteneğini artırmak ve test aşamasında daha güvenilir tahminler elde etmek amacıyla kullanılan tekniklerden biri Test Zamanı Artırma (Test-Time Augmentation-TTA) yöntemidir. TTA, modelin test aşamasında aynı görüntü üzerinde farklı veri artırma işlemleri uygulayarak birden fazla tahminde bulunmasını ve bu tahminlerin ortalamasını alarak daha güvenilir sonuçlar üretmesini sağlar (Howard 2013).

İki aşamada eğitilen model, test verileri üzerinde değerlendirilmiştir. Test sürecinde Test Zamanı Artırma (TTA) yöntemi uygulanmıştır. Test Zamanı Artırma (TTA) yöntemi,

rotasyonlar (0° ile 180° arasında rastgele), yatay ve dikey çevirme (yatay çevirme, dikey çevirme hem yatay hem dikey çevirme ve çevirme yok) ve ölçeklendirmeden ($0,1x$ ila $0,9x$ arasında değişen) oluşmaktadır. Bu işlemler, test verilerindeki her görüntünün farklı varyasyonlarını elde etmek için uygulanmıştır. Her bir görüntüden orijinali de dahil olmak üzere toplam 10 görüntü elde edilmiştir. Model, bu 10 varyasyonun her biri için ayrı ayrı tahminlerde bulunmuş ve sonuçların ortalamasını alarak nihai kararı vermiştir. Bu yaklaşım, modelin farklı veri varyasyonları üzerinde performans değerlendirmesine olanak tanıyarak daha genel ve sağlam sonuçlar elde etmesini sağlamıştır. Örneğin, bir test görüntüsünü orijinal biçiminin yanı sıra döndürülmüş, ölçeklendirilmiş ve çevrilmiş versiyonlarında inceleyen model, tahminlerinde olası sapmayı azaltmıştır.

Bu tez çalışmasında kullanılan yöntemlerin özeti olarak Şekil 3.14'te akış şeması verilmiştir.



Şekil 3.14 Akış Şeması

4. BULGULAR

Bu çalışma Python programlama dili kullanılarak gerçekleştirilmiştir. Derin öğrenme modellerinin tasarlanması için TensorFlow kütüphanesi tercih edilmiştir. Modelin eğitimi ve değerlendirmesi, NVIDIA Tesla T4x2 GPU kullanılarak Kaggle platformunda gerçekleştirilmiştir.

4.1 Eğitim Sonuçları

Bu tez çalışmasında tüm deneylerde Adam optimizasyon algoritması kullanılmıştır (Kingma ve Ba 2014). Aynı zamanda kategorik çapraz entropi kayıp fonksiyonu kullanılmış olup bu tür sınıflandırma problemleri için uygun olduğundan tercih edilmiştir. Eğitim başlangıçta 75 epok (iterasyon) olarak planlanmış olsa da model performansını iyileştirmek için erken durdurma tekniği uygulanmıştır. Bu teknik, ard arda 10 epok boyunca doğrulama kaybında herhangi bir iyileşme gözlenmezse eğitimi sonlandırmaktadır. Eğitim sırasında, öğrenme oranı 0,0005 olarak ayarlanmış ve ReduceLROnPlateau geri çağırma fonksiyonu, doğrulama kaybında herhangi bir iyileşme tespit edilmediğinde öğrenme oranını %20 oranında azaltmıştır.

Sekiz farklı derin öğrenme modeliyle (MobileNetV2, DenseNet169, NASNet, VGG19, ConvNeXt Large, EfficientNetB5, EfficientNetB6 ve EfficientNetV2-S) gerçekleştirilen eğitim sonuçları Çizelge 4.1'de sunulmuştur.

Çizelge 4.1 Derin öğrenme modellerinin eğitim sonuçları

Model	Epok Sayısı	Eğitim Kaybı	Eğitim Doğruluğu (%)	Doğrulama Kaybı	Doğrulama Doğruluğu (%)
MobileNetV2	29	0.5321	80.54	0.6872	78.96
DenseNet169	23	0.2705	90.87	0.6290	80.25
NASNet	26	0.4213	84.36	0.5347	81.46
VGG19	21	0.3787	87.08	0.5144	83.07
ConvNeXt Large	24	0.3361	89.58	0.4556	85.89
EfficientNetB5	22	0.2454	91.78	0.2901	88.01
EfficientNetB6	22	0.2175	92.06	0.2441	90.54
EfficientNetV2-S	23	0.1342	94.63	0.1919	93.00

Çizelge 4.1'de sunulan sonuçlar, farklı derin öğrenme modellerinin eğitim ve doğrulama performansını kapsamlı bir şekilde karşılaştırmaktadır. Bu bağlamda, modellerin epok sayısı, eğitim kaybı, eğitim doğruluğu, doğrulama kaybı ve doğrulama doğruluğu gibi temel metrikler incelenerek, modellerin genelleme yetenekleri değerlendirilebilmektedir.

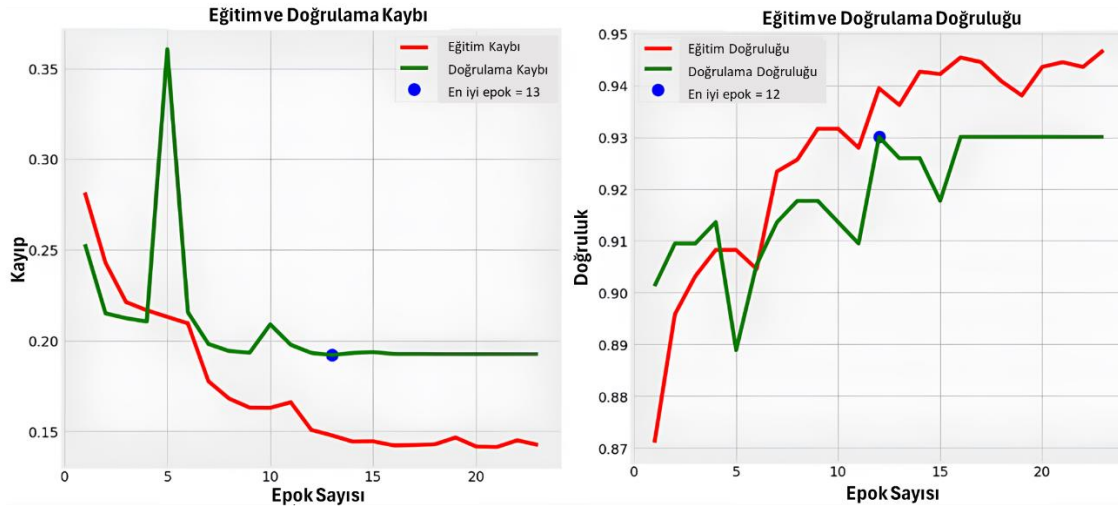
Tablodaki veriler incelendiğinde, EfficientNetV2-S modelinin diğer modellerle kıyaslandığında en düşük eğitim kaybına (0.1342) ve en yüksek eğitim doğruluğuna (%94.63) ulaştığı görülmektedir. Aynı zamanda, bu model doğrulama aşamasında da en düşük doğrulama kaybına (0.1919) ve en yüksek doğrulama doğruluğuna (%93.00) sahip olmasıyla ön plana çıkmıştır. Bu durum, EfficientNetV2-S modelinin eğitim sürecinde başarılı bir öğrenme sağladığını ve doğrulama verisi üzerinde en yüksek başarıyı gösterdiğini ortaya koymaktadır.

Diğer modellerle karşılaştırıldığında, MobileNetV2 modelinin eğitim kaybının 0.5321, eğitim doğruluğunun ise %80.54 olduğu görülmektedir. Bu modelin doğrulama aşamasında en yüksek doğrulama kaybına (0.6872) sahip olduğu ve doğrulama doğruluğunun %78.96 seviyesinde kaldığı gözlemlenmiştir. Bu sonuçlar, MobileNetV2 modelinin başarılı bir eğitim süreci geçirmediğini ve doğrulama aşamasındaki performansının diğer modellere kıyasla düşük kaldığını göstermektedir.

VGG19 modelinin eğitim doğruluğu %87.08 olmasına rağmen doğrulama doğruluğu %83.87 seviyesinde kalmıştır. Modelin doğrulama kaybının 0.5144 olduğu göz önüne alındığında, eğitim sırasında nispeten başarılı olmasına rağmen doğrulama verisinde belirgin bir performans düşüşü yaşadığı görülmektedir. Bu durum, modelin eğitim verisine aşırı uyum sağladığını (overfitting) göstermektedir.

EfficientNetB5 ve EfficientNetB6 modelleri de yüksek doğruluk oranlarına ulaşmış olup, sırasıyla %91.78 ve %92.06 eğitim doğruluğu sağlamıştır. Bu modellerin doğrulama doğrulukları ise %88.01 ve %88.94 olarak ölçülmüştür. Bu durum, EfficientNet tabanlı modellerin genel olarak yüksek doğruluk oranlarına sahip olduğunu ve diğer modellere kıyasla daha başarılı sonuçlar verdiğini göstermektedir.

Özet olarak eğitim sonuçları, EfficientNetV2-S modeli ile en düşük kayıp ve en yüksek doğruluğun elde edildiğini ortaya koymaktadır. Modelin sahip olduğu Fused-MBConv katmanları, geleneksel derin öğrenme modellerine kıyasla erken aşamalarda daha iyi bilgi akışı sağlayarak, düşük seviyeli kenar, doku ve damar yapılarını daha ayrıntılı olarak yakalamasını sağlamıştır. Bu, diyabetik retinopatinin sınıflandırılmasında oldukça önemli olan mikroanevrizmaların ve küçük kanamaların tespitini kolaylaştırmıştır. Böylece, modelin küçük ölçekli anormallikleri daha yüksek doğrulukla yakalaması ve detay düzeyinde daha tutarlı tahminler üretmesi mümkün hale gelmiştir. EfficientNetV2-S modeline ait eğitim grafikleri Şekil 4.1'de sunulmuştur.



Şekil 4.1 EfficientNetV2-S modelinin eğitim grafikleri

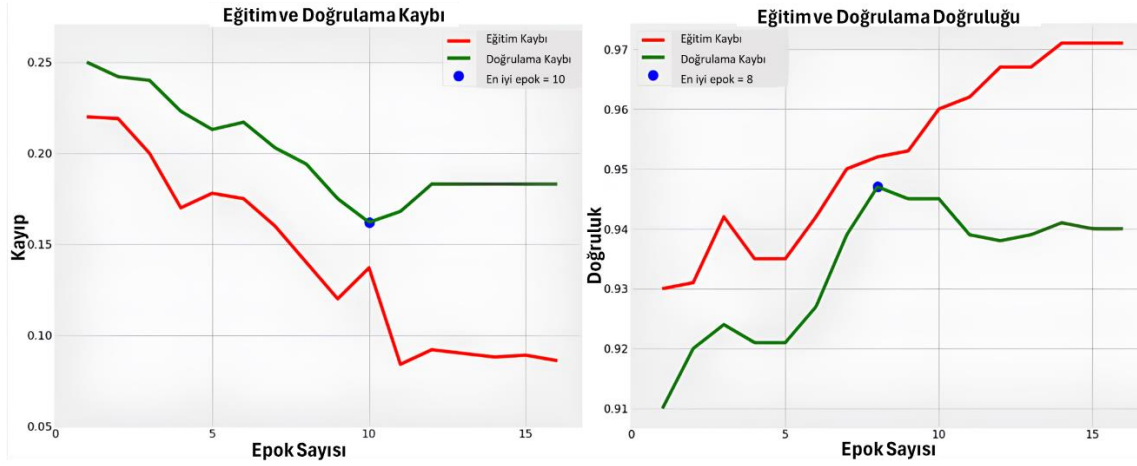
4.2 İnce Ayar Sonuçları

Bu tez çalışmasında, düzenli eğitim sürecinde en iyi performansı gösteren EfficientNetV2-S modeli, daha yüksek doğruluk elde etmek ve modelin başarısını artırmak amacıyla ince ayar (fine-tuning) işlemine tabi tutulmuştur. İnce ayar süreci, önceden eğitilmiş modelin belirli katmanlarının yeniden eğitilebilir hale getirilerek, eğitim verisinden daha fazla öğrenme sağlanmasına dayanmaktadır. Bu işlemle modelin doğrulama verisi üzerindeki performansını artırarak, daha başarılı bir sınıflandırma performansı elde etmesi hedeflenmiştir.

İnce ayar sonucunda, EfficientNetV2-S modeli 10. epokta 0.162 doğrulama kaybına

ulaşarak en düşük hata seviyesine erişmiştir. Ayrıca, modelin doğrulama doğruluğu 8. epokta %94.7'ye ulaşarak en yüksek seviyeye çıkmıştır. Bu değerler, modelin ince ayar süreci boyunca istikrarlı bir şekilde öğrenmeye devam ettiğini ve belirli bir noktada en iyi performansa ulaştığını göstermektedir.

Gerçekleştirilen ince ayar süreci, EfficientNetV2-S modelinin eğitim ve doğrulama performansını önemli ölçüde iyileştirmiştir. Model, doğrulama doğruluğunu artırırken doğrulama kaybını düşürerek, daha yüksek bir başarı oranı yakalamıştır. İnce Ayar sonucunda EfficientNetV2-S modeline ait eğitim grafikleri Şekil 4.2'de verilmiştir.



Şekil 4.2 EfficientNetV2-S modelinin ince ayar eğitim grafikleri

Bu tez çalışmasında iki aşamalı model eğitiminin toplam süresi yaklaşık 6 saat sürmüştür. Her iki eğitim aşamasında da ortalama GPU bellek kullanımı 12,8 GB olarak ölçülmüştür.

4.3 Test Sonuçları

Derin öğrenme modellerinin başarısını anlamak ve klinik ortamda uygulanabilirliğini değerlendirmek için eğitim sonrası test sürecine geçilmiştir. Bu aşamada, model daha önce hiç görmediği ve herhangi bir işleme tabi tutulmayan test verisi üzerinde çalıştırılarak sınıflandırma başarısı detaylı olarak analiz edilmiştir.

Test verileri Test Zamanı Artırma (TTA) yöntemi ile test edilmiştir. Bu yöntem

sayesinde, modelin genelleme kapasitesi test aşamasında geliştirilmiş ve daha güvenilir sonuçlar elde edilmiştir. Test veri kümesindeki tüm görüntüler için çıktı oluşturma da dahil olmak üzere TTA süreci yaklaşık 316 saniye sürmüştür. TTA sürecindeki her görüntü için birden fazla artırılmış görüntü işlendiğinden, çıkarım süresi standart tek geçişli tahminlere göre nispeten daha uzundur. Bununla birlikte, modelin hafif yapısı ve az sayıda parametreye sahip olması nedeniyle, büyük ve karmaşık modellere kıyasla zaman olarak iyi bir performansa sahiptir. Ek olarak, TTA yönteminin uygulanması, çıkarım süresindeki artışa rağmen, modelin tahminlerinin kararlılığını ve genelleştirme yeteneğini artırmakta, bu da pratik uygulamalarda daha güvenilir sonuçlar elde edilmesini sağlamaktadır.

Modelin performansını değerlendirirken çeşitli metrikler kullanılmıştır. Bu metrikler 4 farklı değerle hesaplanmaktadır. Bunlardan ilki Doğru Pozitif (True Positive-TP) modelin, gerçekten hastalıklı olan bir görüntüyü doğru bir şekilde hastalıklı olarak sınıflandırmasıdır. Yani, modelin doğru teşhis ettiği hasta bireyleri ifade etmektedir. Bir diğeri Doğru Negatif (True Negative-TN) modelin, sağlıklı bir bireyin görüntüsünü doğru bir şekilde sağlıklı olarak sınıflandırmasıdır. Bu değer, modelin yanlış teşhislerden kaçındığını göstermektedir. Üçüncüsü Yanlış Pozitif (False Positive - FP) modelin, aslında sağlıklı bir bireyi yanlışlıkla hastalıklı olarak sınıflandırmasıdır. Sonuncusu ise Yanlış Negatif (False Negative-FN) modelin, gerçekten hastalıklı bir bireyin görüntüsünü sağlıklı olarak sınıflandırmasıdır.

Test aşamasında elde edilen sonuçlar, modelin gerçek dünyada ne kadar güvenilir bir şekilde kullanılabileceğini anlamak açısından oldukça önemlidir. Sonraki alt başlıklarda, test aşamasında hesaplanan tüm performans metrikleri ayrıntılı olarak ele alınmıştır.

4.3.1 Karışıklık Matrisi (Confusion Matrix)

Makine öğrenimi ve derin öğrenme modellerinin sınıflandırma başarısını değerlendirmek için karışıklık matrisi (confusion matrix) en yaygın kullanılan yöntemlerden biridir. Karışıklık matrisi, modelin gerçek etiketlerle tahmin edilen etiketler arasındaki ilişkiyi görselleştirerek, modelin hangi sınıflarda başarılı olduğunu ve hangi sınıflarda hata yaptığını detaylı bir şekilde incelemeye yardımcı olur.

Karışıklık matrisi, sınıflandırma görevlerinde modelin doğru ve yanlış tahminlerinin ayrıntılı bir tablosunu sunar (Çizelge 4.2). Matris dört temel bileşenden oluşur: Doğru Pozitifler (TP), Yanlış Pozitifler (FP), Doğru Negatifler (TN) ve Yanlış Negatifler (FN) (Heydarian vd. 2022). Bu matris, modelin yanlış pozitif ve yanlış negatif tahmin oranları hakkında daha net bir değerlendirme yapılmasını sağlar.

Çizelge 4.2 Karışıklık Matrisi

		Tahmin Edilen Sınıf	
		Pozitif (T)	Negatif (F)
Gerçek Sınıf	Pozitif (P)	TP	FN
	Negatif (N)	FP	TN

Çizelge 4.3'te, test verileri TTA yöntemi kullanılarak değerlendirilmiş ve ilgili karışıklık matrisi sunulmuştur.

Çizelge 4.3 Test edilen CNN modelinin karışıklık matrisi

		Tahmin Edilen Sınıf				
		Normal	Hafif	Orta	Şiddetli	Proliferatif
Gerçek Sınıf	Normal	151	0	1	0	0
	Hafif	2	21	4	0	0
	Orta	2	3	88	0	1
	Şiddetli	0	0	2	9	1
	Proliferatif	0	1	6	1	18

Çizelge 4.3 incelendiğinde model, 151 adet gerçek normal görüntünün tamamına yakını doğru şekilde sınıflandırmış ve sadece 1 adet normal görüntüyü yanlış tahmin ederek orta sınıfa atamıştır. Bu durum, modelin normal sınıfındaki yüksek doğruluk oranını göstermektedir.

Hafif diyabetik retinopati sınıfındaki 27 örnekten 21'i doğru sınıflandırılmıştır. Ancak 4 görüntü orta, 2 görüntü ise normal olarak sınıflandırılmıştır. Hafif sınıfında genel olarak başarılı bir performans gözlemlenmektedir

Model, 88 görüntüyü doğru şekilde orta sınıf olarak sınıflandırmıştır. Ancak 3 görüntü hafif, 2 görüntü normal, 1 görüntü ise proliferatif olarak yanlış sınıflandırılmıştır. Modelin orta sınıfta genel olarak yüksek bir başarıya sahip olduğu görülmektedir.

Model, 9 görüntüyü doğru bir şekilde şiddetli olarak sınıflandırmıştır. Ancak 2 görüntü orta, 1 görüntü proliferatif olarak sınıflandırılmıştır. Şiddetli sınıfın sınırlı sayıda örneğe sahip olmasına rağmen, model burada başarılı bir sınıflandırma gerçekleştirmiştir.

Model, 18 proliferatif görüntüyü doğru şekilde sınıflandırmıştır. Modelin proliferatif vakaları genellikle iyi tahmin ettiği, ancak bazı vakaları daha az şiddetli olarak sınıflandırdığı görülmektedir.

Bu matris, yapılan analizlerin yanı sıra modelin sınıflandırma performansını değerlendirmek için kullanılan hassasiyet, duyarlılık, özgüllük, doğruluk ve F1 puanı gibi ölçümleri hesaplamak için kullanılmaktadır (Tan ve Le 2019).

4.3.2 Hassasiyet (Precision)

Hassasiyet, gerçek pozitif (TP) tahminlerin toplam pozitif tahminlere (gerçek pozitifler ve yanlış pozitifler) oranı olarak tanımlanır. Bu metrik modelin yanlış pozitifleri (FP) en aza indirme yeteneğini göstererek pozitif vakaların model tarafından ne kadar doğru tahmin edildiğini yansıtır (Sokolova ve Lapalme 2009). Hassasiyet, özellikle yanlış pozitiflerin önlenmesinin önemli olduğu tıbbi teşhis gibi durumlarda oldukça önemli bir metriktir. Bu metriğe ait denklem 4.1’de verilmiştir.

$$Hassasiyet = \frac{TP}{TP + FP} \quad (4.1)$$

4.3.3 Duyarlılık (Recall)

Duyarlılık, bir sınıflandırma modelinin gerçek pozitif örnekleri doğru bir şekilde tespit etme yeteneğini ölçen temel metriklerden biridir. Tıbbi teşhis, sahtekarlık tespiti ve güvenlik sistemleri gibi önemli alanlarda, duyarlılığın yüksek olması gerekmektedir. Duyarlılık, gerçek pozitif tahminlerin toplam gerçek pozitiflere (gerçek pozitiflerin ve

yanlış negatiflerin toplamı) oranıdır. Bu ölçüm, modelin ciddi hastalık vakaları olan sınıfları ne kadar doğru bir şekilde tanımlayabildiğini gösterir (Goutte ve Gaussier 2005). Denklem 4.2' duyarlılık metriğinin formülü verilmiştir.

$$Duyarlılık = \frac{TP}{TP + FN} \quad (4.2)$$

4.3.4 Özgüllük (Specificity)

Özgüllük (specificity), bir sınıflandırma modelinin gerçek negatif örnekleri doğru bir şekilde tespit etme yeteneğini ölçen bir metriktir. Modelin, negatif sınıfa ait örnekleri ne kadar başarılı bir şekilde tanıyabildiğini gösterir Özgüllük, gerçek negatif tahminlerin toplam sayısının, gerçek negatiflerin ve yanlış pozitiflerin toplamına oranıdır. Bu ölçüm, modelin hedef sınıfa ait olmayan görüntüleri güvenilir bir şekilde ayırt etme yeteneğini gösterir (Trevethan 2017). Özgüllük denklemi, denklem 4.3'te verilmiştir.

$$Özgüllük = \frac{TN}{TN + FP} \quad (4.3)$$

4.3.5 Doğruluk (Accuracy)

Doğruluk (accuracy), bir sınıflandırma modelinin genel başarısını ölçen en temel metriktir. Toplam doğru sınıflandırma sayısının toplam tahmin sayısına oranı olarak tanımlanır. Başka bir deyişle, model tarafından yapılan tüm tahminler arasında doğru tahmin edilen sınıfların oranını gösterir. Matematiksel ifadesi Denklem 4.4'te verilmiştir.

$$Doğruluk = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.4)$$

4.3.6 F1-Skoru (F1-Score)

F1 Skoru, hassasiyet ve duyarlılık arasında bir denge sağlayan ve modelin performansını her iki metrik açısından yansıtan bir ölçümdür. Özellikle dengesiz veri setlerinde modelin genel performansının daha iyi analiz edilmesini sağlar (Fujino vd. 2008). Denklem 4.5'te

F1-Skoru metriğinin matematiksel ifadesi yer almaktadır.

$$F1\ Skoru = 2 \times \frac{Hassasiyet \times Duyarlilik}{Hassasiyet + Duyarlilik} \quad (4.5)$$

Bu metrikler, modelin test sürecinde, hiç görmediği görüntüler üzerindeki sınıflandırma performansını değerlendirmek için kullanılmıştır. Eğitilen modelin test görüntüleri üzerindeki sınıflandırma performansına ilişkin sonuçlar Çizelge 4.4'te sunulmuştur.

Çizelge 4.4 Test edilen CNN modelinin sınıflandırma performansı sonuçları

Sınıf	Hassasiyet (%)	Duyarlılık (%)	Özgüllük (%)	F1 Skoru (%)	Doğruluk (%)
Normal	97.42	99.34	97.48	98.37	98.39
Hafif	84.00	77.78	98.59	80.77	96.79
Orta	87.13	93.62	94.01	90.26	93.89
Şiddetli	90.00	75.00	99.67	81.82	98.71
Proliferatif	90.00	69.23	99.30	78.26	96.79
Genel	92.24	92.28	96.77	92.07	92.28

Sınıflandırma performansı sonuçları (Çizelge 4.4) incelendiğinde, modelin diyabetik retinopati aşamalarını sınıflandırmada yüksek doğruluk oranları ile başarılı bir performans sergilediği görülmektedir.

Modelin genel hassasiyet değeri %92.24 olarak hesaplanmıştır. Normal sınıf için hassasiyet %97.42 gibi oldukça yüksek bir değer alırken, tüm sınıflar için hassasiyet oranları tatmin edici seviyelerdedir. Bu sonuç, modelin sınıflar arasında net bir ayırım yapabildiğini ve yüksek doğruluk oranlarına sahip olduğunu göstermektedir. Modelin genel duyarlılık değeri %92.28 gibi yüksek bir değer olarak hesaplanmıştır. Genel özgüllük değerinin %96.77 olduğu görülmektedir. Bu yüksek özgüllük oranı, modelin yanlış pozitif tahmin yapma olasılığının düşük olduğunu ve büyük ölçüde doğru negatif tahminlerde bulunduğunu göstermektedir. Tüm sınıflarda özgüllük oranlarının yüksek

olmasından, modelin başarılı bir sınıflandırma performansına sahip olduğu anlaşılmaktadır. F1 skoru, hassasiyet ve duyarlılık arasındaki dengeyi sağlayan bir metriktir ve modelin genel performansını ölçmek açısından büyük bir öneme sahiptir. Modelin genel F1 skoru %92.07 olarak hesaplanmıştır. Bu yüksek F1 skoru, modelin sınıflandırma başarısının dengeli olduğunu göstermektedir. Doğruluk metriği, modelin tüm tahminlerdeki genel doğruluğunu ifade eder. Modelin genel doğruluk değeri %92.28 olarak hesaplanmıştır. Bu oran, modelin test verisi üzerindeki performansının oldukça başarılı olduğunu ve veri üzerinde güvenilir sonuçlar ürettiğini göstermektedir.

Şiddetli ve proliferatif sınıfların en düşük Duyarlılık ve F1 puanlarına sahip olduğu göze çarpmaktadır. Bu sonuç, bu iki kategori arasındaki yüksek benzerlikle ilişkilendirilmektedir. Bu da modelin bu sınıfları doğru bir şekilde ayırt etmesini daha zor hale getirmiştir. Elde edilen sonuçlar modelin genel olarak yüksek performans sergilediğini ve sınıflar arasında başarılı bir ayırım yapabildiğini göstermektedir. Hassasiyet, duyarlılık, özgüllük, F1 skoru ve doğruluk gibi temel metriklerin tamamında tutarlı ve yüksek sonuçlar elde edilmiştir.

4.3.7 ROC Eğrisi ve AUC Değeri

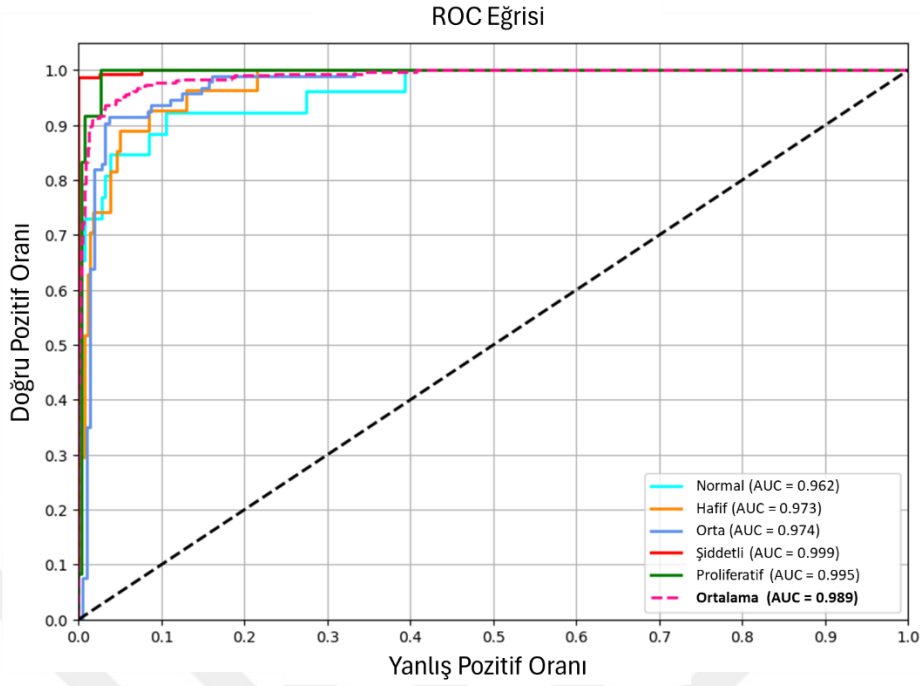
ROC (Receiver Operating Characteristic-Alicı İşletim Karakteristiği) eğrisi, sınıflandırma modellerinin performansını değerlendirmek için yaygın olarak kullanılan bir yöntemdir. ROC eğrisi, modelin gerçek pozitif oranını (TPR) ve yanlış pozitif oranını (FPR) grafiksel olarak temsil eder. Bu eğri, modelin bir sınıflandırma problemindeki ayırt edici yeteneğini görselleştirir.

ROC eğrisi, modelin sınıflandırma performansını belirlerken sınıfların nasıl ayrıştırıldığını ölçer. Eğrinin sol üst köşeye yakın olması, modelin daha başarılı bir sınıflandırma yaptığını gösterir. Eğer ROC eğrisi, 45 derecelik doğrusal bir çizgi şeklinde ilerliyorsa, modelin tahminlerinin rastgele olduğu anlaşılır ve bu durumda modelin ayırt edici gücü olmadığı kabul edilir. Bu nedenle, ROC eğrisinin eğimi ne kadar fazla ise, modelin pozitif ve negatif sınıfları ayırt etme başarısı da o kadar yüksektir (Fawcett 2006).

ROC eğrisi ile birlikte değerlendirilen bir diğer önemli metrik ise AUC (Area Under the Curve-Eğri Altındaki Alan) değeridir. AUC metriği, ROC eğrisinin altında kalan alanı temsil eder ve modelin genel sınıflandırma performansını tek bir değerle ifade eder. AUC değeri 0.5 ile 1 arasında değişmektedir ve modelin ayırt edicilik yeteneğini ölçmek için kullanılır. AUC değeri 0.5 ise modelin tahminleri tamamen rastgeledir ve sınıfları ayırt etme yeteneği yoktur. 0.5 ile 0.6 arasındaysa modelin sınıflandırma becerisi zayıftır. 0.6 ile 0.8 arasındaysa model ortalama bir sınıflandırma performansına sahiptir. 0.8 ile 0.9 arasında olan AUC değeri modelin iyi bir sınıflandırma yeteneği olduğunu göstermektedir. 0.9 ile 1 arasında olan AUC değeri ise modelin sınıflandırma performansının yüksek olduğunu ortaya koymaktadır.

AUC metriği, özellikle veri setlerinde sınıf dengesizliği bulunduğunda önemli bir değerlendirme ölçütüdür. Doğruluk (accuracy) gibi metrikler, dengesiz veri kümelerinde yanıltıcı sonuçlar verebilirken, AUC değeri modelin genel performansını daha sağlıklı bir şekilde ölçebilmektedir. Özellikle sağlık alanındaki tahmin modellerinde, yanlış negatif ve yanlış pozitif hataların kritik sonuçlar doğurabileceği göz önüne alındığında, ROC-AUC metriğinin yüksek olması, modelin güvenilirliğini artıran önemli bir faktör olarak değerlendirilmektedir.

Model performansı, her sınıf için ROC eğrileri ve karşılık gelen AUC (değerleri kullanılarak değerlendirilmiştir. (Şekil 4.3).



Şekil 4.3 ROC Eğrisi

Grafikte, her sınıf için ayrı ROC eğrileri çizilmiş ve bu eğrilerin altında kalan alanı temsil eden AUC değerleri hesaplanmıştır. Modelin her bir sınıf için hesaplanan AUC değerleri şu şekildedir:

- Normal sınıfı için AUC değeri = 0.962
- Hafif sınıfı için AUC değeri = 0.973
- Orta sınıfı için AUC değeri = 0.974
- Şiddetli sınıfı için AUC değeri = 0.999
- Proliferatif sınıfı için AUC değeri = 0.995
- Genel AUC değeri = 0.989

Bu değerler incelendiğinde, tüm sınıflar için AUC değerlerinin 0.96'nın üzerinde olduğu ve modelin yüksek sınıflandırma başarısına sahip olduğu görülmektedir. ROC eğrisinin sol üst köşeye yakın olması, modelin yanlış pozitif oranını düşük tutarak yüksek doğru pozitif oranına sahip olduğunu göstermektedir. Eğrinin diyagonal çizgiden (sınıflandırma çizgisi) ne kadar uzak olduğu, modelin ayırt etme gücünü göstermektedir. Tüm sınıflar için ROC eğrilerinin diyagonal çizgiden belirgin şekilde uzaklaştığı ve sol üst köşeye yakın olduğu görülmektedir. Bu durum, modelin tüm sınıfları başarıyla ayırt ettiğini ve genel sınıflandırma performansının yüksek olduğunu kanıtlamaktadır.

4.4 İşlenmiş ve İşlenmemiş Görüntüler Üzerinde Model Performans Analizi

Önerilen model, ön işleme adımlarının önemini değerlendirmek için ham görüntülerle yeniden eğitilmiştir. Herhangi bir veri büyütme veya görüntü işleme aşaması uygulanmayan işlenmemiş görüntüler, Çizelge 4.5'te gösterildiği gibi iki aşamalı eğitim sürecinin ardından farklı metrikler bakımından karşılaştırılmıştır.

Çizelge 4.5 İşlenmiş ve işlenmemiş görüntülerin test sonuçları

	Hassasiyet	Duyarlılık	Özgüllük	F1 Puanı	Doğruluk
	(%)	(%)	(%)	(%)	(%)
İşlenmemiş Görüntüler	84.92	85.19	94.59	83.41	85.26
İşlenmiş Görüntüler	92.24	92.28	96.77	92.07	92.28

İşlenmemiş görüntülerle eğitilen modelin hassasiyet değeri %84.92, duyarlılık değeri %85.19, özgüllük değeri %94.59, F1 skoru %83.41 ve doğruluk değeri %85.26 olarak hesaplanmıştır. Buna karşılık, işlenmiş görüntüler kullanılarak eğitilen modelin hassasiyet değeri %92.24, duyarlılık değeri %92.28, özgüllük değeri %96.77, F1 puanı %92.07 ve doğruluk değeri %92.28 olarak hesaplanmıştır. Bu değerler, işlenmiş görüntüler üzerinde eğitilen modelin belirgin şekilde daha başarılı bir sınıflandırma gerçekleştirdiğini göstermektedir. Bu da görüntü işleme adımlarının eğitim sürecine ne kadar önemli bir katkı sağladığını açıkça göstermektedir.

4.5 Literatürdeki Yöntemlerle Modelin Karşılaştırmalı Performans Analizi

Bu tez çalışmasında önerilen modelin, literatürdeki farklı derin öğrenme tabanlı yöntemlerle karşılaştırmalı analizi yapılmıştır. Çizelge 4.6'da, APTOS veri seti üzerinde test edilen farklı modellerin hassasiyet, duyarlılık, özgüllük, F1 skoru ve doğruluk gibi temel performans metriklerine göre kıyaslanması verilmiştir.

Çizelge 4.6 APTOS veri setinde kullanılan farklı yöntemlerin karşılaştırmalı sonuçları

Referans	Model	Hassasiyet (%)	Duyarlılık (%)	Özgüllük (%)	F1-Skoru (%)	Doğruluk (%)
Al-Antary ve Arafa (2021)	MSA-Net	-	91.00	90.50	-	84.60
Menaouer vd. (2022)	Hibrit Model	94.66	95.00	-	94.00	90.60
Mondal vd. (2023)	EDLDR	76.00	82.00	-	-	86.08
Kassani vd. (2019)	Xception	-	88.24	87.00	-	83.09
Kallel ve Ehtioui (2024)	InceptionV3	92.00	85.40	-	86.60	96.88
Önerilen Model	EfficientNetV2-S	92.24	92.28	96.77	92.07	92.28

Çizelge 4.6'daki verilere göre, bu tez çalışmasında önerilen yöntem hassasiyet (%92,24), duyarlılık (%92,28), özgüllük (%96,77), F1-Skoru (%92,07) ve doğruluk (92,28) gibi temel metriklerde rekabetçi ve dengeli bir performans sergilemektedir. Literatürdeki çalışmalar değerlendirildiğinde, genellikle belirli bir metrik yüksek değer alırken, diğer metriklerde belirgin düşüşler yaşandığı gözlemlenmektedir. Bu durum, modellerin genel performans açısından tutarsızlık gösterdiğine ve belirli sınıflar üzerinde aşırı uyum sağlayarak genelleştirme yeteneklerinin sınırlı olabileceğine işaret etmektedir.

Örneğin, Menaouer vd. (2022) yüksek hassasiyet ve duyarlılık sağlarken, bu tez çalışmasında önerilen model genel olarak dengeli bir performans sağlayarak F1 Skoru ve doğruluk alanlarında üstünlük göstermektedir. Benzer şekilde Kallel ve Ehtioui yaptıkları çalışmada %96,88 oranında doğruluk elde etseler de diğer metrikler bakımından geride kaldıkları görülmektedir.

Özgüllük açısından yapılan değerlendirmede, EfficientNetV2-S modeli %96,77 ile en yüksek değeri elde etmiştir. Literatürdeki diğer çalışmalar dikkate alındığında, bu tez çalışmasında önerilen model tüm temel metriklerde üstün performans sergilemiş olup, özellikle özgüllük açısından %96,77 değeri ile en yüksek başarı oranını elde etmiştir. Bu sonuç, yanlış pozitiflik oranlarını başarılı bir şekilde en aza indirdiğini göstermektedir.

Yanlış pozitif oranlarının düşük olması, özellikle tıbbi teşhis sistemlerinde önemli bir faktör olup, önerilen modelin klinik uygulanabilirliğini artırmaktadır.

Sonuç olarak, önerilen model diyabetik retinopati aşamalarının sınıflandırılmasında, yalnızca belirli bir metrikte değil, tüm temel performans göstergelerinde tutarlıdır ve yüksek başarı elde etmiştir.



5. TARTIŞMA ve SONUÇ

Bu tez çalışmasında, APTOS 2019 veri setinden retina görüntüleri kullanılarak diyabetik retinopatiyi normal, hafif, orta, şiddetli ve proliferatif olmak üzere beş farklı sınıfta tespit etmek için derin öğrenme tabanlı bir yöntem sunularak diyabetik retinopatinin erken evrelerde tespit edilerek görme kaybının önüne geçilmesi ve literatüre önemli bir katkı sağlanması amaçlanmıştır. Bu transfer öğrenme tabanlı yaklaşımda, sekiz farklı derin öğrenme modeli yeniden tasarlanmış ve eğitilmiştir. Eğitim sonuçlarına göre, en iyi performansı gösteren EfficientNetV2-S modeline ince ayar yapılmış ve yeniden eğitilmiştir.

Veri setinde bulunan sınıf dağılımındaki dengesizlik, nadir görülen şiddetli ve proliferatif aşamalarda modelin performansını doğrudan etkileyebilecek bir faktördür. Bu sorunu en aza indirmek için, veri artırma (döndürme, ölçeklendirme, parlaklık artıma-azaltma, aynalama), sınıf ağırlıkları ve kayıp fonksiyonuna dayalı çeşitli teknikler uygulanmıştır. Bu sayede literatürde sıklıkla karşılaşılan veri yetersizliği sorunu başarılı bir şekilde yönetilmiş ve model az sayıda veri bulunan sınıflarda bile yüksek bir sınıflandırma performansı göstermiştir. Ek olarak, veri setinde yer alan görüntülerin kalitesindeki farklılıklar önemli zorluklardan biridir. Farklı cihazlardan elde edilen görüntülerden ve çekim koşullarındaki değişikliklerden kaynaklanan bu kalite farklılıkları, geliştirilen görüntü işleme teknikleri ile ortadan kaldırılmış ve görüntülerin okunabilirliği artırılmıştır.

75 iterasyon olarak planlanan eğitim sürecinde, aşırı öğrenmeyi önlemek için EarlyStopping ve ReduceLROnPlateau gibi fonksiyonlar kullanılarak, model için daha dengeli bir öğrenme süreci sağlanmıştır. Kullanılan hesaplama kaynakları bakımından EfficientNetV2-S modeli, diğer büyük ölçekli derin öğrenme modellerine kıyasla daha verimli çalışmış ve düşük hesaplama maliyeti ile geniş ölçüde uygulanabilirliğini artırmıştır.

Bu tez çalışmasında transfer öğrenme tabanlı geliştirilen hibrit model daha önce modelin hiç görmediği ve herhangi bir ön işlem uygulanmayan test verileri ile test edilmiştir. Test

verileri, modelin daha genelleştirilmiş ve sağlam tahminler yapmasını sağlamak için Test Zamanı Artırma (TTA) yöntemi kullanılarak test edilmiştir (Ioffe ve Szegedy 2015). TTA yönteminin, veri setinde yer alan hafif ve şiddetli sınıfların başarı oranlarını önemli ölçüde iyileştirdiği gözlenmiştir. Test sonuçlarına göre çalışmada geliştirilen model %92.24 hassasiyet, %92.28 duyarlılık, %96.77 özgüllük, %92.07 F1-skoru ve %92.28 doğruluk elde etmiştir. Temek metriklerdeki bu başarı oranları modelin önceki çalışmalara göre daha tutarlı ve daha yüksek bir sınıflandırma performansına sahip olduğunu göstermektedir. Literatürde yer alan diğer çalışmalarla karşılaştırıldığında, geliştirilen model özellikle özgüllük açısından en yüksek performansı göstermiştir. Geliştirilen model %98.99 AUC değeri elde etmiştir. Genel ortalama AUC değerinin 0.989 gibi oldukça yüksek bir seviyede olması, modelin tüm sınıflar için çok yüksek bir ayırım gücüne sahip olduğunu ve diyabetik retinopati evrelerini hassas bir şekilde sınıflandırdığını kanıtlamaktadır. AUC değeri 1'e yaklaştıkça modelin daha iyi bir performans gösterdiği dikkate alındığında, bu sonuç modelin diyabetik retinopati teşhisinde başarılı bir şekilde kullanılabileceğini ortaya koymaktadır.

Bu tez çalışması, diyabetik retinopati tanısında hem model başarısı hem de metodolojik yenilikler açısından önemli bir adım olup, literatürdeki güncel yaklaşımlara göre daha sağlam, tutarlı ve genelleştirilebilir bir çözüm sunmaktadır. Elde edilen %92.24 hassasiyet, %92.28 duyarlılık, %96.77 özgüllük, %92.07 F1-skoru, %92.28 doğruluk ve %98.99 AUC değeri, geliştirilen modelin klinik uygulamalarda potansiyel olarak kullanılabileceğini göstermektedir.

6. KAYNAKLAR

- Al-Antary M T, Arafa Y, 2021, Multi-Scale Attention Network for Diabetic Retinopathy Classification, IEEE Access, 9, 54190-54200.
- Asri H, Mousannif H, Moatassime H A, Noel T, 2016, Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis, Procedia Computer Science, 83, 1064-1069.
- Association A D, 2004, Gestational diabetes mellitus, Diabetes care, 27, 88-90.
- Bayram F, Eleyan A, 2022, COVID-19 detection on chest radiographs using feature fusion based deep learning, Signal, Image and Video Processing, 16, 6, 1455-1462.
- Contreras-Luján E E, García-Guerrero E E, López-Bonilla O R, Tlelo-Cuautle E, López-Mancilla D, Inzunza-González E, 2022, Evaluation of Machine Learning Algorithms for Early Diagnosis of Deep Venous Thrombosis, Mathematical and Computational Applications, 27, 2, 24.
- Dao Q T, Trinh H Q, Nguyen V A, 2023, An effective and comprehensible method to detect and evaluate retinal damage due to diabetes complications, PeerJ Computer Science, 9, e1585.
- Emmert-Streib F, Yang Z, Feng H, Tripathi S, Dehmer M, 2020, An Introductory Review of Deep Learning for Prediction Models With Big Data, Frontiers in Artificial Intelligence, 3,
- Fawcett T, 2006, An introduction to ROC analysis, Pattern Recognition Letters, 27, 8, 861-874.
- Fong D S, Aiello L P, Ferris III F L, Klein R, 2004, Diabetic Retinopathy, Diabetes care, 27, 10, 2540-2553.
- Fujino A, Isozaki H, Suzuki J, 2008, Multi-label text categorization with model combination based on f1-score maximization, Proceedings of the Third International Joint Conference on Natural Language Processing: Volume-II,
- Gedraite E S, Hadad M, 2011, Investigation on the effect of a Gaussian Blur in image

- filtering and segmentation, Proceedings ELMAR-2011, 14-16 Sept. 2011, 393-396.
- Goutte C, Gaussier E, 2005, A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation, Advances in Information Retrieval, 2005//, Berlin, Heidelberg, 345-359.
- He H, Garcia E A, 2009, Learning from Imbalanced Data, IEEE Transactions on Knowledge and Data Engineering, 21, 9, 1263-1284.
- Heaton J, 2018, Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning, Genetic Programming and Evolvable Machines, 19, 1, 305-307.
- Heydarian M, Doyle T E, Samavi R, 2022, MLCM: Multi-Label Confusion Matrix, IEEE Access, 10, 19083-19095.
- Houssein E H, Emam M M, Ali A A, 2022, An optimized deep learning architecture for breast cancer diagnosis based on improved marine predators algorithm, Neural Computing and Applications, 34, 20, 18015-18033.
- Howard A G, 2013, Some Improvements on Deep Convolutional Neural Network Based Image Classification,
- Ioffe S, Szegedy C, 2015, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, Proceedings of the 32nd International Conference on Machine Learning, Proceedings of Machine Learning Research, 448--456.
- Javid A M, Das S, Skoglund M, Chatterjee S, 2021, A ReLU Dense Layer to Improve the Performance of Neural Networks, ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 6-11 June 2021, 2810-2814.
- Jia Y, Chen G, Chi H, 2024, Retinal fundus image super-resolution based on generative adversarial network guided with vascular structure prior, Scientific Reports, 14, 1, 22786.
- Jojoa Acosta M F, Caballero Tovar L Y, Garcia-Zapirain M B, Percybrooks W S, 2021, Melanoma diagnosis using deep learning techniques on dermoscopic images,

- BMC Medical Imaging, 21, 1, 6.
- Kallel F, Echtioui A, 2024, Retinal fundus image classification for diabetic retinopathy using transfer learning technique, *Signal, Image and Video Processing*, 18, 2, 1143-1153.
- Karapinar Senturk Z, 2020, Early diagnosis of Parkinson's disease using machine learning algorithms, *Medical Hypotheses*, 138, 109603.
- Kassani S H, Kassani P H, Khazaeinezhad R, Wesolowski M J, Schneider K A, Deters R, 2019, Diabetic Retinopathy Classification Using a Modified Xception Architecture, 2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 10-12 Dec. 2019, 1-6.
- Khan A Q, Sun G, Khalid M, Farrash M, Bilal A, 2024, Multi-Deep Learning Approach With Transfer Learning for 7-Stages Diabetic Retinopathy Classification, *International Journal of Imaging Systems and Technology*, 34, 6, e23213.
- Kingma D P, Ba J L, 2014, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980,
- Krešo I, Oršić M, Bevandić P, Šegvić S, 2018, Robust semantic segmentation with ladder-densenet models, arXiv preprint arXiv:1806.03465,
- Lecun Y, Bottou L, Bengio Y, Haffner P, 1998, Gradient-based learning applied to document recognition, *Proceedings of the IEEE*, 86, 11, 2278-2324.
- Li X, Ji J, Li J, He S, Zhou Q, 2021, Research on Image Denoising Based on Median Filter, 2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), 18-20 June 2021, 528-531.
- Li Y, Deng L, Yang X, Liu Z, Zhao X, Huang F vd., 2019, Early diagnosis of gastric cancer based on deep learning combined with the spectral-spatial classification method, *Biomedical Optics Express*, 10, 10, 4999-5014.
- Liu S, Liu S, Cai W, Pujol S, Kikinis R, Feng D, 2014, Early diagnosis of Alzheimer's disease with deep learning, 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), 29 April-2 May 2014, 1015-1018.

- Liu Y, Jain A, Eng C, Way D H, Lee K, Bui P vd., 2020, A deep learning system for differential diagnosis of skin diseases, *Nature Medicine*, 26, 6, 900-908.
- Liu Z, Mao H, Wu C-Y, Feichtenhofer C, Darrell T, Xie S, 2022, A convnet for the 2020s, *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 11976-11986.
- Manalı D, Demirel H, Eleyan A, 2024, Deep Learning Based Breast Cancer Detection Using Decision Fusion, *Computers*, 13, 11, 294.
- Mascarenhas S, Agarwal M, 2021, A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for Image Classification, 2021 International Conference on Disruptive Technologies for Multi-Disciplinary Research and Applications (CENTCON), 19-21 Nov. 2021, 96-99.
- Menaouer B, Dermane Z, El Houda Kebir N, Matta N, 2022, Diabetic Retinopathy Classification Using Hybrid Deep Learning Approach, *SN Computer Science*, 3, 5, 357.
- Mondal S S, Mandal N, Singh K K, Singh A, Izonin I, 2023, EDLDR: An Ensemble Deep Learning Technique for Detection and Classification of Diabetic Retinopathy, *Diagnostics*, 13, 1, 124.
- Naji M A, Filali S E, Aarika K, Benlahmar E L H, Abdelouhahid R A, Debauche O, 2021, Machine Learning Algorithms For Breast Cancer Prediction And Diagnosis, *Procedia Computer Science*, 191, 487-492.
- Nawaz I M, Rezzola S, Cancarini A, Russo A, Costagliola C, Semeraro F vd., 2019, Human vitreous in proliferative diabetic retinopathy: Characterization and translational implications, *Progress in Retinal and Eye Research*, 72, 100756.
- Park S, Kwak N, 2017, Analysis on the Dropout Effect in Convolutional Neural Networks, *Computer Vision – ACCV 2016*, 2017, Cham, 189-204.
- Raschka S, Mirjalili V, 2019, *Python machine learning: Machine learning and deep learning with Python, scikit-learn, and TensorFlow 2*, Packt publishing ltd,
- Reddy G T, Bhattacharya S, Ramakrishnan S S, Chowdhary C L, Hakak S, Kaluri R vd., 2020, An Ensemble based Machine Learning model for Diabetic Retinopathy

- Classification, 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), 24-25 Feb. 2020, 1-6.
- Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L-C, 2018, Mobilenetv2: Inverted residuals and linear bottlenecks, Proceedings of the IEEE conference on computer vision and pattern recognition, 4510-4520.
- Sharma A, Shinde S, Shaikh I I, Vyas M, Rani S, 2021, Machine Learning Approach for Detection of Diabetic Retinopathy with Improved Pre-Processing, 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), 19-20 Feb. 2021, 517-522.
- Sharma R, Kamra A, 2023, A Review on CLAHE Based Enhancement Techniques, 2023 6th International Conference on Contemporary Computing and Informatics (IC3I), 14-16 Sept. 2023, 321-325.
- Sharma S, Sharma S, Athaiya A, 2017, Activation functions in neural networks, Towards Data Sci, 6, 12, 310-316.
- Shorten C, Khoshgoftaar T M, 2019, A survey on Image Data Augmentation for Deep Learning, Journal of Big Data, 6, 1, 60.
- So A, Hooshyar D, Park K W, Lim H S, 2017, Early Diagnosis of Dementia from Clinical Data by Machine Learning Techniques, Applied Sciences, 7, 7, 651.
- Sokolova M, Lapalme G, 2009, A systematic analysis of performance measures for classification tasks, Information Processing & Management, 45, 4, 427-437.
- Sumathy B, Chakrabarty A, Gupta S, Hishan S S, Raj B, Gulati K vd., 2022, Prediction of diabetic retinopathy using health records with machine learning classifiers and data science, International Journal of Reliable and Quality E-Healthcare (IJRQEH), 11, 2, 1-16.
- Tan M, Le Q, 2019, EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, Proceedings of the 36th International Conference on Machine Learning, Proceedings of Machine Learning Research, 6105--6114.
- Tan M, Le Q, 2021, EfficientNetV2: Smaller Models and Faster Training, Proceedings of the 38th International Conference on Machine Learning, Proceedings of

Machine Learning Research, 10096--10106.

Tang J, Kern T S, 2011, Inflammation in diabetic retinopathy, Progress in Retinal and Eye Research, 30, 5, 343-358.

Trevethan R, 2017, Sensitivity, specificity, and predictive values: foundations, pliabilities, and pitfalls in research and practice, Frontiers in public health, 5, 307.

Wang L, 2022, Deep Learning Techniques to Diagnose Lung Cancer, Cancers, 14, 22, 5569.

Weiss K, Khoshgoftaar T M, Wang D, 2016, A survey of transfer learning, Journal of Big Data, 3, 1, 9.

Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H vd., 2021, A Comprehensive Survey on Transfer Learning, Proceedings of the IEEE, 109, 1, 43-76.

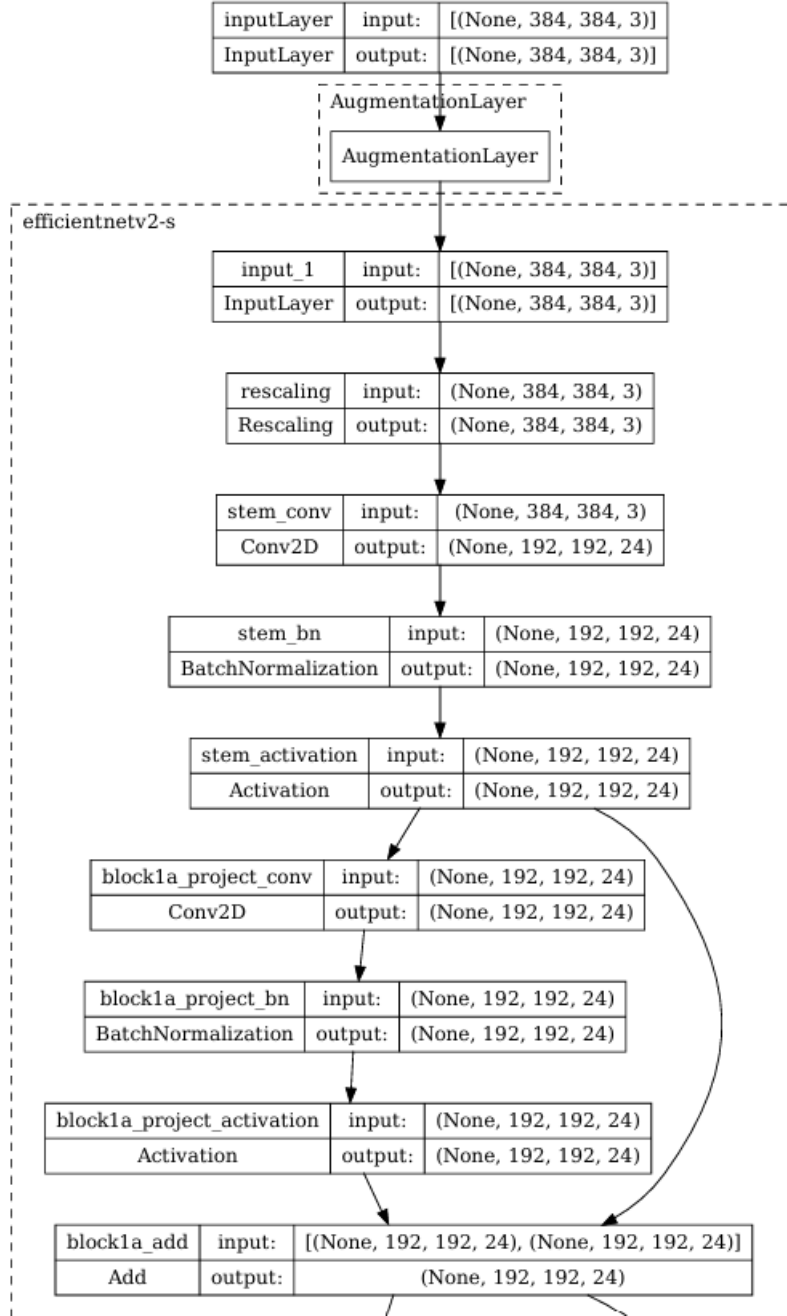
Zoph B, Vasudevan V, Shlens J, Le Q V, 2018, Learning transferable architectures for scalable image recognition, Proceedings of the IEEE conference on computer vision and pattern recognition, 8697-8710.

İnternet Kaynakları

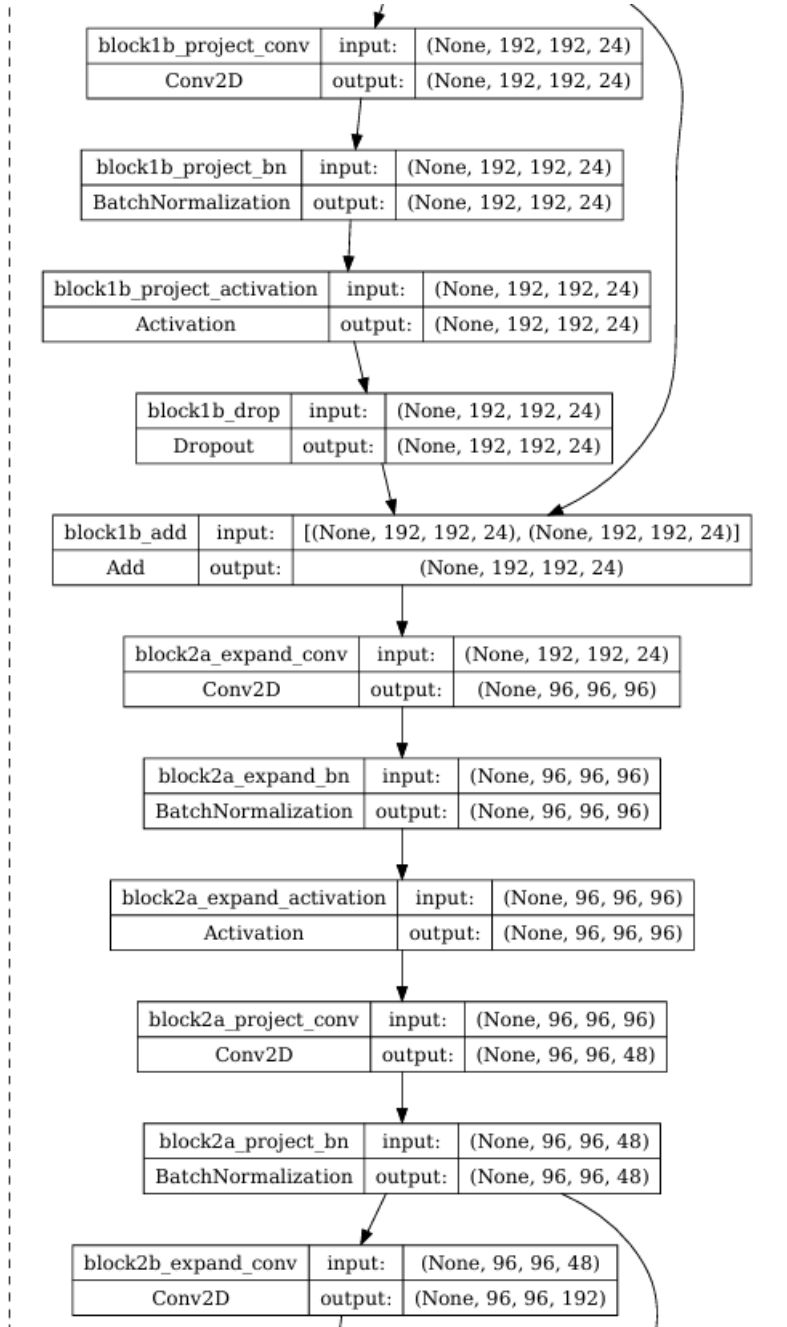
- 1) <https://kaggle.com/competitions/aptos2019-blindness-detection>, 28.06.2019
- 2) <https://zilliz.com/glossary/convolutional-neural-network>, 15.03.2024
- 3) <https://gelecegiyazanlar.turkcell.com.tr/konu/egitim/derin-ogrenme-cnn/evrisim-katmani>, 06.09.2022
- 4) <https://medium.com/deep-learning-from-deepest/katmanlar%C4%B1-anlamak-e6b33371f0f8>, 26.08.2021

EKLER

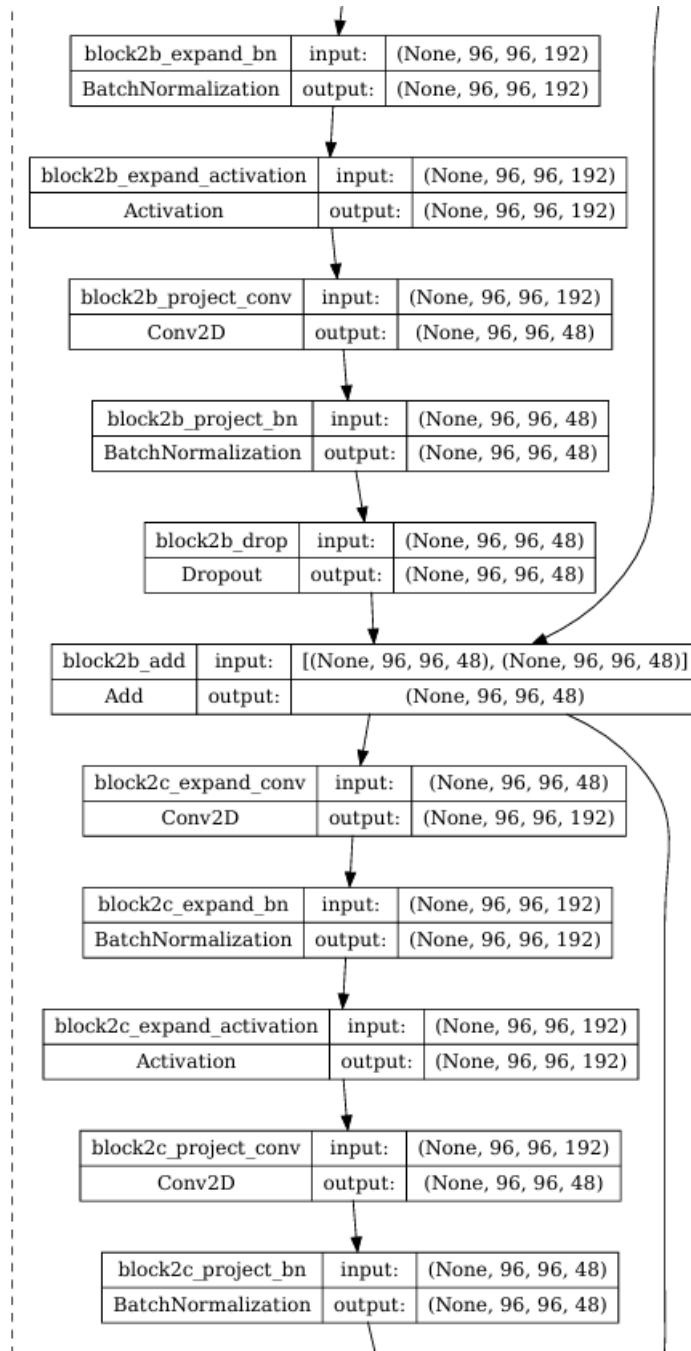
EK 1. Geliştirilen Modelin Mimarisi



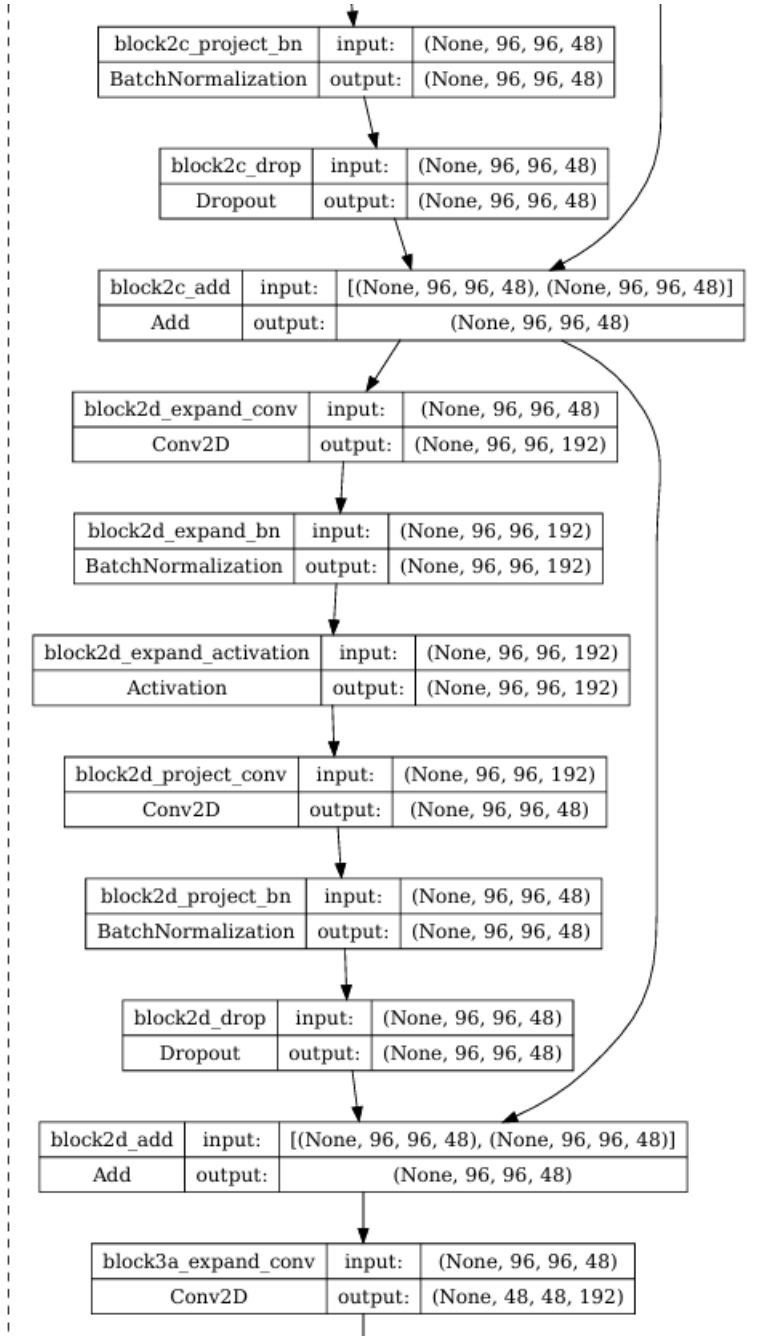
EK 1. (Devam) Geliştirilen Modelin Mimarisi



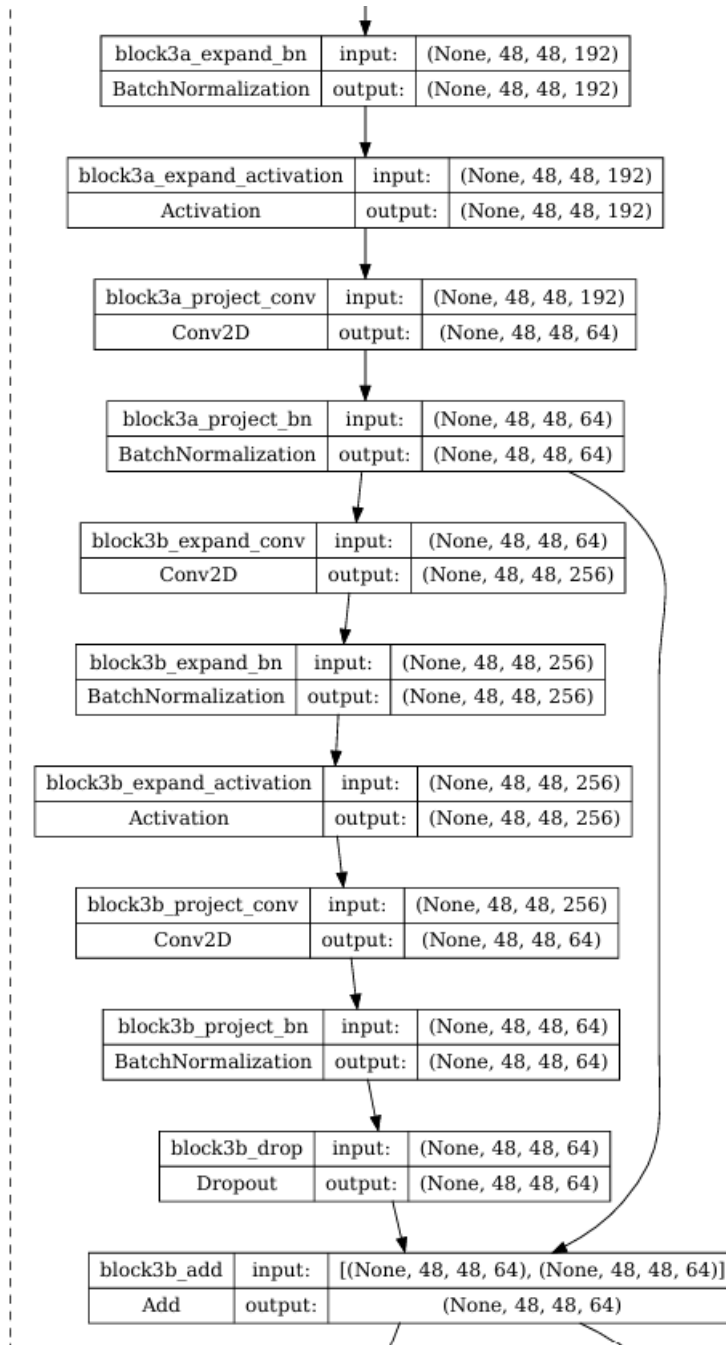
EK 1. (Devam) Geliştirilen Modelin Mimarisi



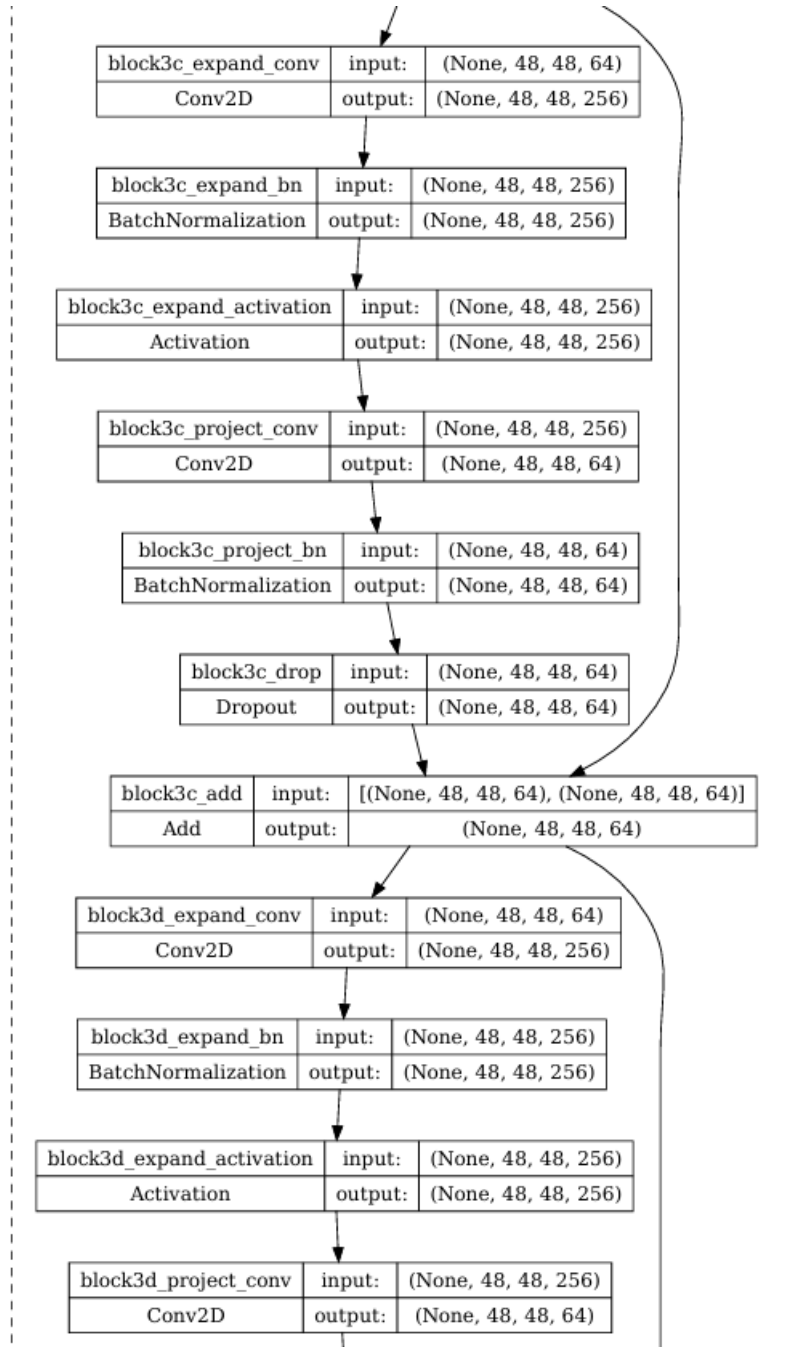
EK 1. (Devam) Geliştirilen Modelin Mimarisi



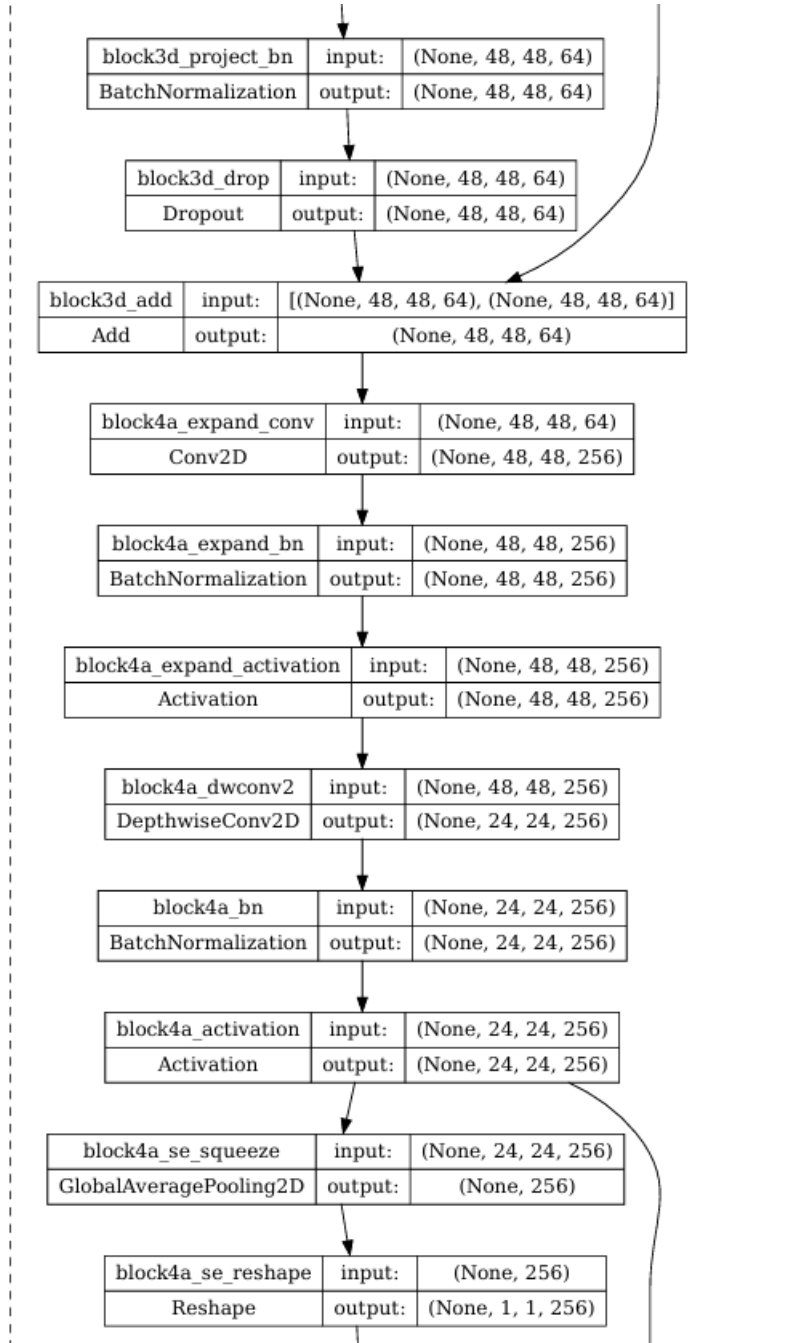
EK 1. (Devam) Geliştirilen Modelin Mimarisi



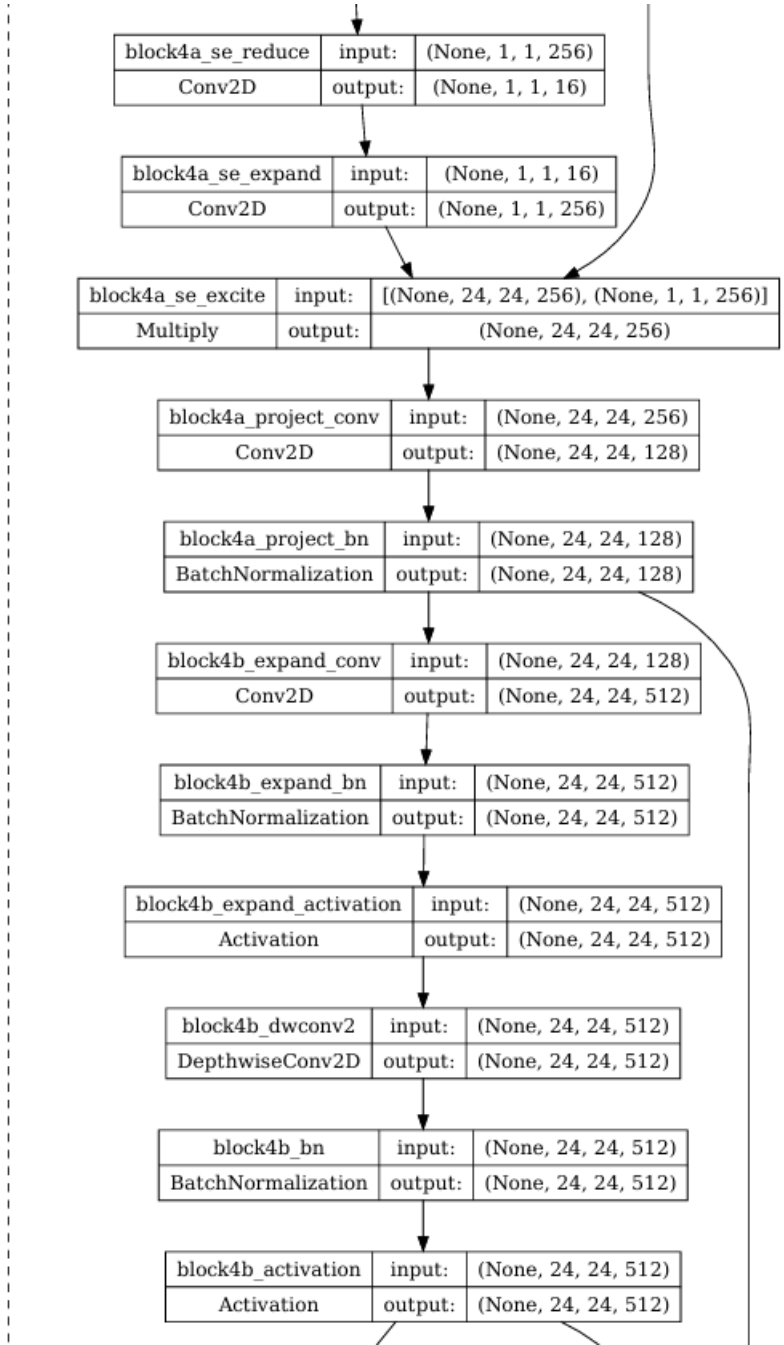
EK 1. (Devam) Geliştirilen Modelin Mimarisi



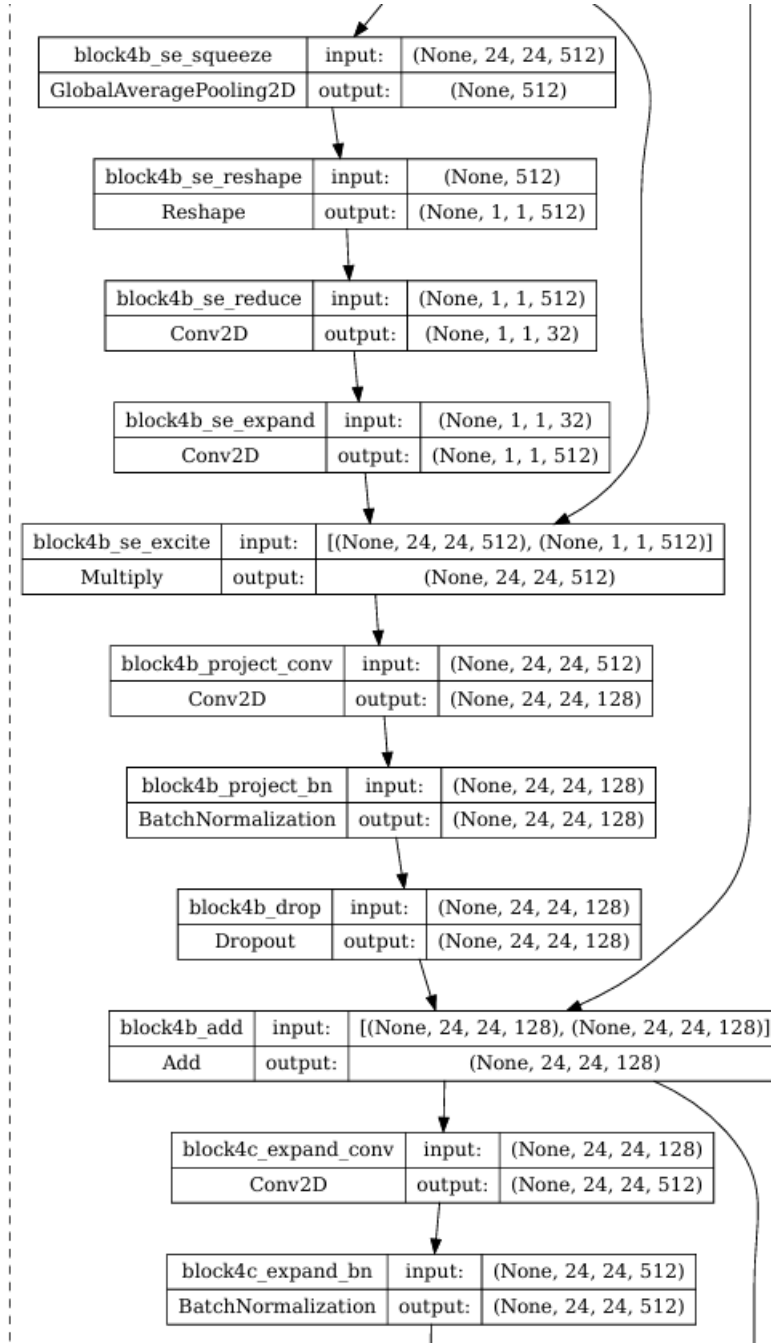
EK 1. (Devam) Geliştirilen Modelin Mimarisi



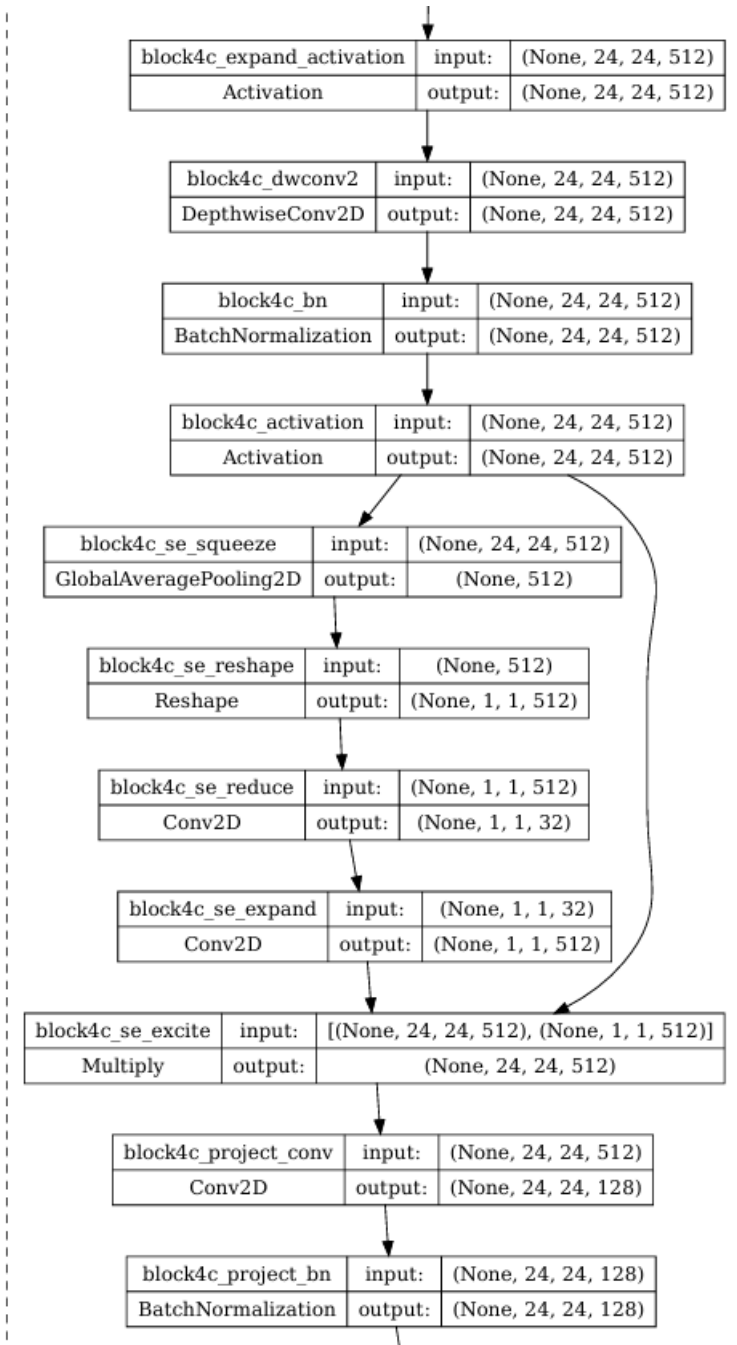
EK 1. (Devam) Geliştirilen Modelin Mimarisi



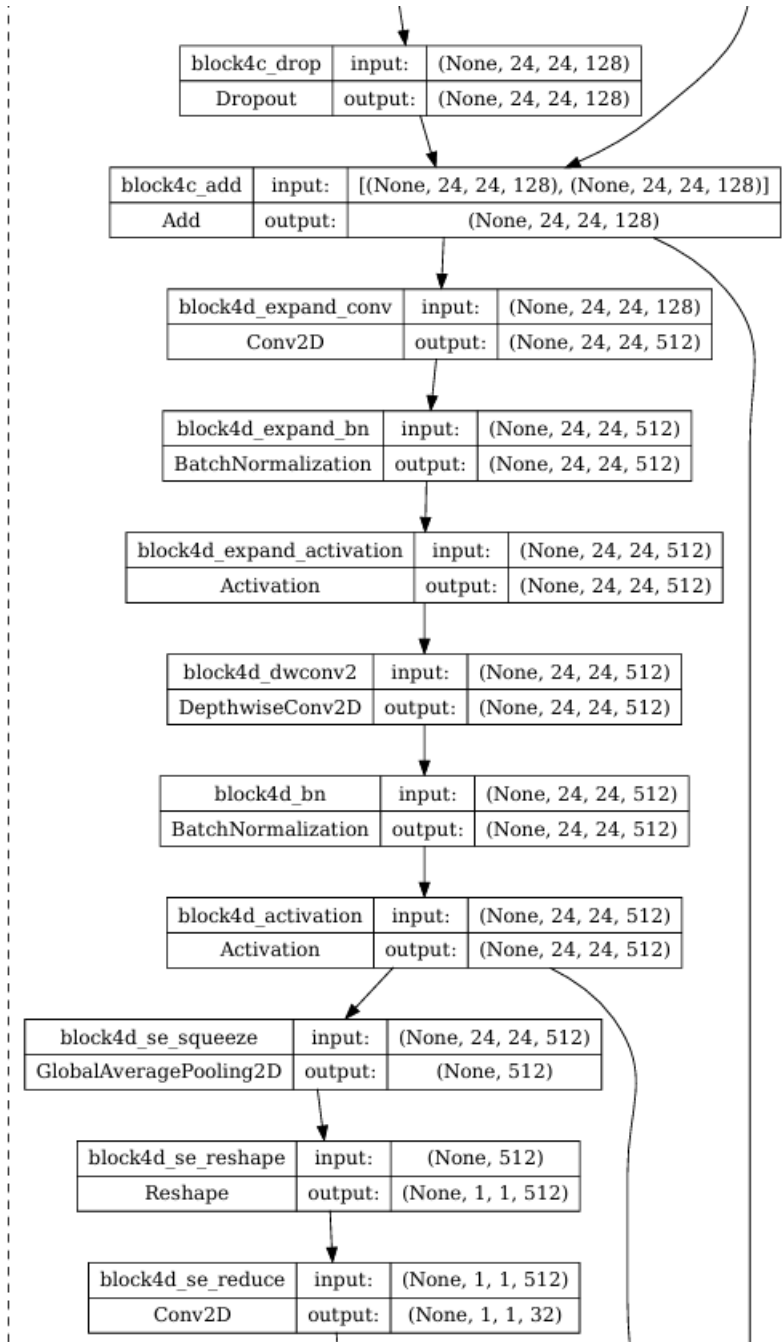
EK 1. (Devam) Geliştirilen Modelin Mimarisi



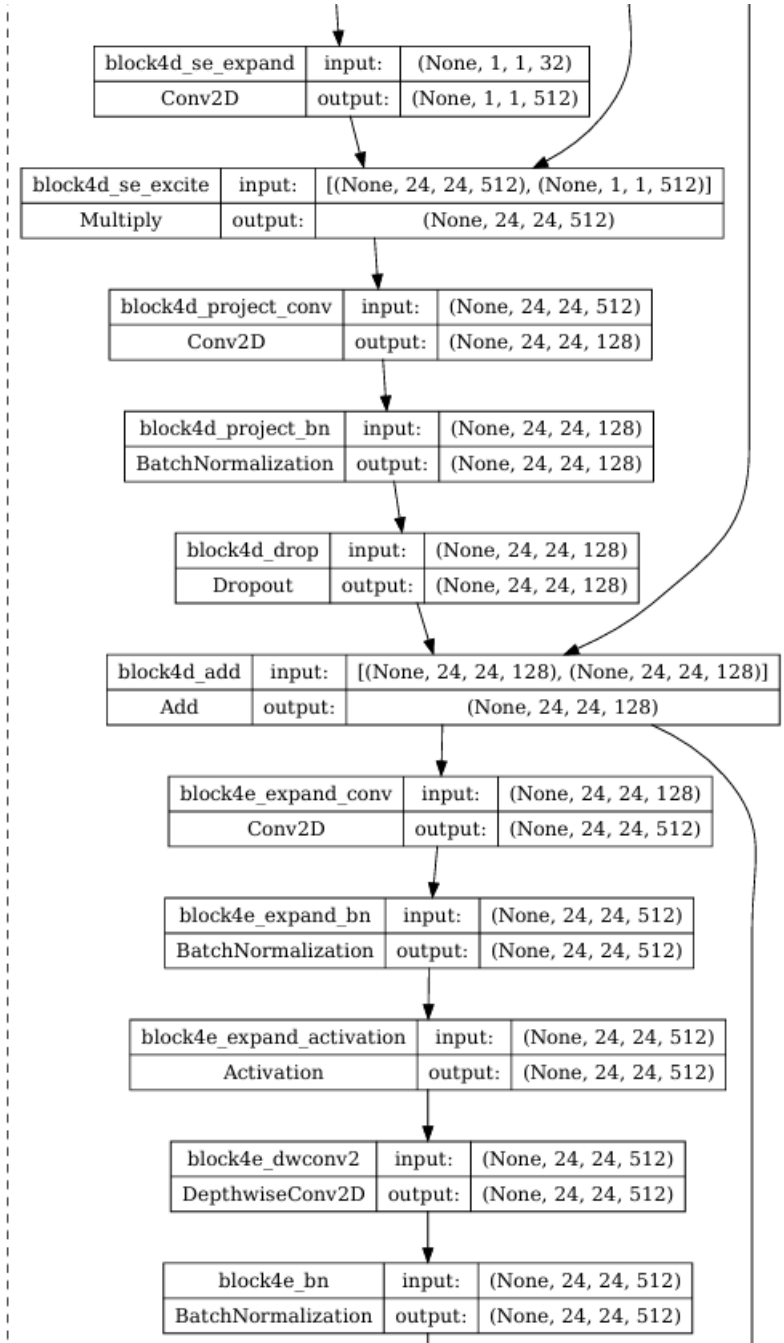
EK 1. (Devam) Geliştirilen Modelin Mimarisi



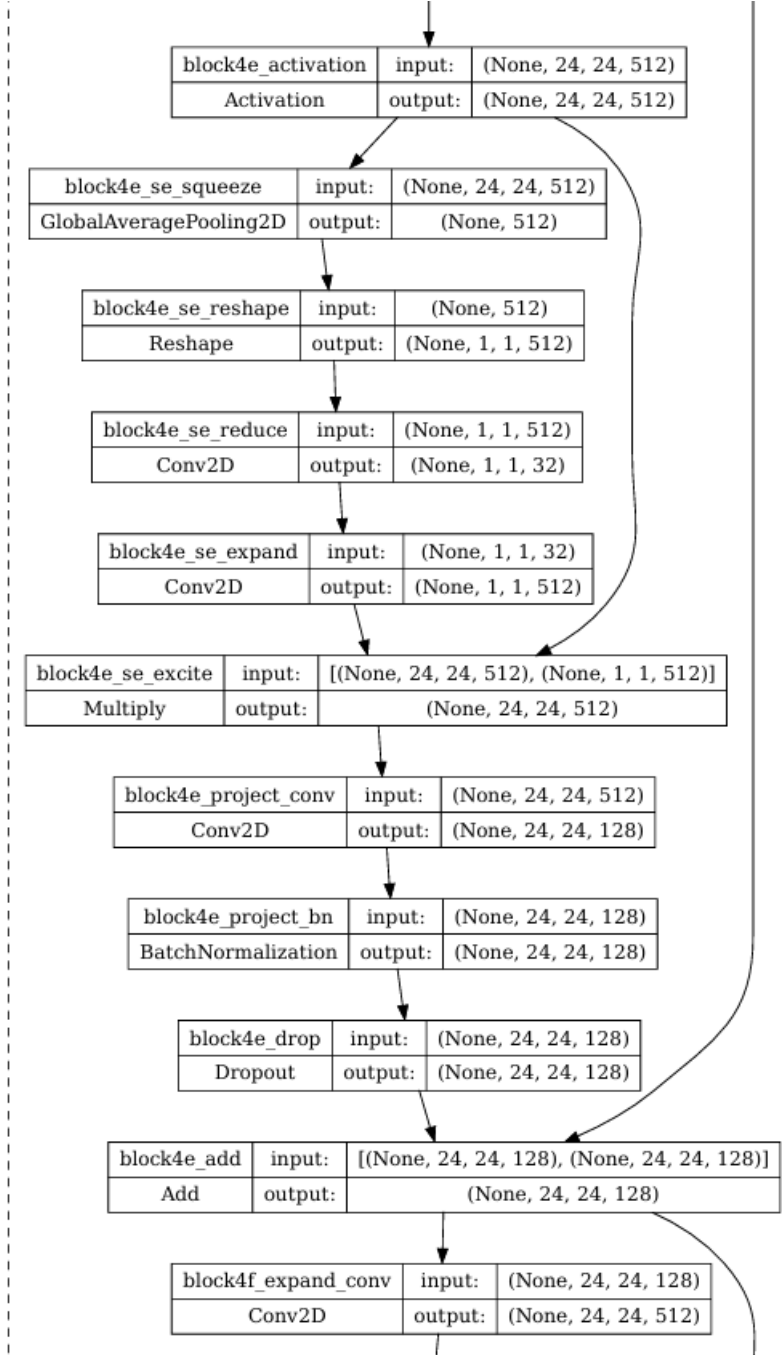
EK 1. (Devam) Geliştirilen Modelin Mimarisi



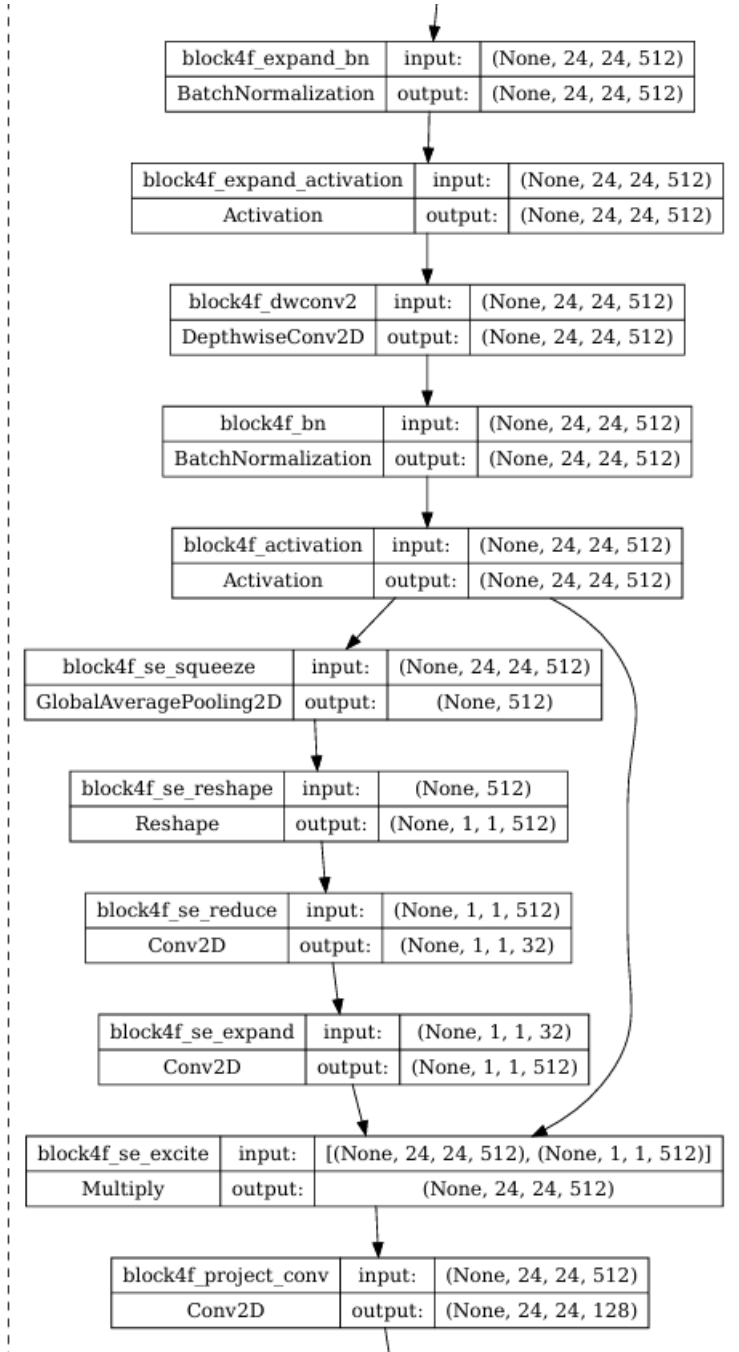
EK 1. (Devam) Geliştirilen Modelin Mimarisi



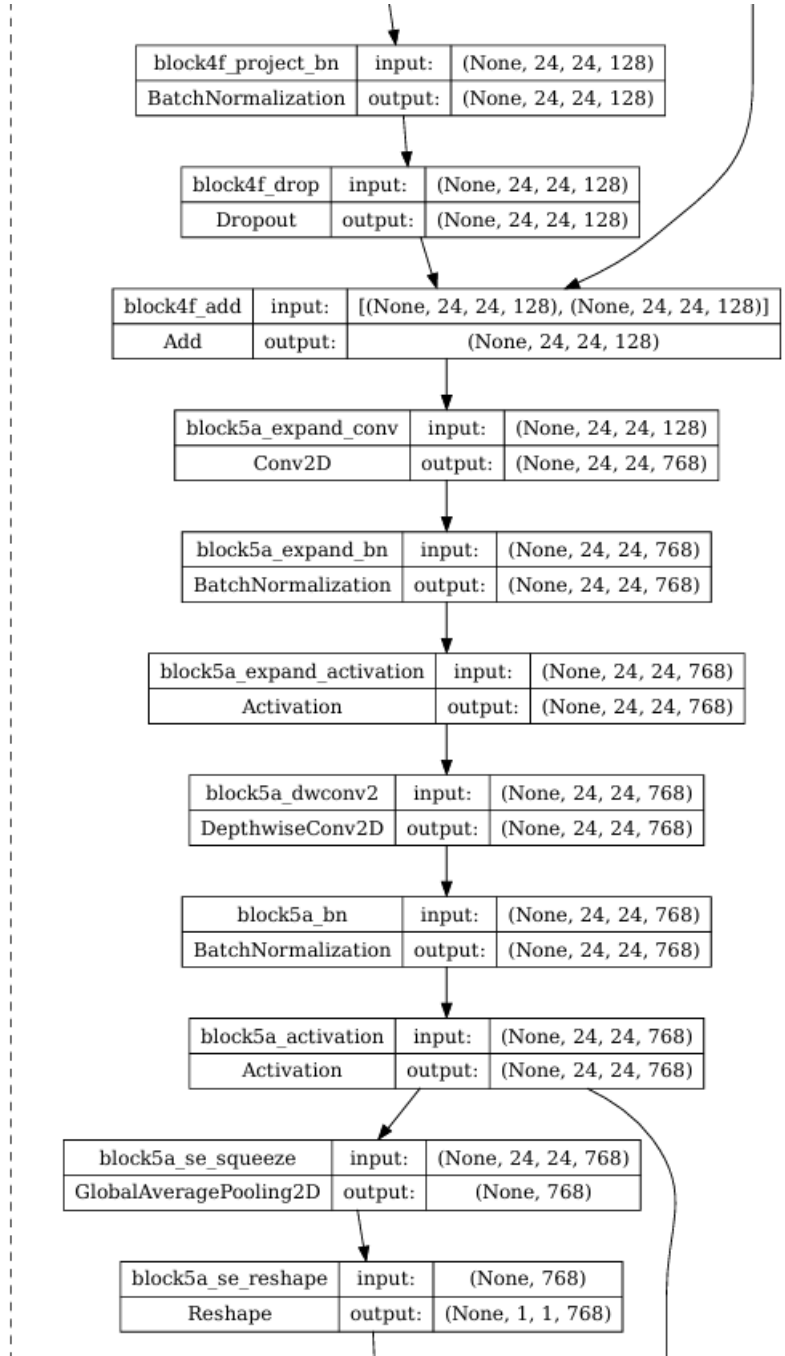
EK 1. (Devam) Geliştirilen Modelin Mimarisi



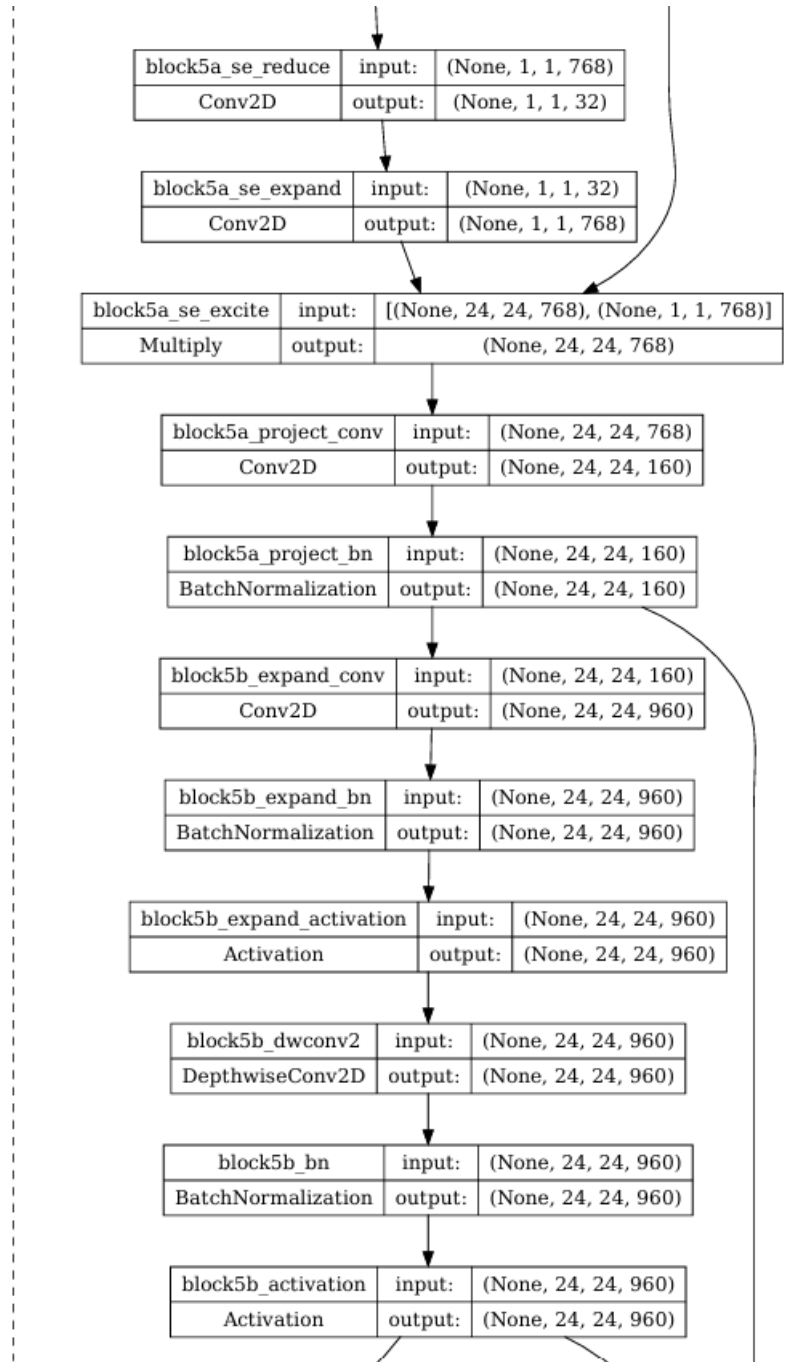
EK 1. (Devam) Geliştirilen Modelin Mimarisi



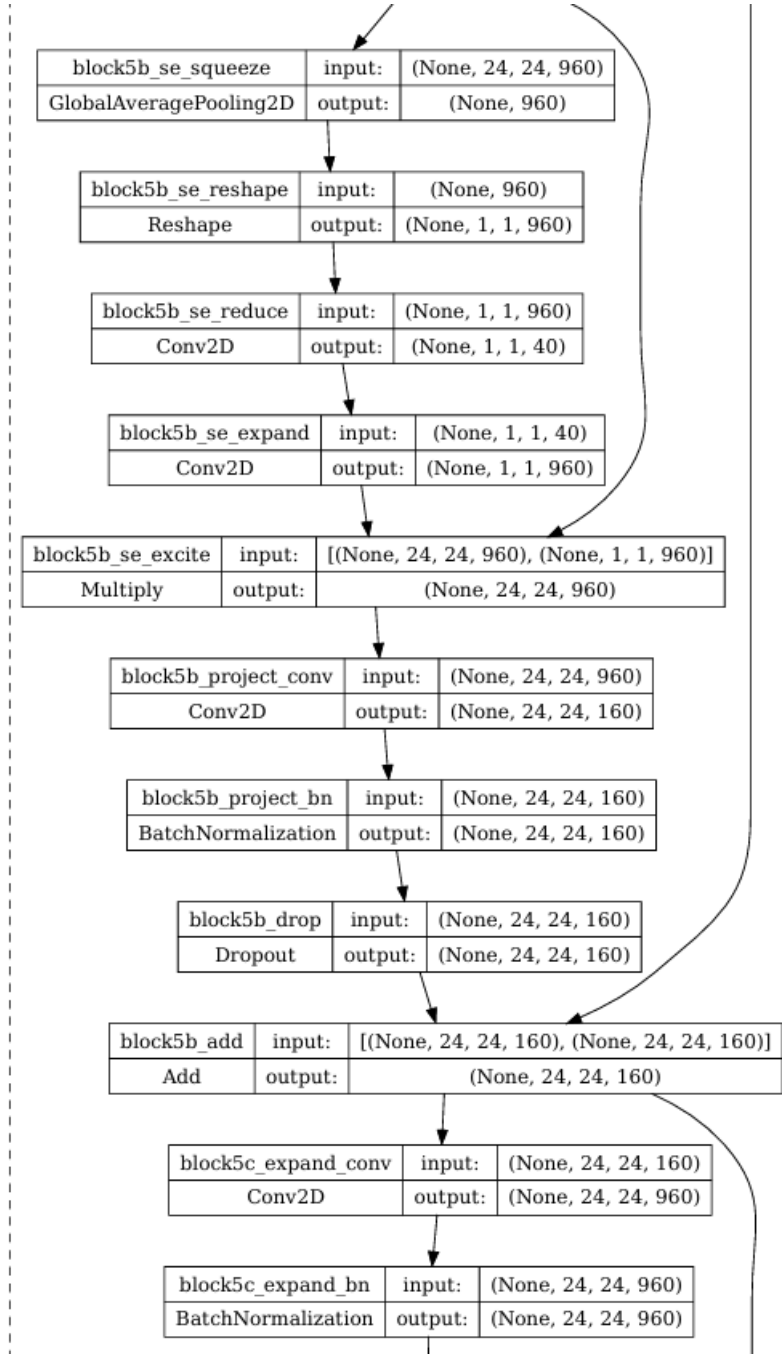
EK 1. (Devam) Geliştirilen Modelin Mimarisi



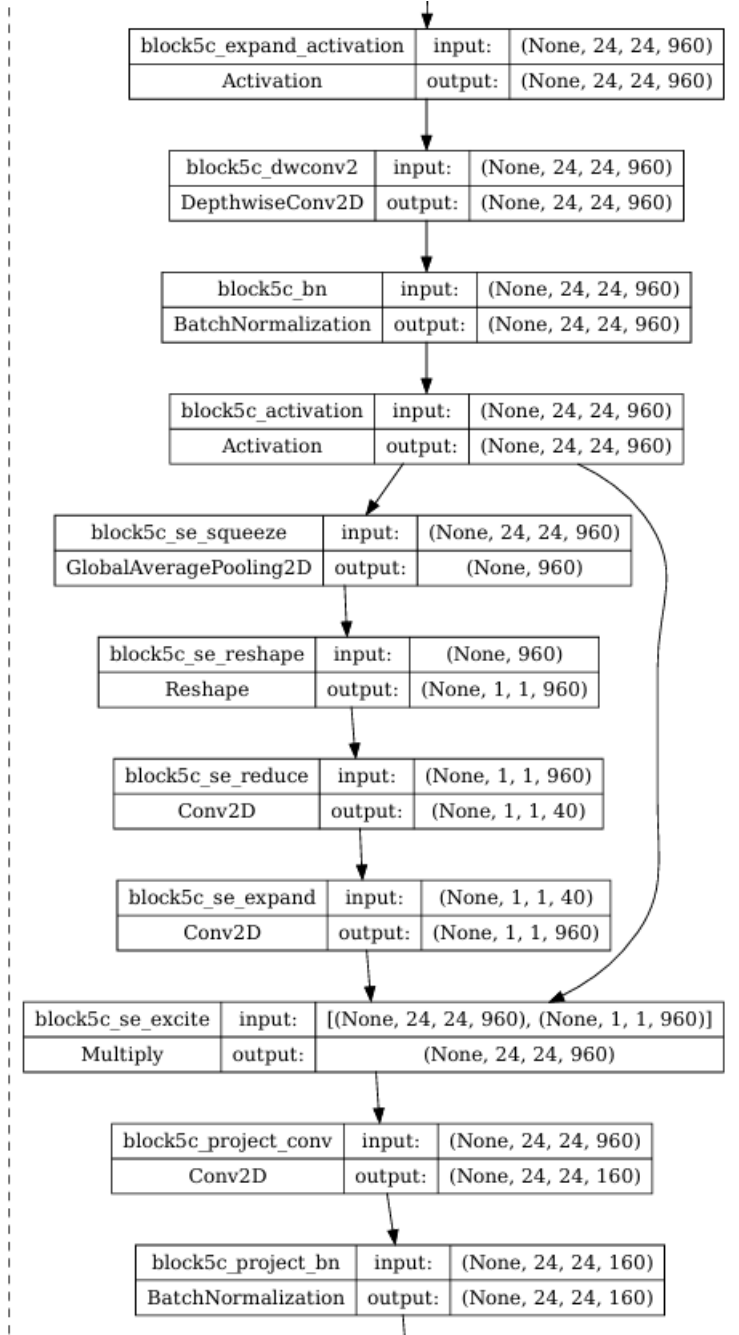
EK 1. (Devam) Geliştirilen Modelin Mimarisi



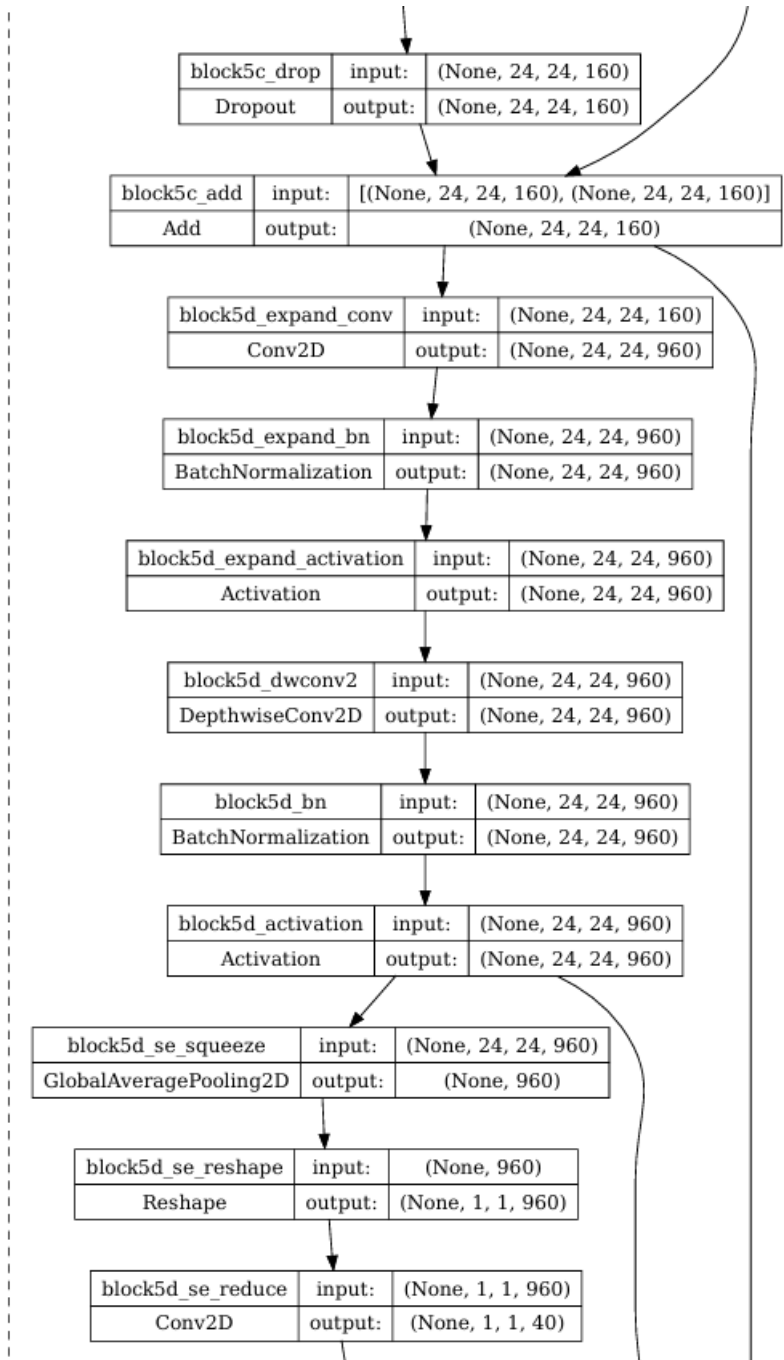
EK 1. (Devam) Geliştirilen Modelin Mimarisi



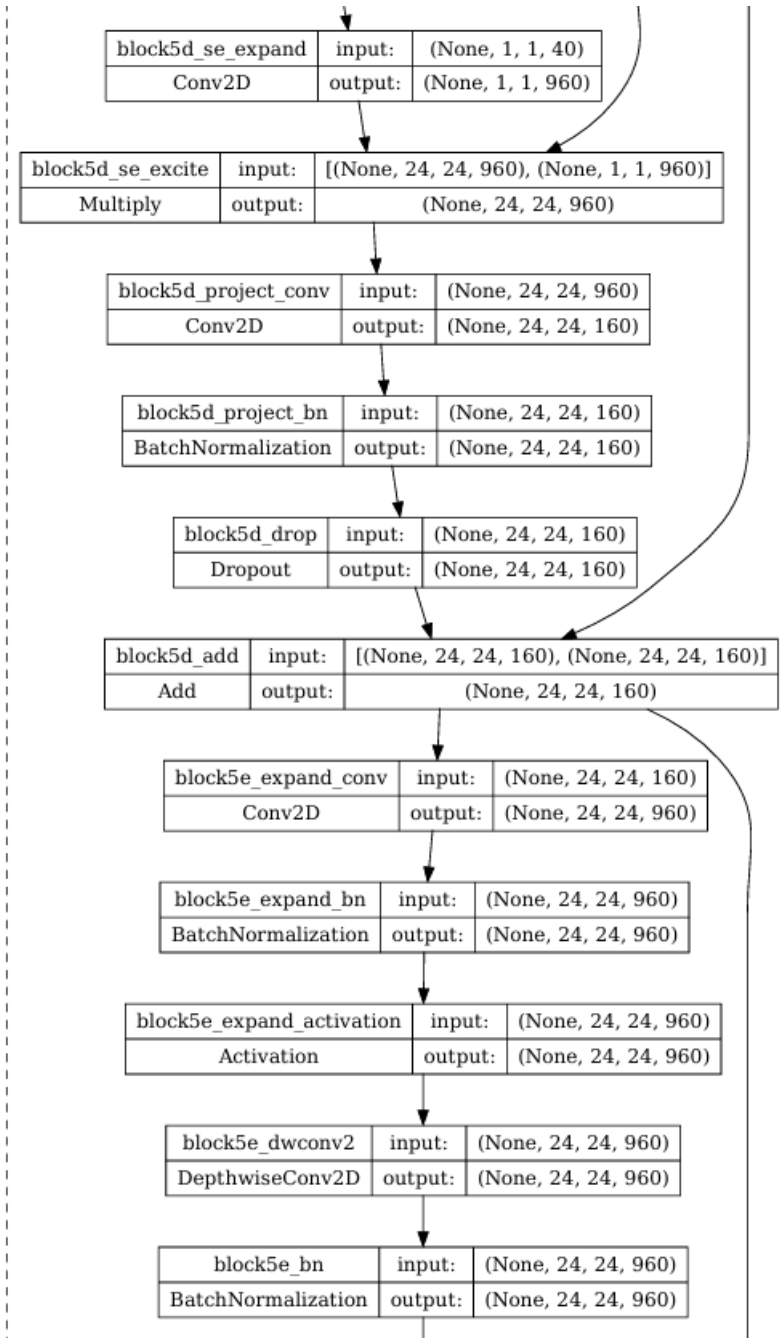
EK 1. (Devam) Geliştirilen Modelin Mimarisi



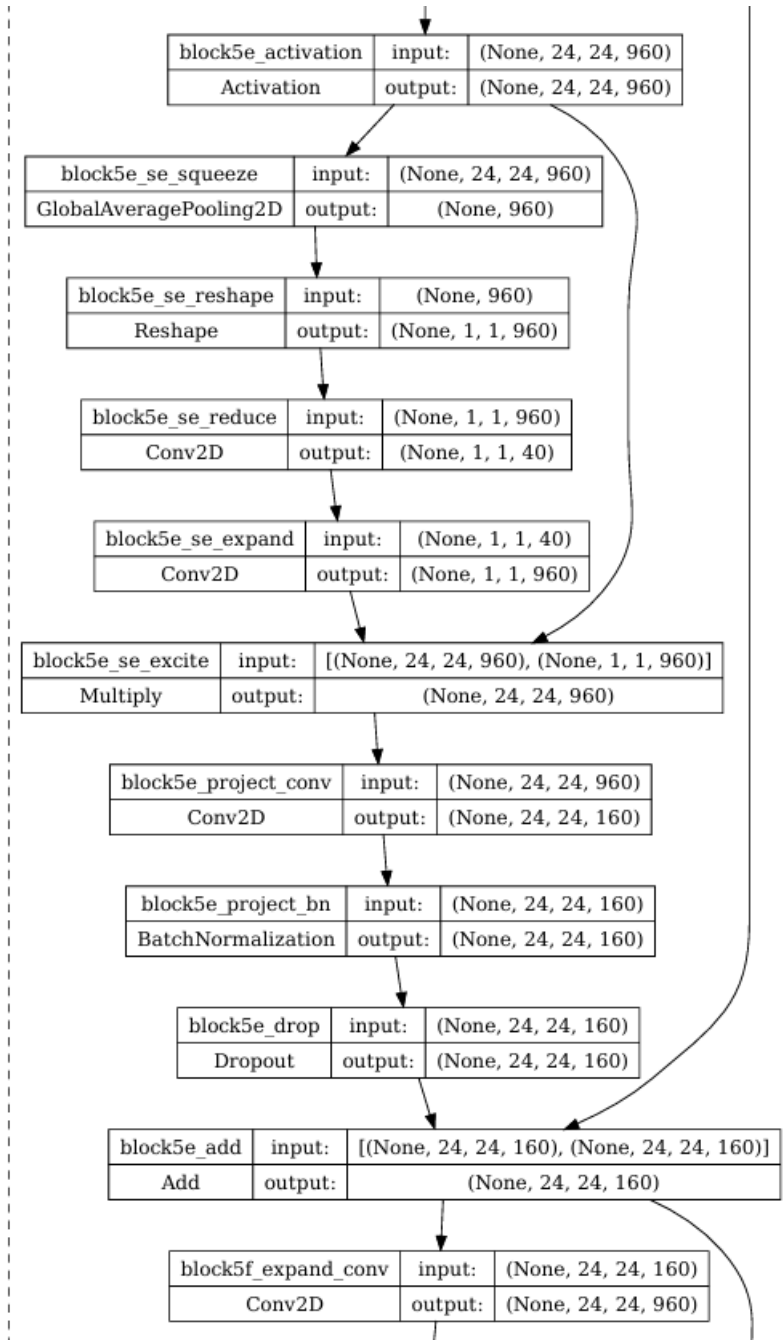
EK 1. (Devam) Geliştirilen Modelin Mimarisi



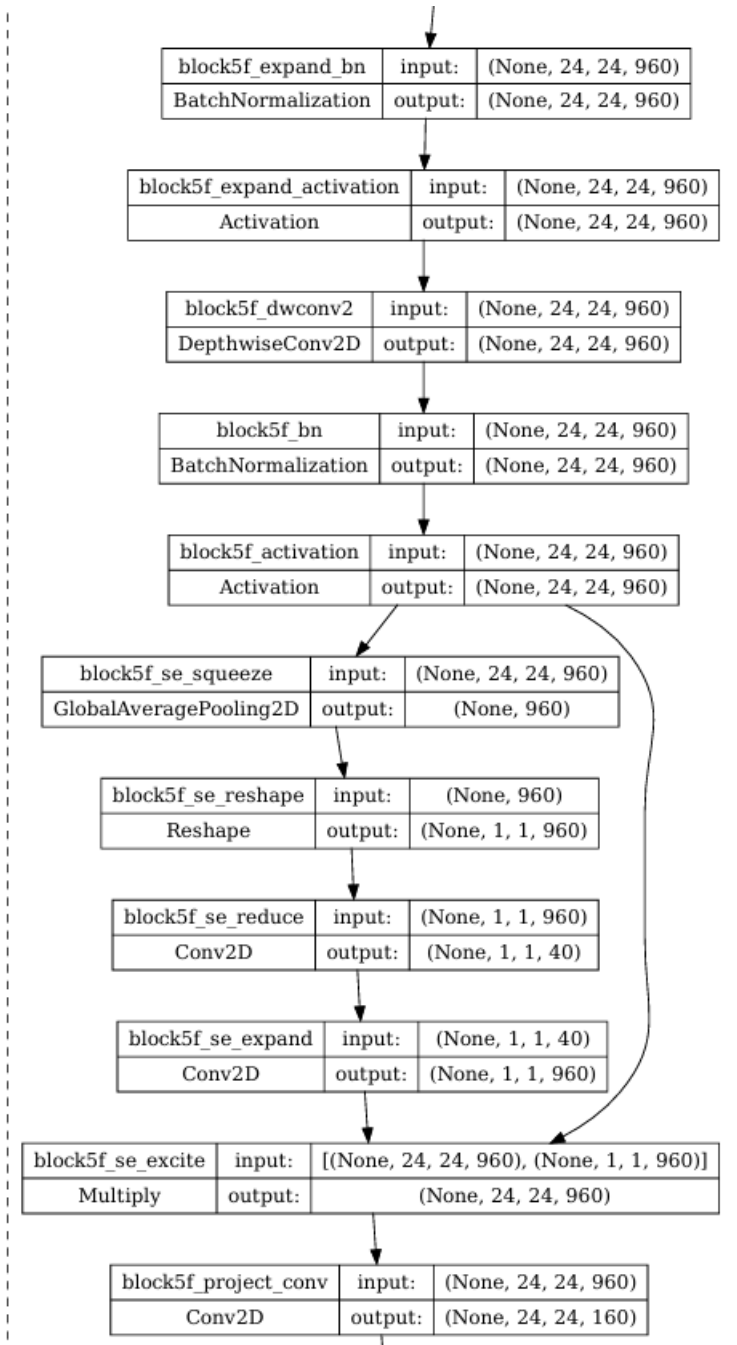
EK 1. (Devam) Geliştirilen Modelin Mimarisi



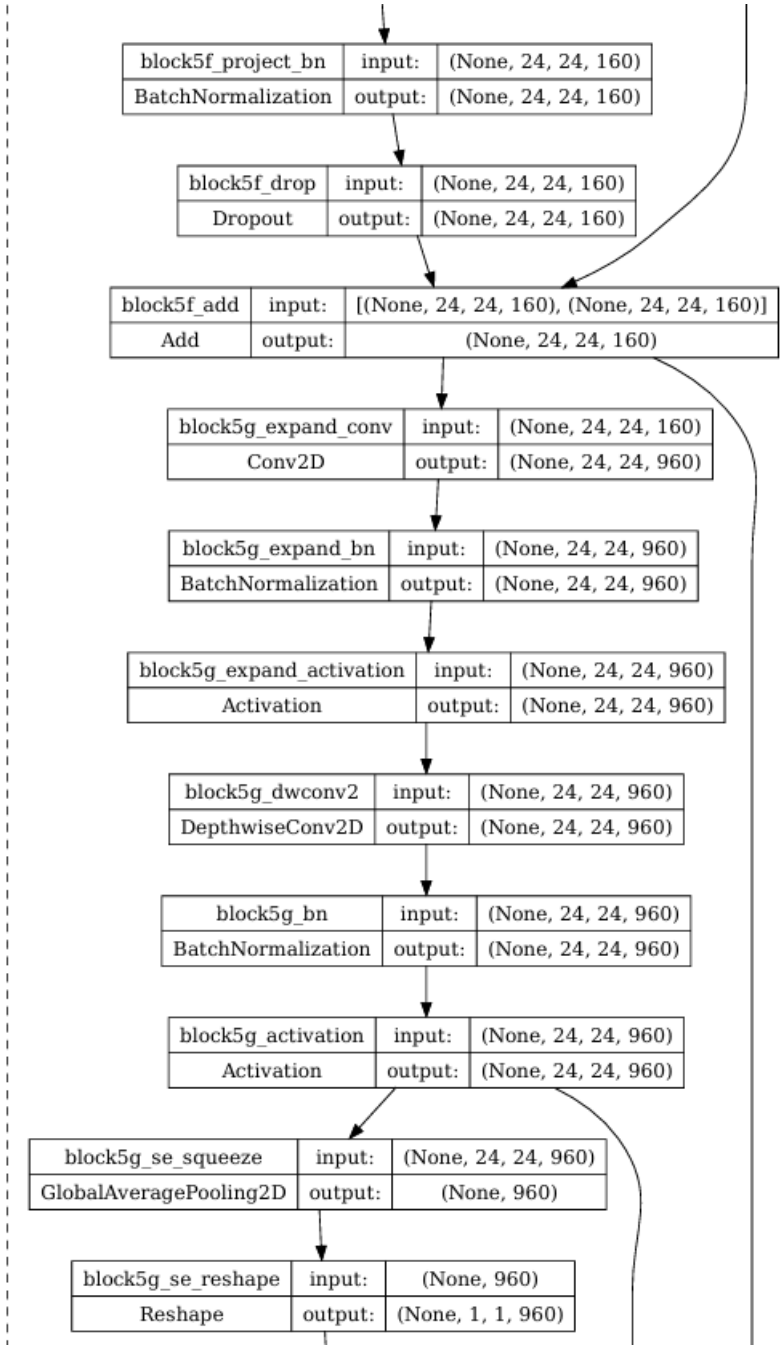
EK 1. (Devam) Geliştirilen Modelin Mimarisi



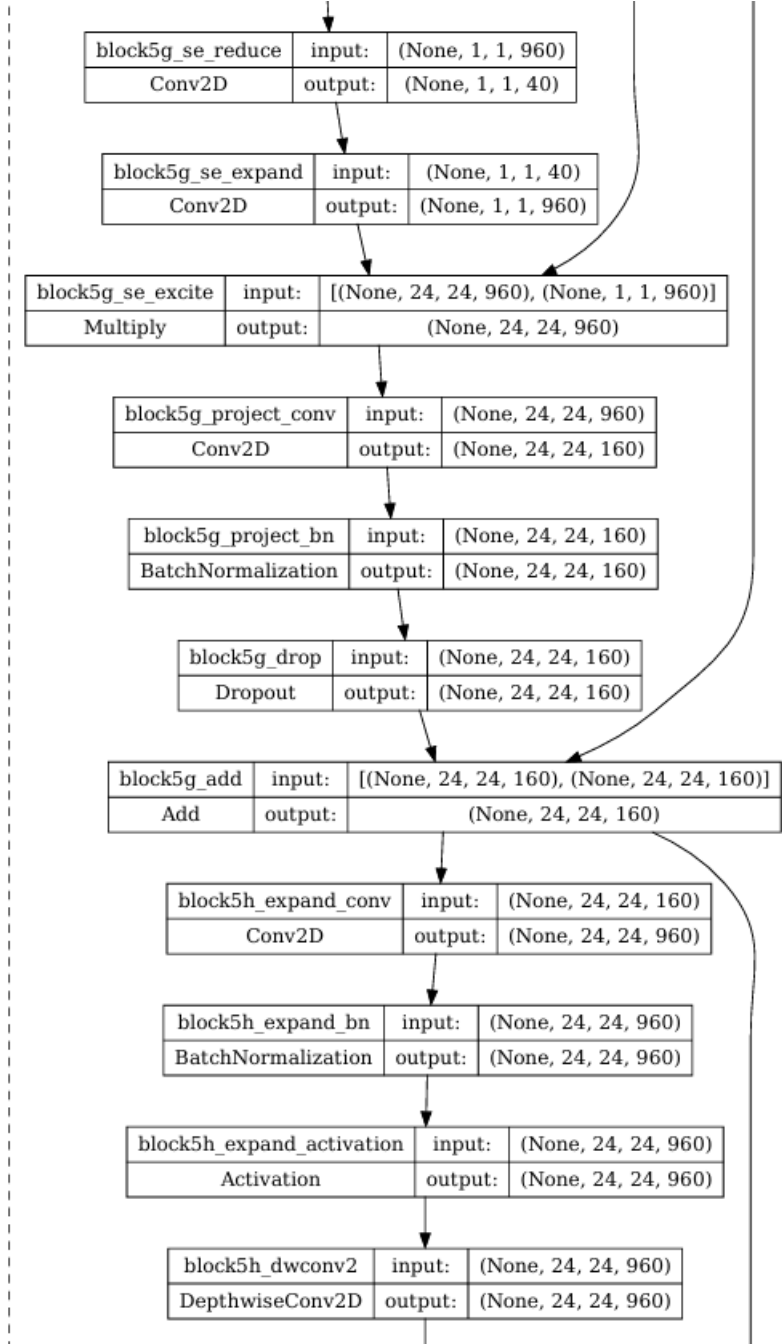
EK 1. (Devam) Geliştirilen Modelin Mimarisi



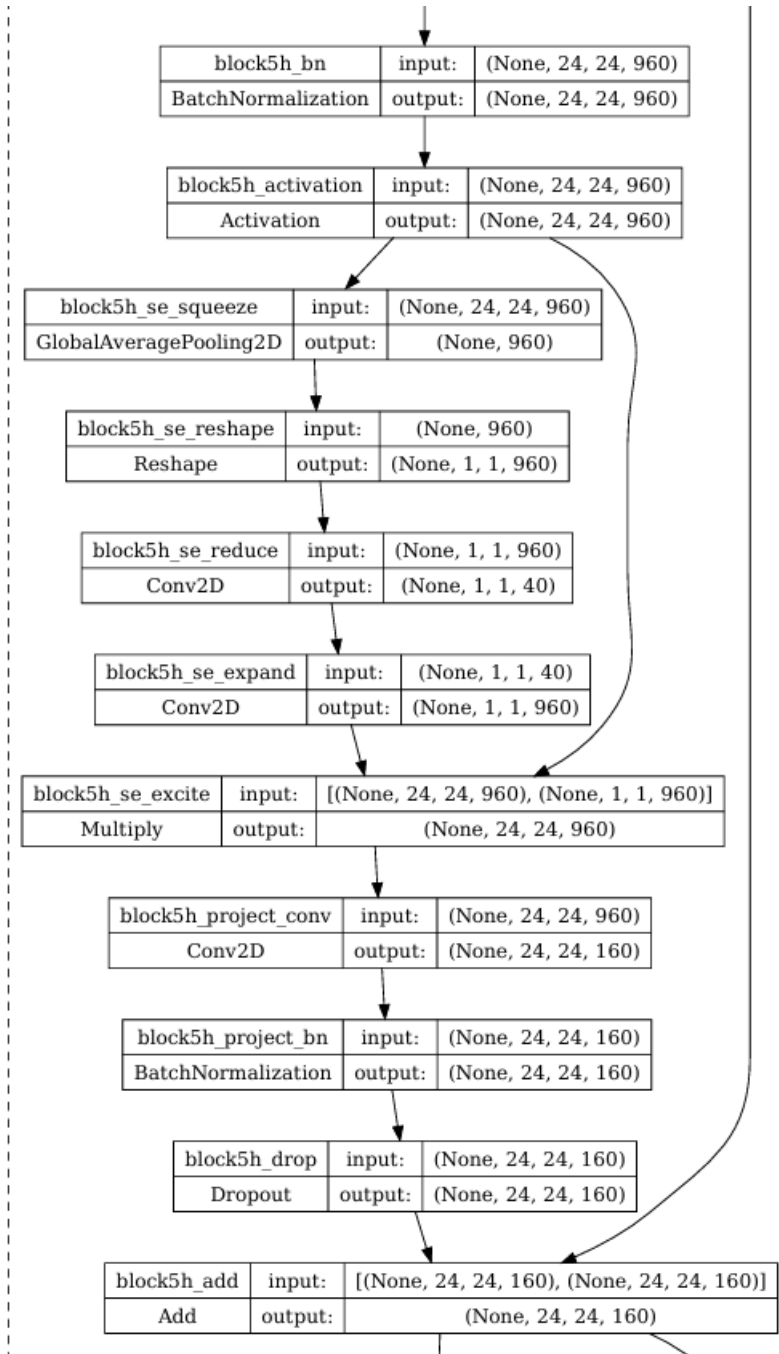
EK 1. (Devam) Geliştirilen Modelin Mimarisi



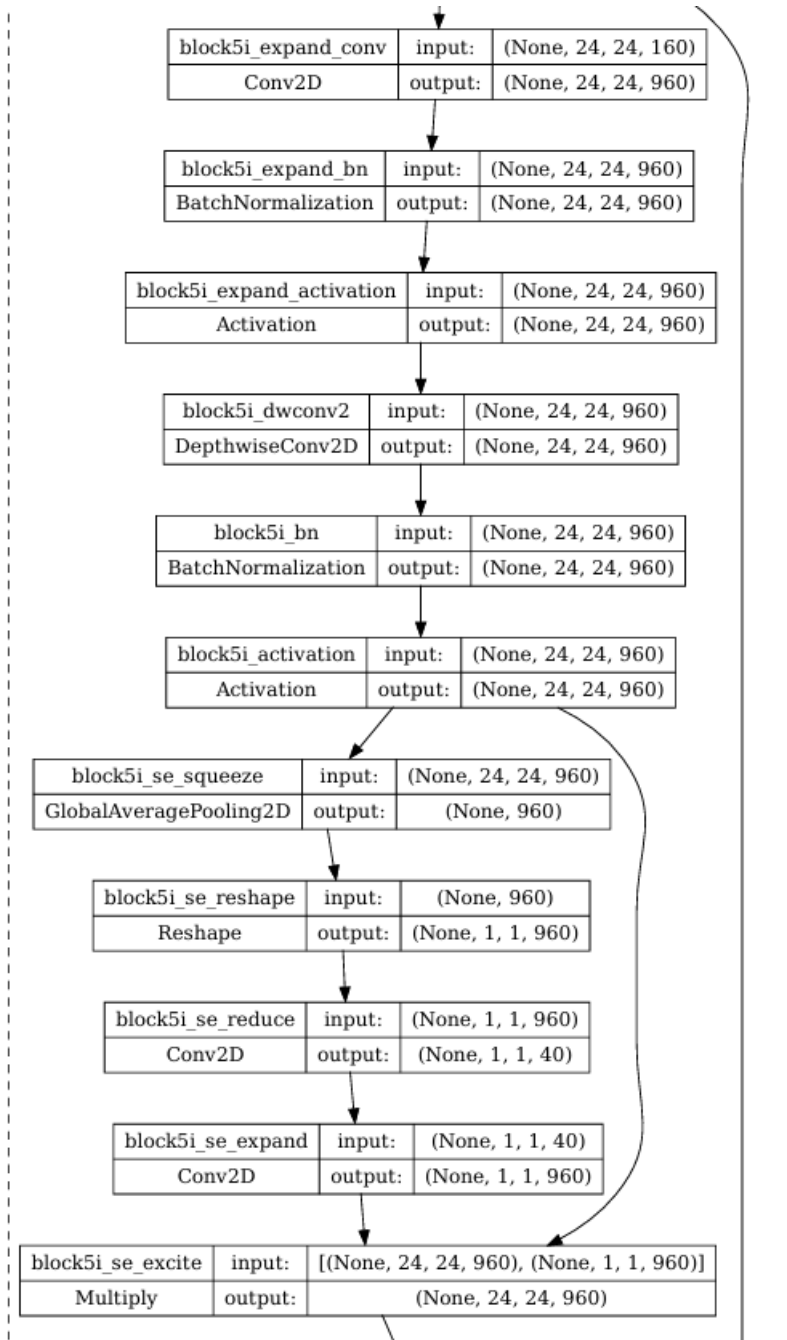
EK 1. (Devam) Geliştirilen Modelin Mimarisi



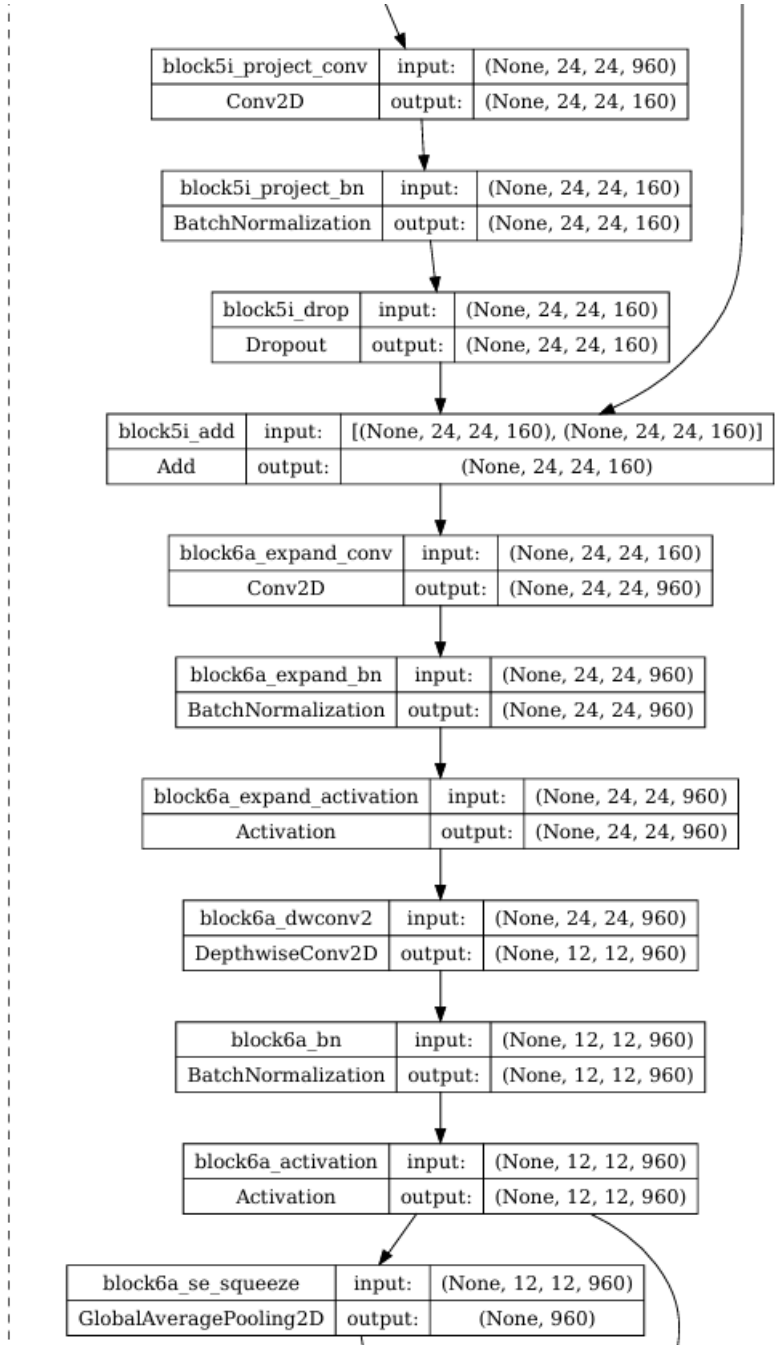
EK 1. (Devam) Geliştirilen Modelin Mimarisi



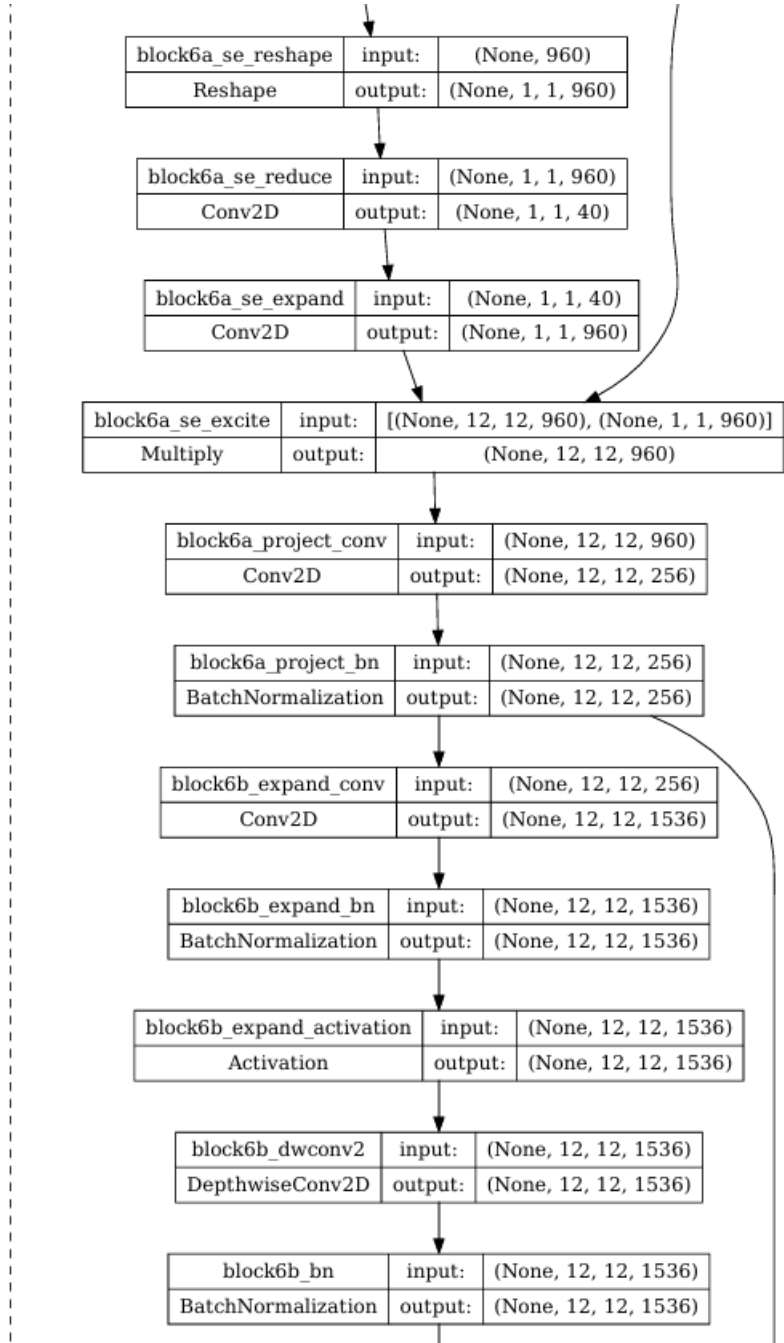
EK 1. (Devam) Geliştirilen Modelin Mimarisi



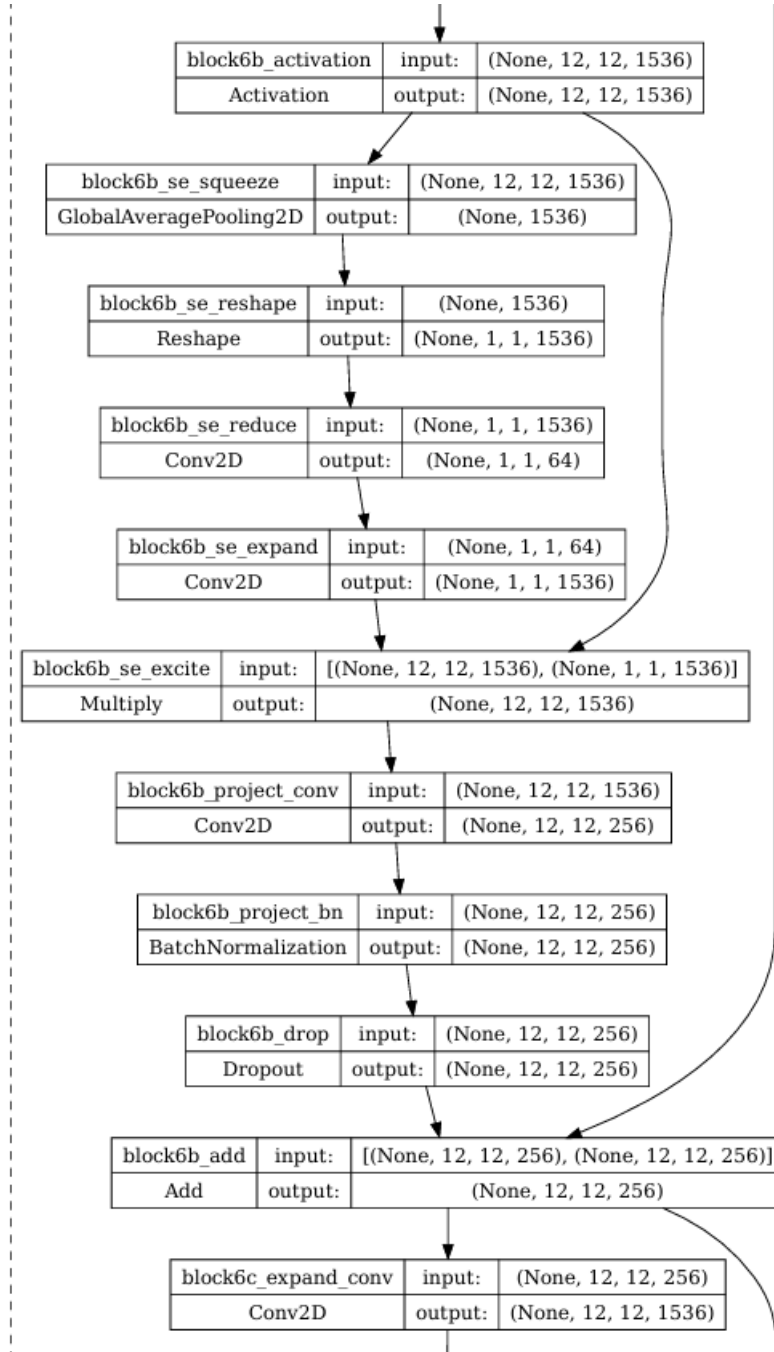
EK 1. (Devam) Geliştirilen Modelin Mimarisi



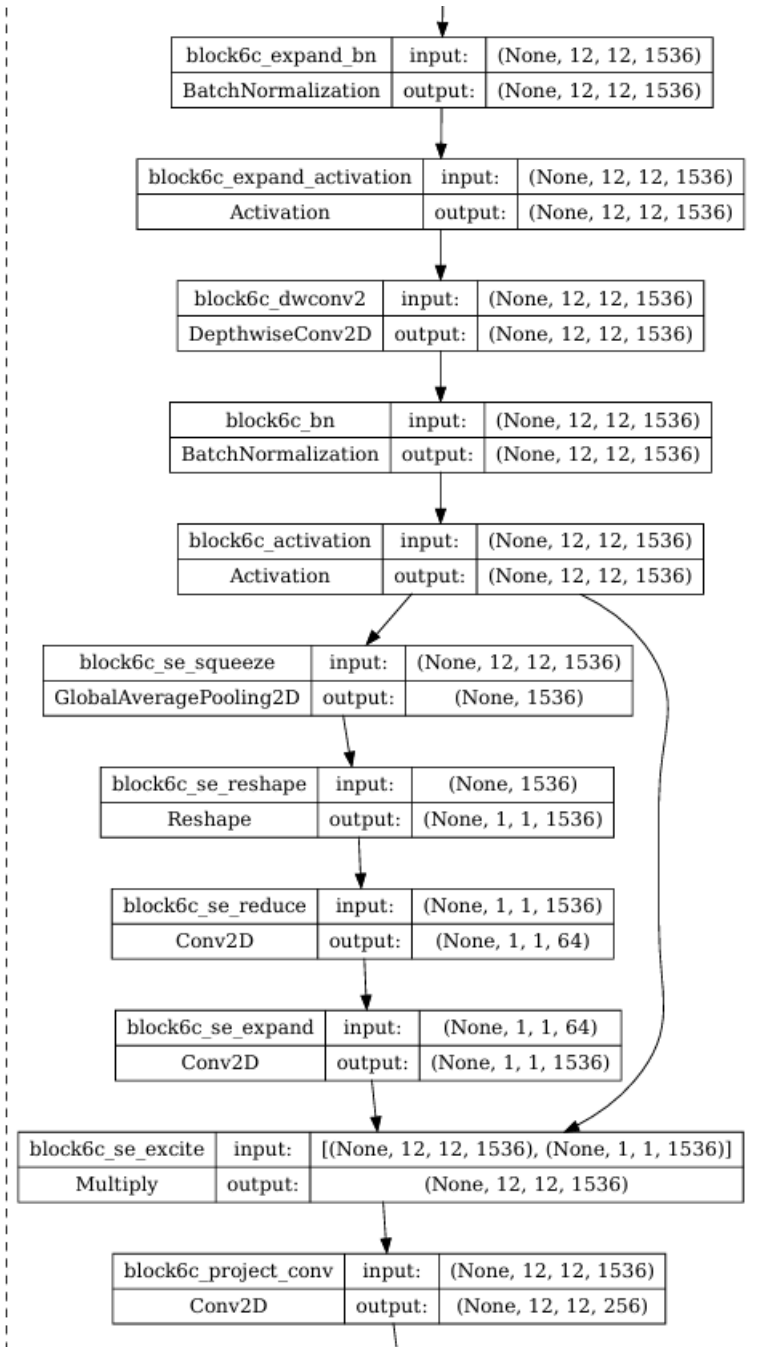
EK 1. (Devam) Geliştirilen Modelin Mimarisi



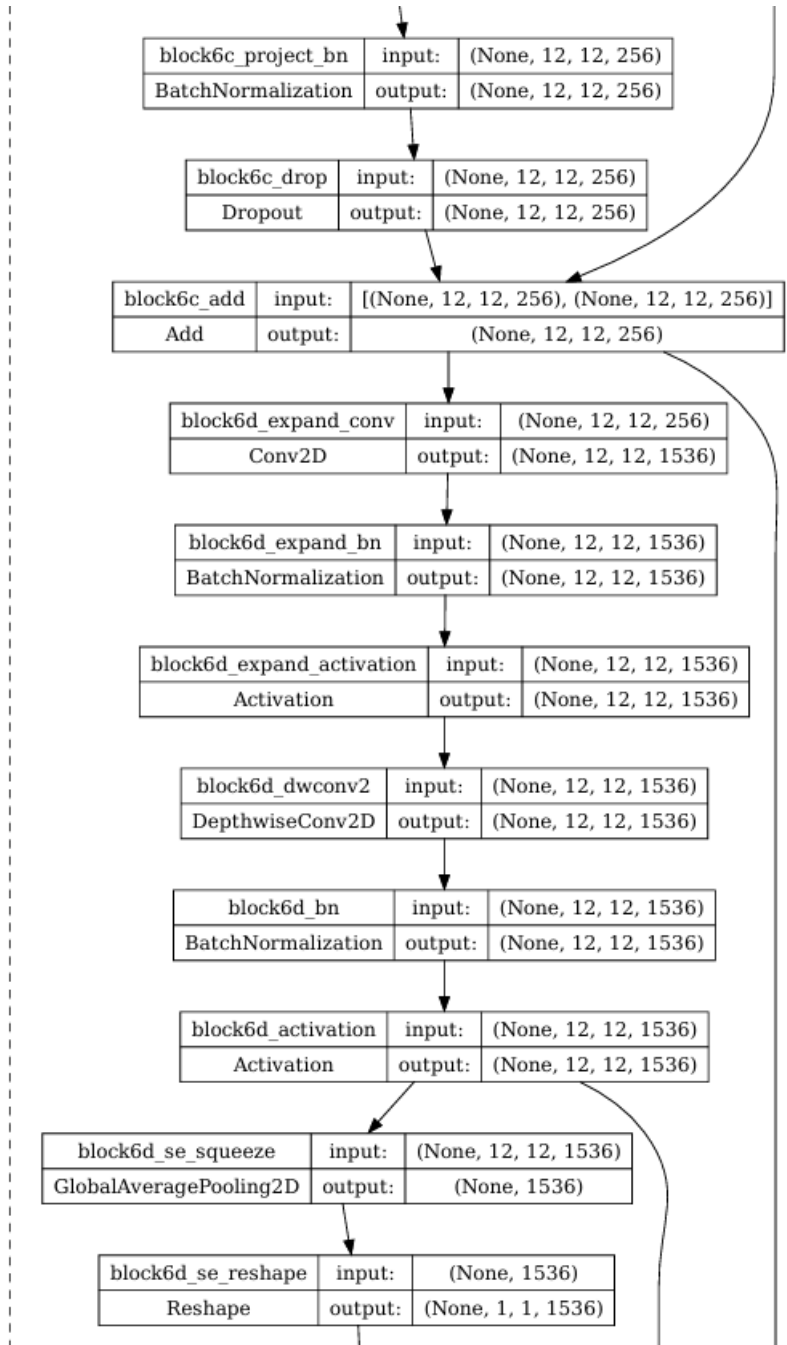
EK 1. (Devam) Geliştirilen Modelin Mimarisi



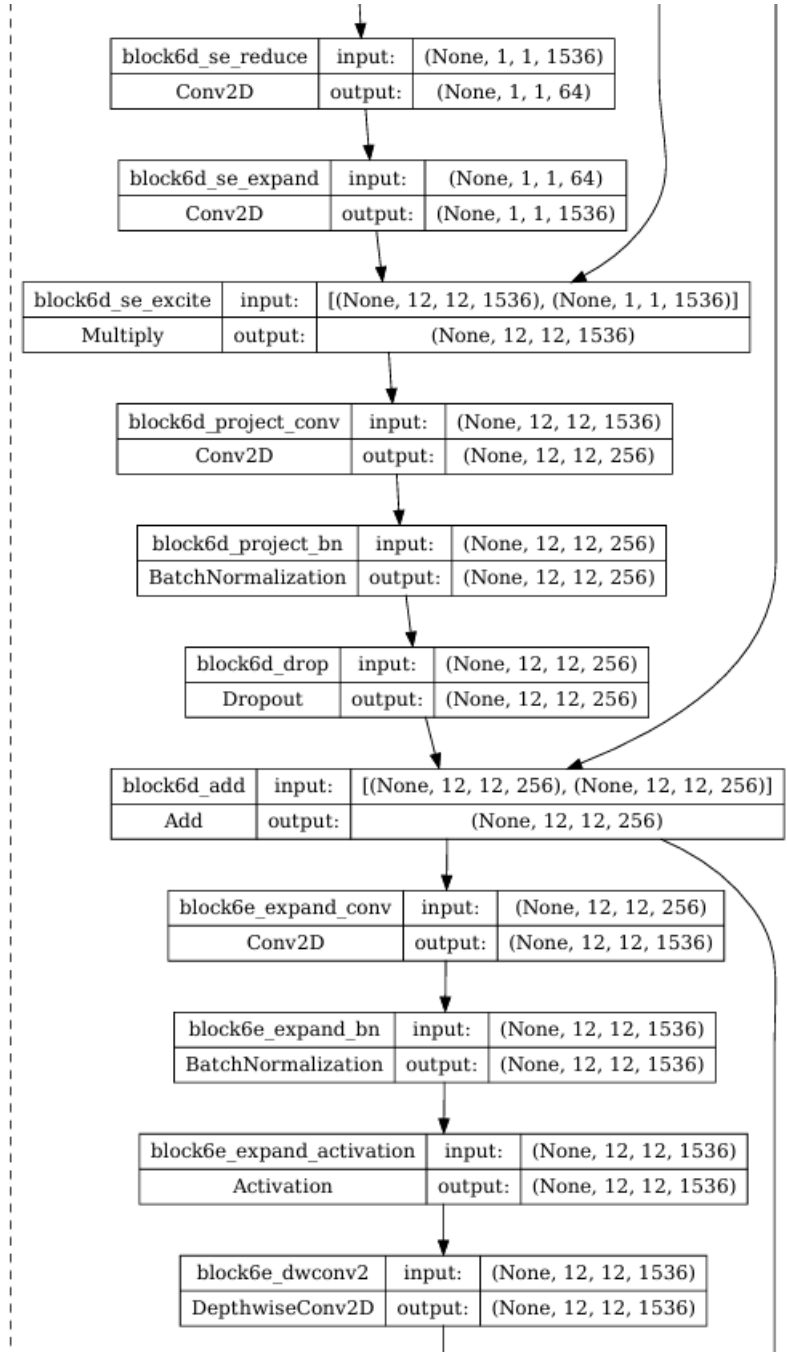
EK 1. (Devam) Geliştirilen Modelin Mimarisi



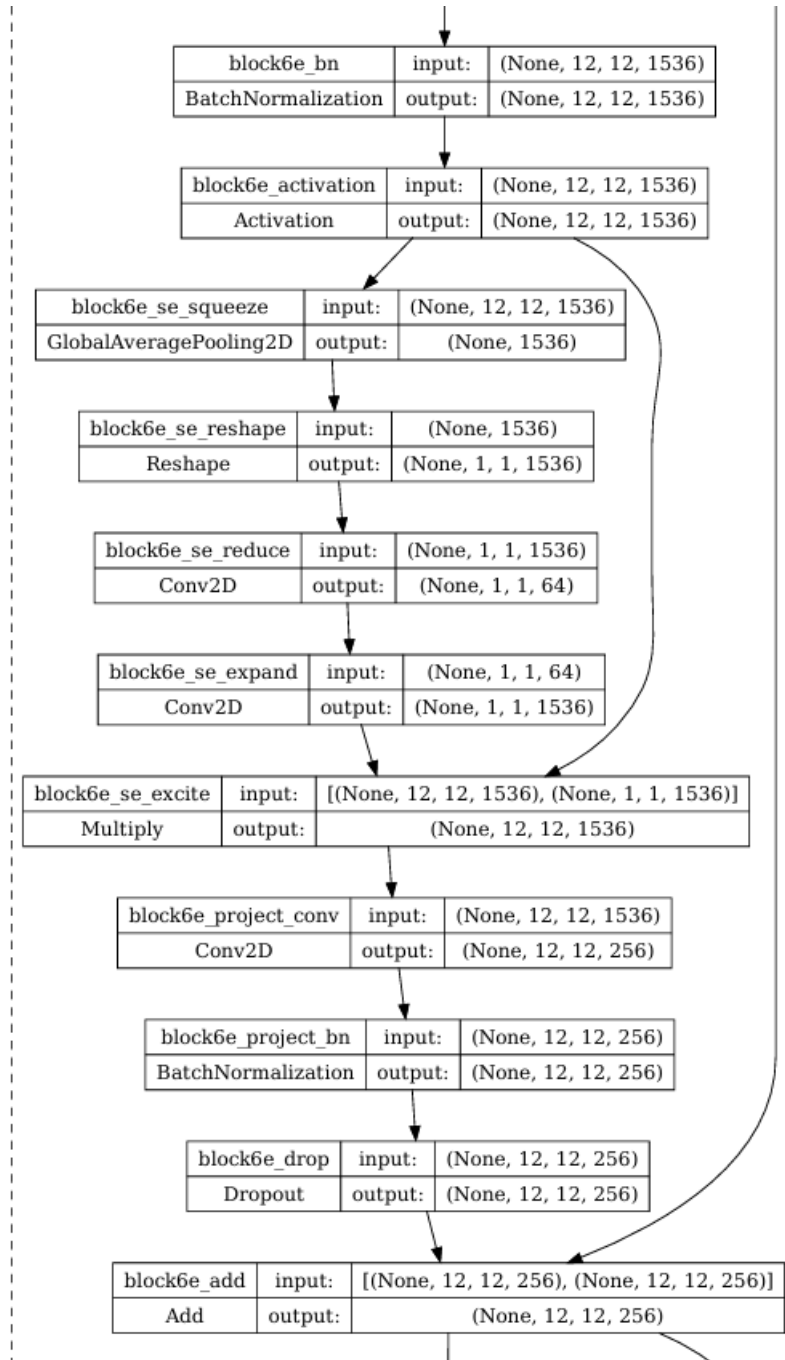
EK 1. (Devam) Geliştirilen Modelin Mimarisi



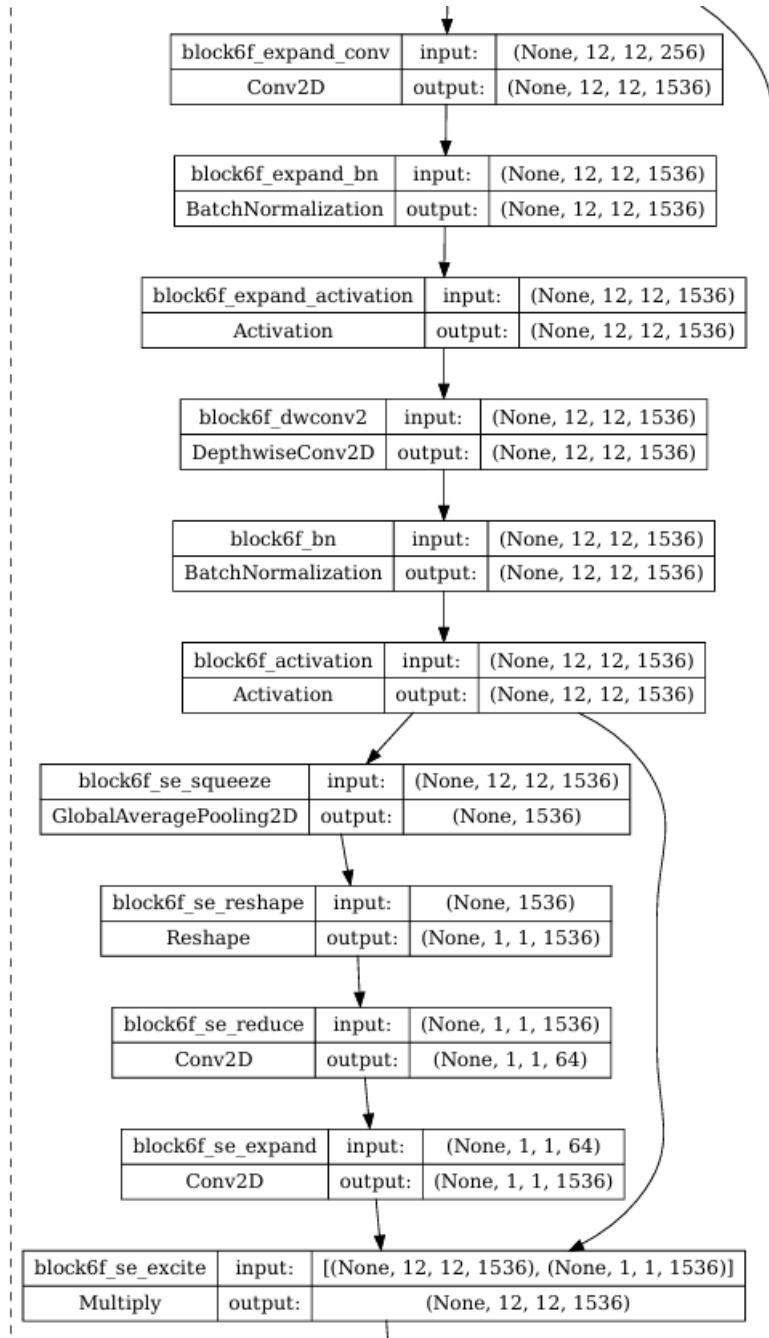
EK 1. (Devam) Geliştirilen Modelin Mimarisi



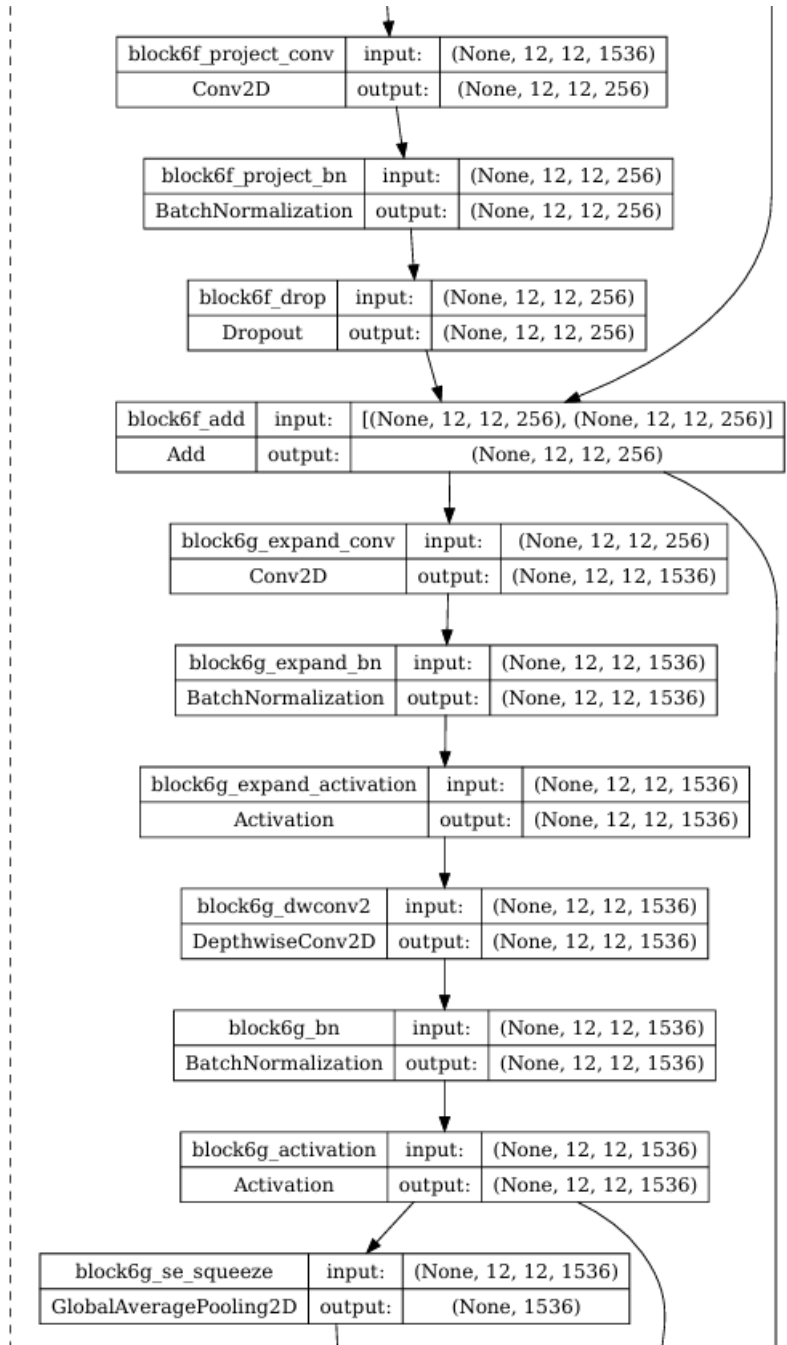
EK 1. (Devam) Geliştirilen Modelin Mimarisi



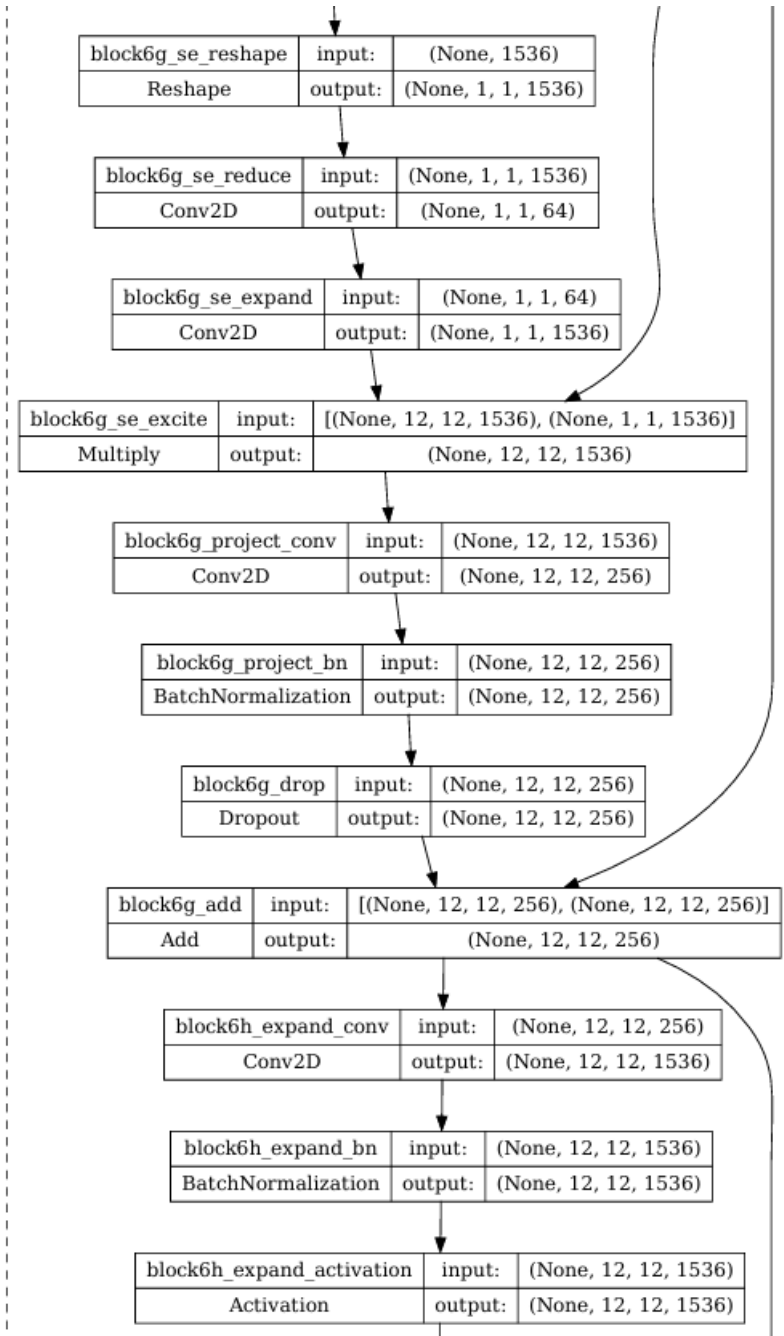
EK 1. (Devam) Geliştirilen Modelin Mimarisi



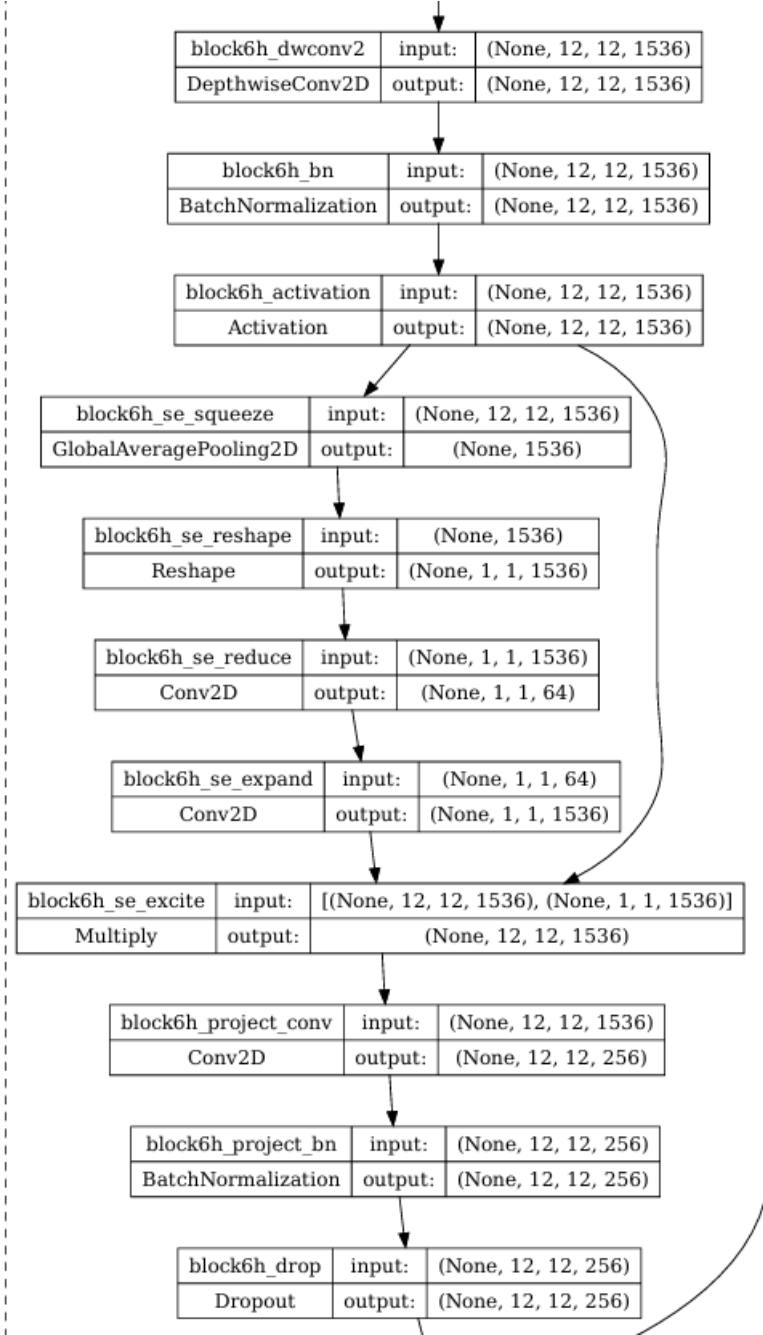
EK 1. (Devam) Geliştirilen Modelin Mimarisi



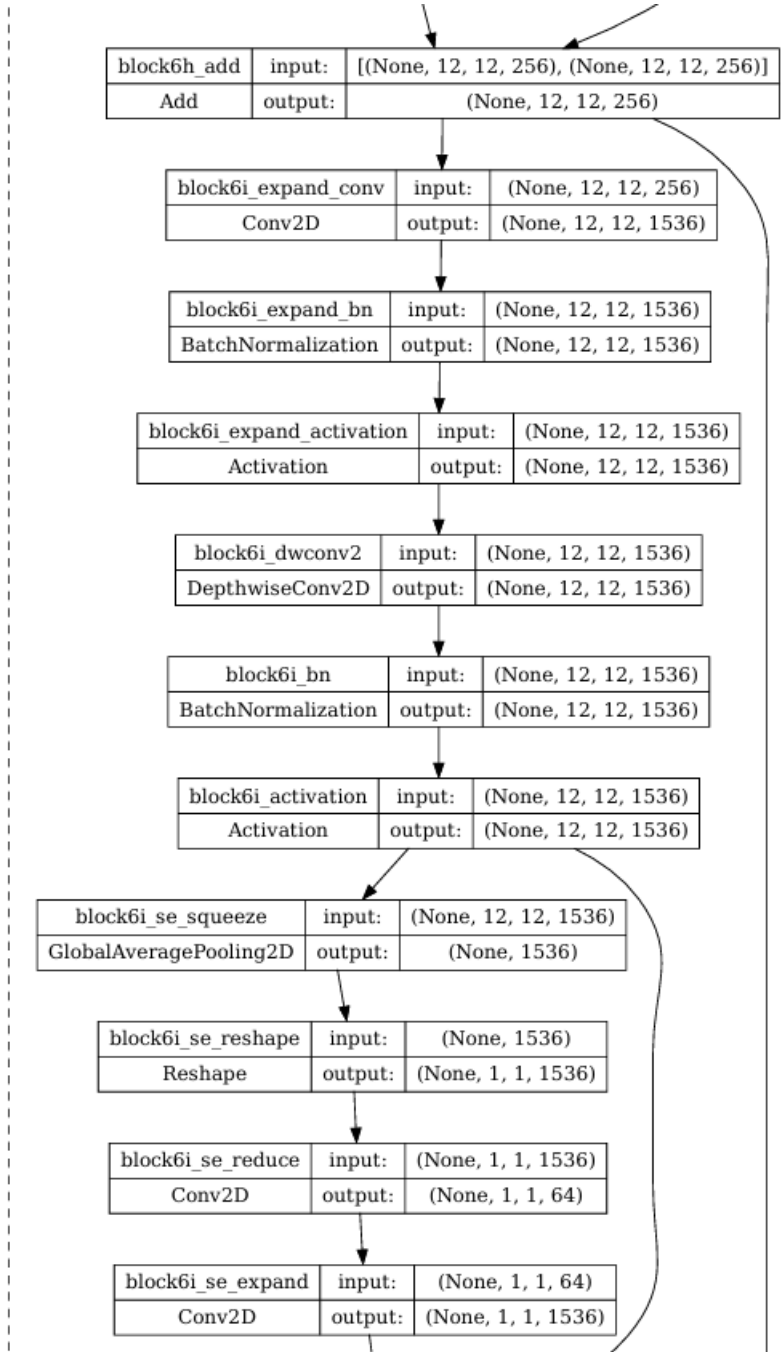
EK 1. (Devam) Geliştirilen Modelin Mimarisi



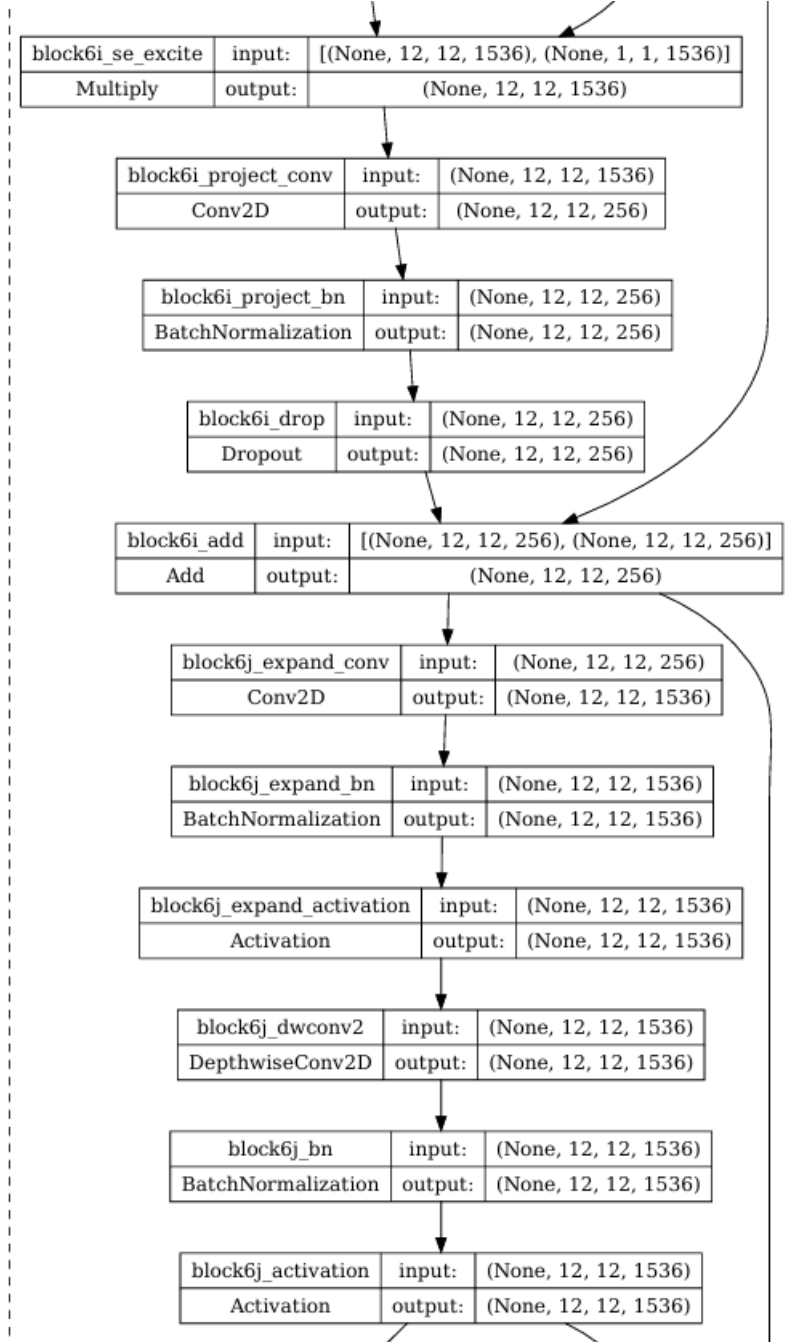
EK 1. (Devam) Geliştirilen Modelin Mimarisi



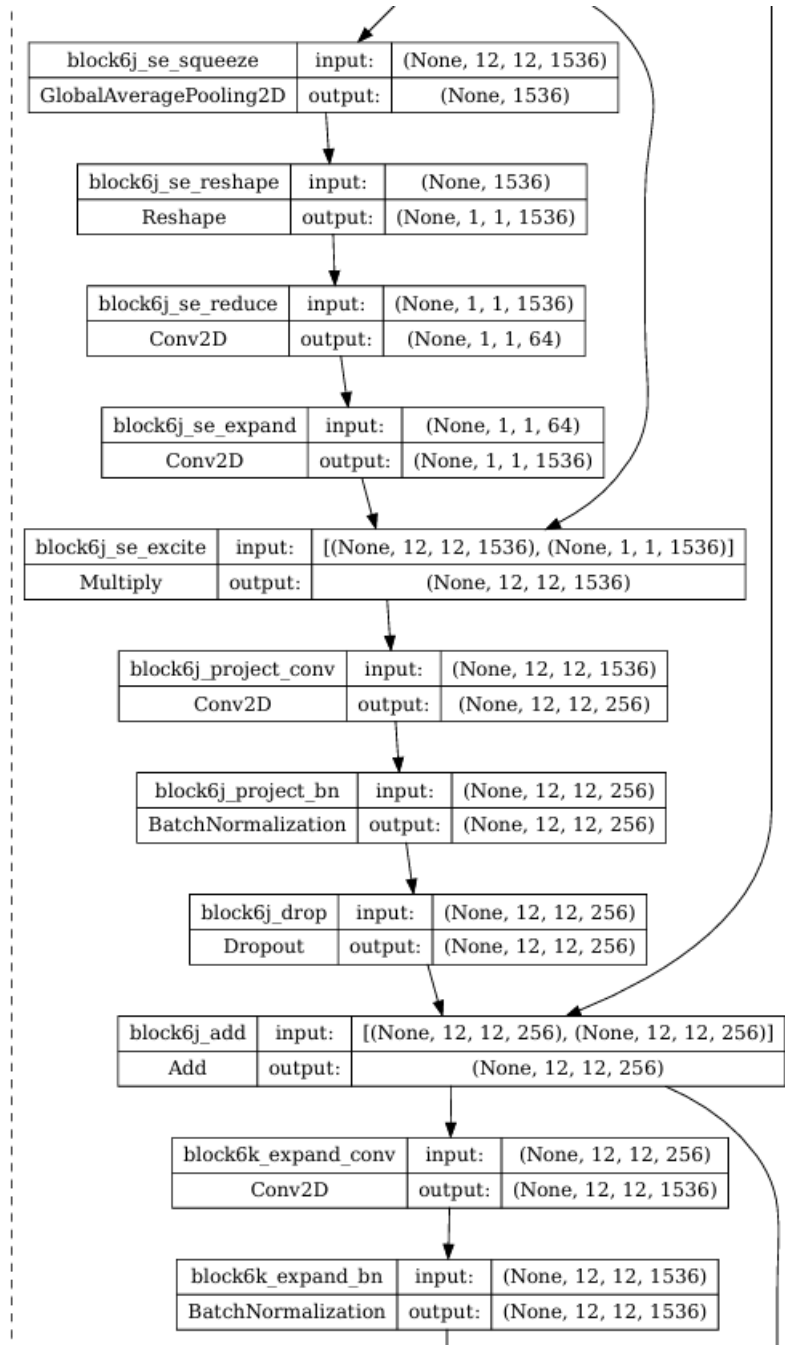
EK 1. (Devam) Geliştirilen Modelin Mimarisi



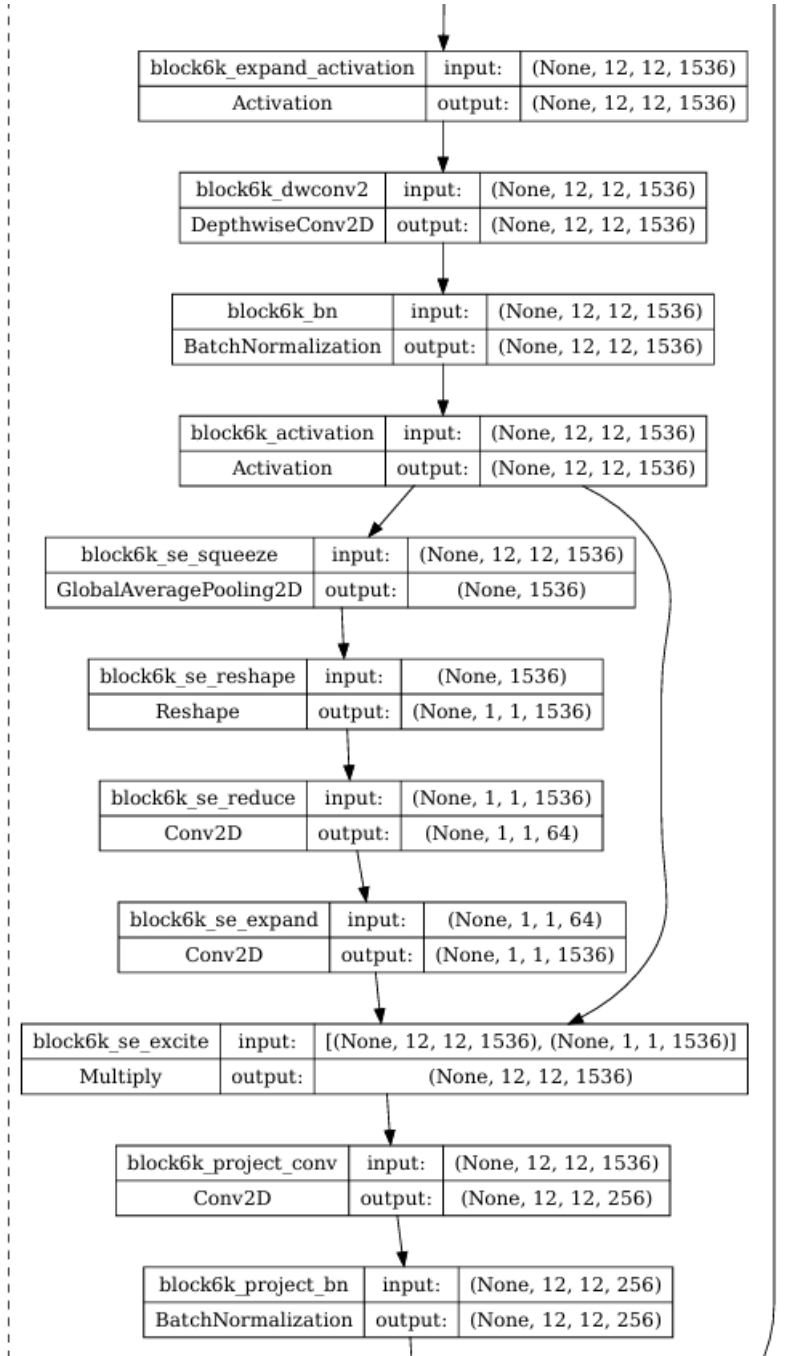
EK 1. (Devam) Geliştirilen Modelin Mimarisi



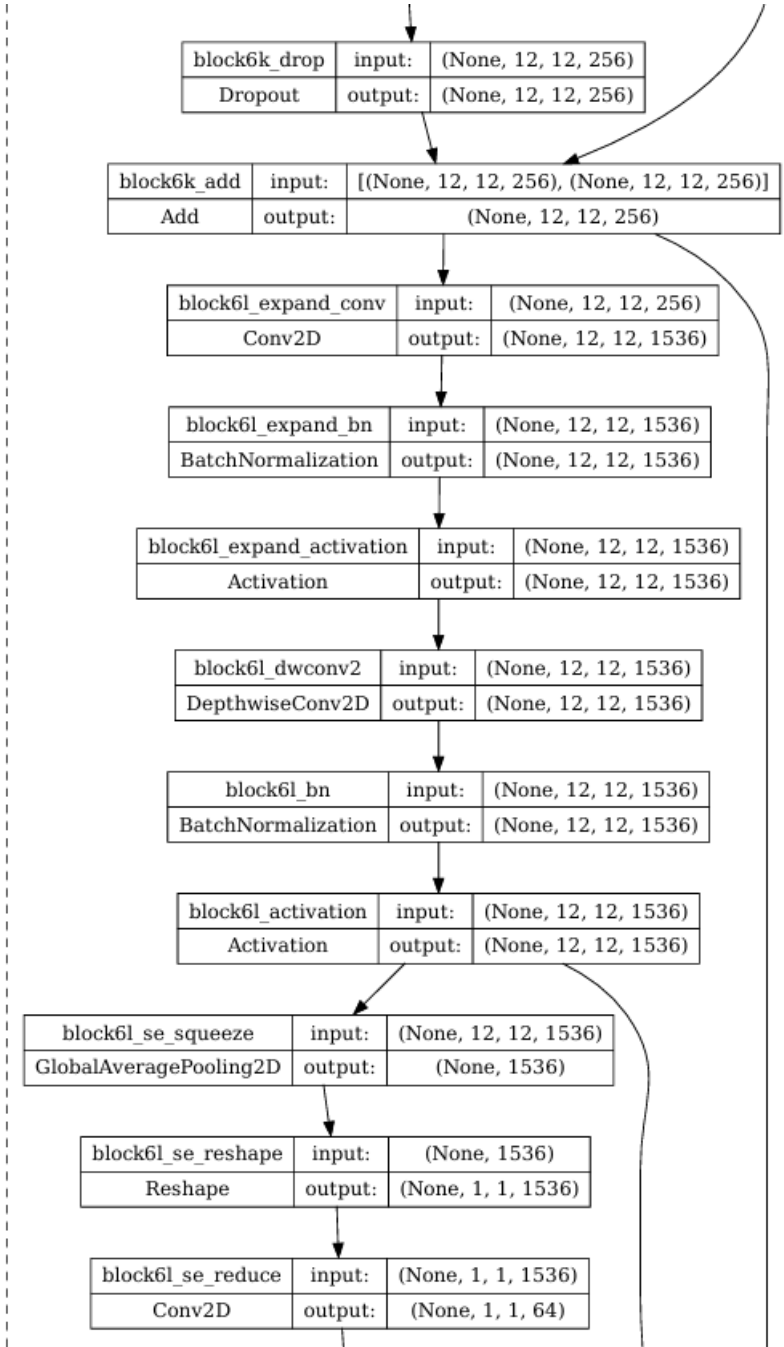
EK 1. (Devam) Geliştirilen Modelin Mimarisi



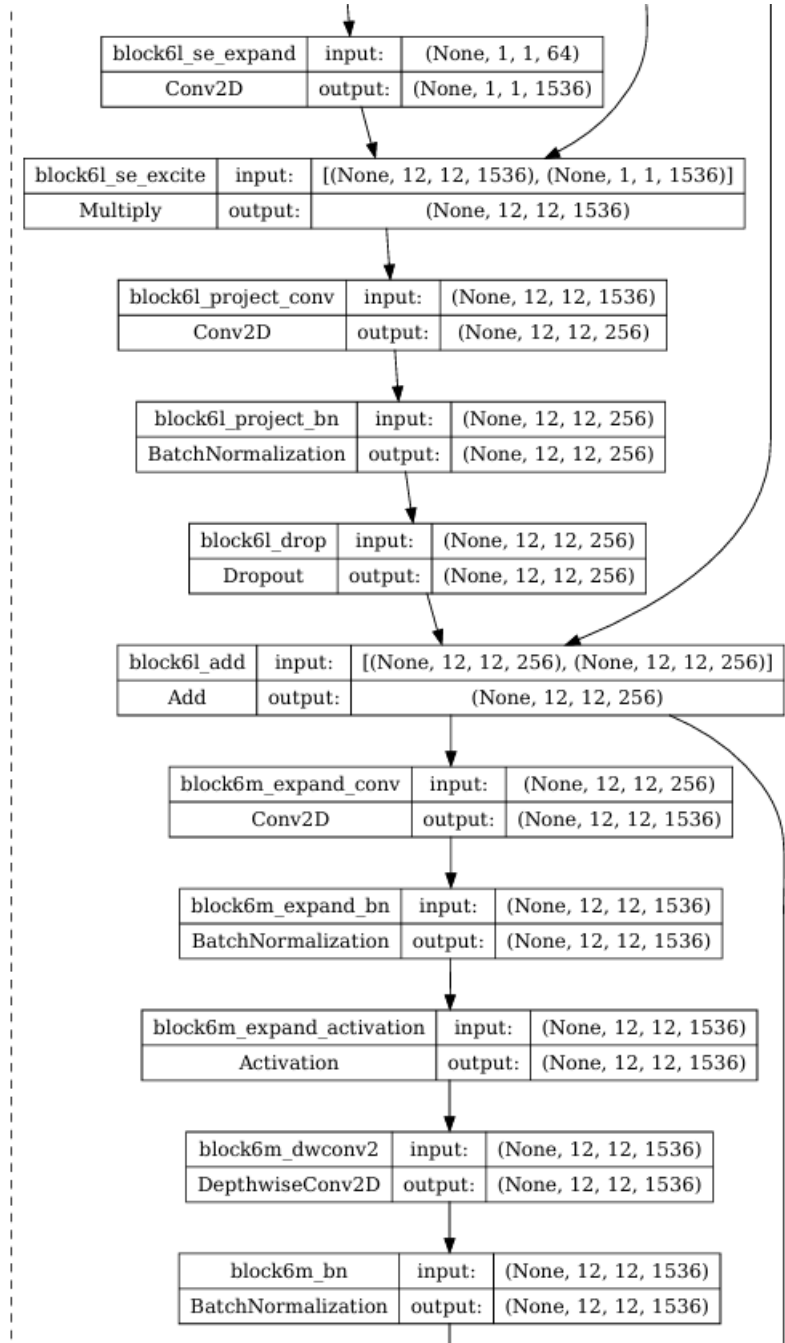
EK 1. (Devam) Geliştirilen Modelin Mimarisi



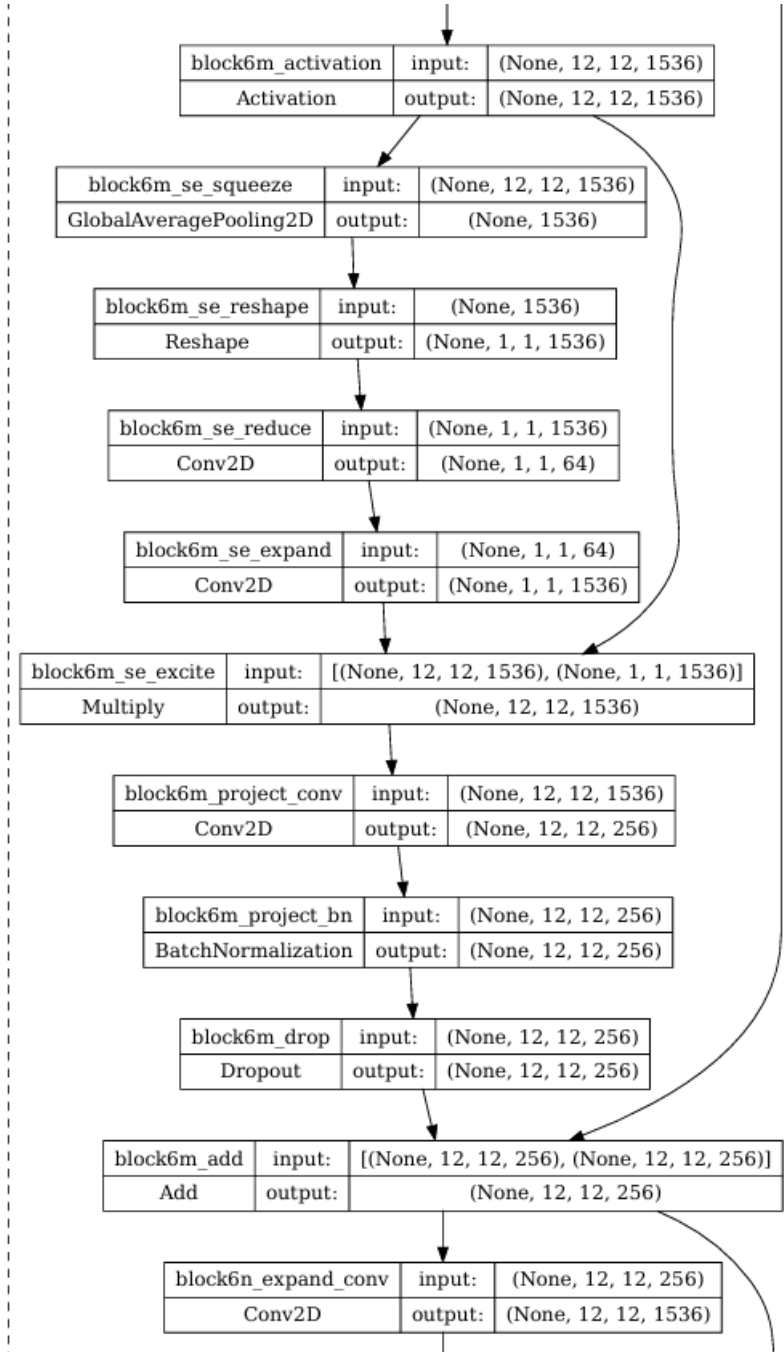
EK 1. (Devam) Geliştirilen Modelin Mimarisi



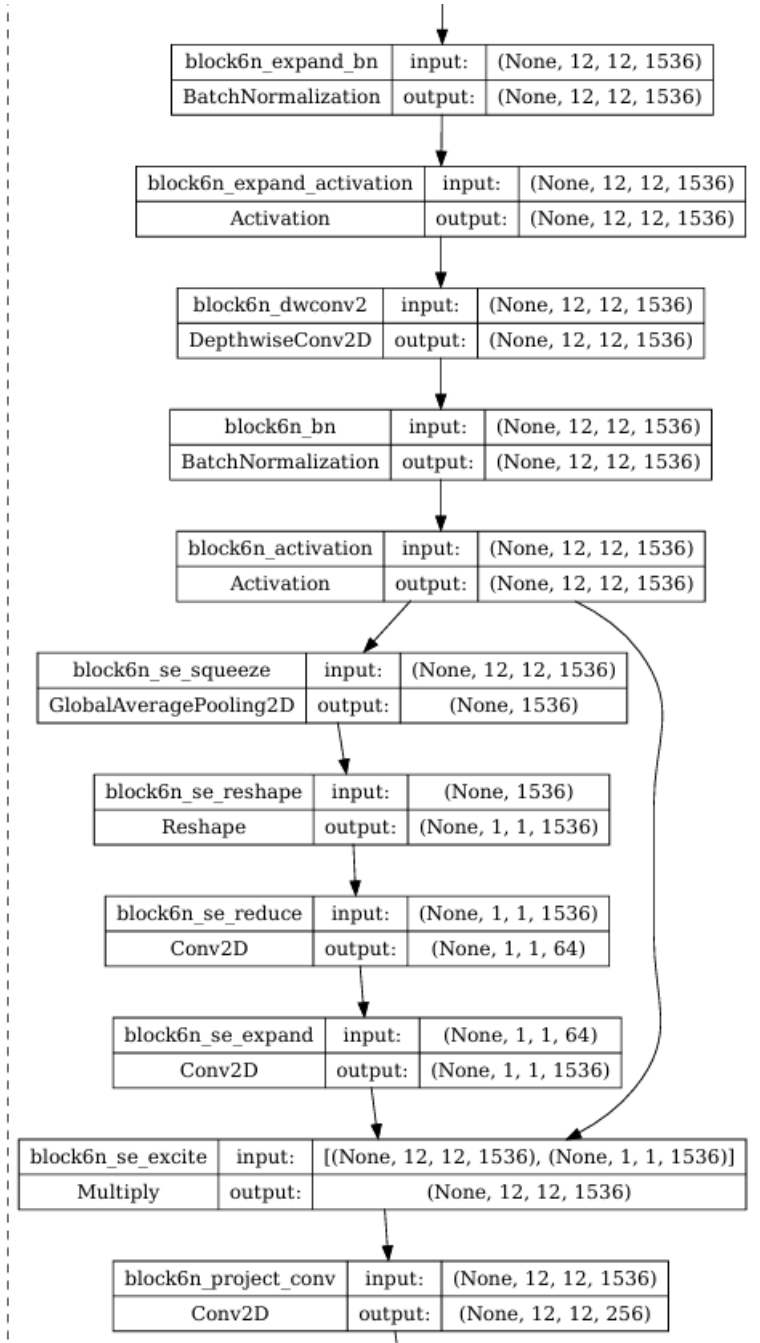
EK 1. (Devam) Geliştirilen Modelin Mimarisi



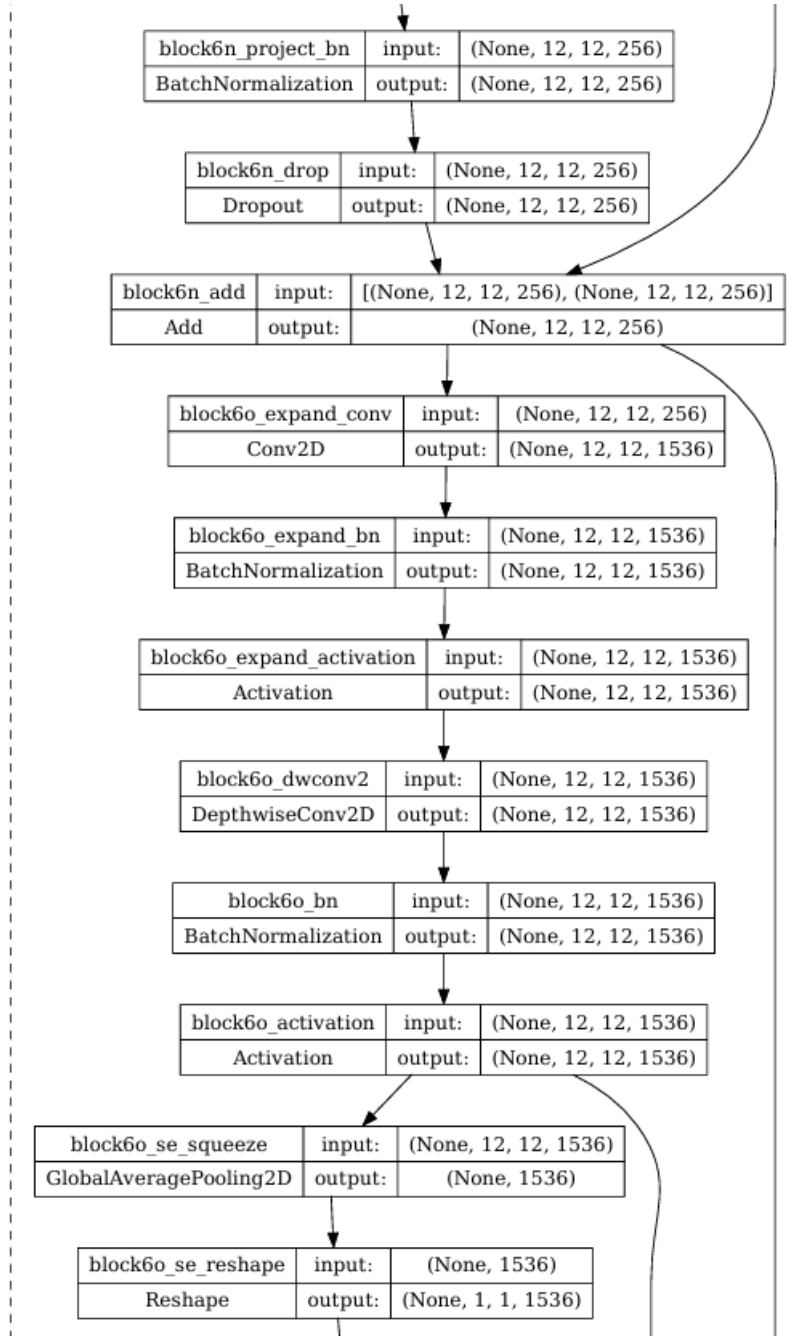
EK 1. (Devam) Geliştirilen Modelin Mimarisi



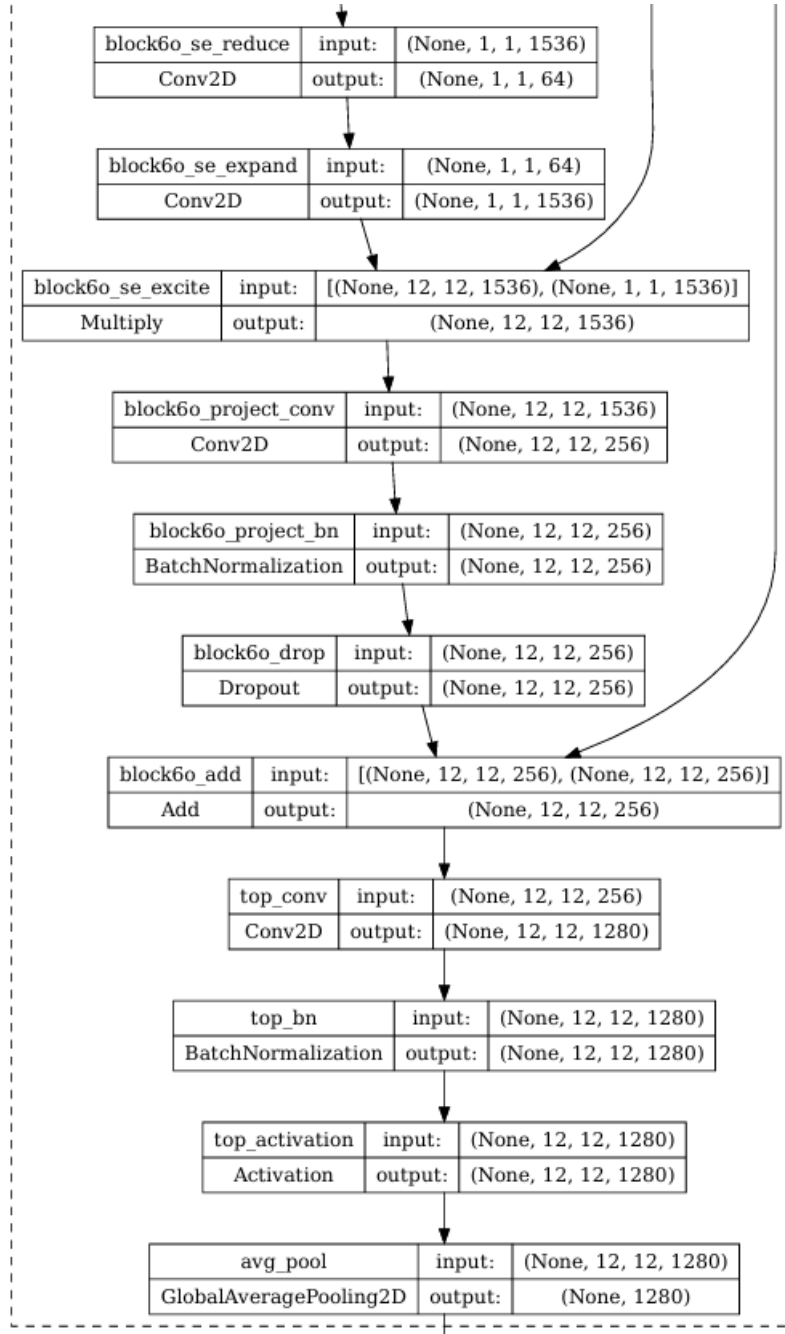
EK 1. (Devam) Geliştirilen Modelin Mimarisi



EK 1. (Devam) Geliştirilen Modelin Mimarisi



EK 1. (Devam) Geliştirilen Modelin Mimarisi



EK 1. (Devam) Geliştirilen Modelin Mimarisi

