

BURSA TEKNİK ÜNİVERSİTESİ ❖ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

**FİZİKSEL TABANLI İŞLEME MATERYALLERİ İÇİN KABARTI
DOKULARININ YAPAY SİNİR AĞLARI İLE SENTEZİ**



YÜKSEK LİSANS TEZİ

Muhammed Ömer Faruk SELVİ

Bilgisayar Mühendisliği Anabilim Dalı

HAZİRAN 2022

**T.C.
BURSA TEKNİK ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

**FİZİKSEL TABANLI İŞLEME MATERYALLERİ İÇİN
KABARTI DOKULARININ YAPAY SİNİR AĞLARI İLE
SENTEZİ**

YÜKSEK LİSANS TEZİ

**Muhammed Ömer Faruk SELVİ
(20376482019)**

Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı: Dr. Öğr. Üyesi Erdem YAVUZ

HAZİRAN 2022



20.04.2016 tarihli Resmi Gazete’de yayımlanan Lisansüstü Eğitim ve Öğretim Yönetmeliğinin 9/2 ve 22/2 maddeleri gereğince; Bu Lisansüstü teze, Bursa Teknik Üniversitesi’nin abonesi olduğu intihal yazılım programı kullanılarak Lisansüstü Eğitim Enstitüsü’nün belirlemiş olduğu ölçütlere uygun rapor alınmıştır.

İNTİHAL BEYANI

Bu tezde görsel, işitsel ve yazılı biçimde sunulan tüm bilgi ve sonuçların akademik ve etik kurallara uyularak tarafımdan elde edildiğini, tez içinde yer alan ancak bu çalışmaya özgü olmayan tüm sonuç ve bilgileri tezde kaynak göstererek belgelediğimi, aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiğimi beyan ederim.

Öğrencinin Adı Soyadı: Muhammed Ömer Faruk SELVİ

İmzası :





Aileme,

ÖNSÖZ

Başta bu tez çalışmamda olmak üzere; ihtiyaç duyduğum her konuda bana desteğini veren ve gittiğim yolda beni yönlendiren tez danışmanım Dr. Öğr. Üyesi Erdem YAVUZ hocama, bölüm başkanımız Prof. Dr. Turgay Tugay BİLGİN hocama ve tüm bölüm öğretim üyesi hocalarıma teşekkürlerimi sunarım.

Haziran 2022

Muhammed Ömer Faruk SELVİ
(Araştırma Görevlisi)

İÇİNDEKİLER

Sayfa

ÖNSÖZ.....	vi
İÇİNDEKİLER.....	vii
KISALTMALAR.....	ix
SEMBOLLER.....	x
ÇİZELGE LİSTESİ.....	xi
ŞEKİL LİSTESİ.....	xii
ÖZET.....	xiii
SUMMARY.....	xv
1. GİRİŞ.....	1
1.1 Tezin Amacı.....	4
1.2 Literatür Taraması.....	5
2. FİZİKSEL TABANLI GRAFİK İŞLEME VE MATERYAL DOKULARI...11	11
2.1 Fiziksel Tabanlı Grafik İşleme.....	11
2.2 Fiziksel Tabanlı Grafik İşlemede Kullanılan Materyal Dokuları.....	11
2.2.1 Renk dokusu.....	11
2.2.2 Kabartı dokuları (Bump map texture).....	11
2.2.2.1 Yükseklik dokusu (Height map texture).....	11
2.2.2.2 Normal dokusu (Normal map texture).....	12
2.2.3 Metaliklik dokusu (Metalness map texture).....	14
2.2.4 Pürüzlülük dokusu (Roughness map texture).....	14
3. YAPAY SİNİR AĞLARI (YSA).....15	15
3.1 Yapay Sinir Ağları'nın Aşamaları.....	16
3.1.1 İleri besleme.....	16
3.1.2 Geri yayılım.....	17
3.2 Yapay Sinir Ağları Modelleri.....	18
3.2.1 Ayırıcı modeller (Discriminators).....	18
3.2.2 Üretici modeller (Generators).....	19
3.3 Yapay Sinir Ağları ile Görüntü İşleme.....	19
3.3.1 Evrimsel sinir ağları (ESA, CNN).....	19
3.3.2 Evrimsel sinir ağları modelleri.....	20
3.3.2.1 LeNet.....	20
3.3.2.2 AlexNet.....	20
3.3.2.3 VGG Net.....	21
3.3.2.4 ResNet.....	21
3.3.2.5 Inception Net (GoogLeNet).....	22
3.3.2.6 DenseNet.....	22
3.3.2.7 Otokodlayıcılar (Autoencoders).....	22
4. MATERYAL VE YÖNTEM.....23	23

4.1 Veri Seti Büyütme (Data Augmentation).....	23
4.1.1 Görüntü veri setlerinin büyütülmesi.....	23
4.1.2 Döndürme.....	24
4.1.3 Bükme (Shearing).....	24
4.1.4 Yakınlaştırma.....	25
4.1.5 Yatay çevirme.....	25
4.1.6 Dikey çevirme.....	25
4.1.7 Doldurma modu.....	25
4.1.8 Parlaklık aralığı.....	25
4.2 Deneysel Çalışmalar.....	26
4.2.1 Standart otokodlayıcı.....	28
4.2.2 Kimlik eşlemeli ResNet tipi otokodlayıcı.....	31
4.2.3 Evrişim katmanlı ResNet tipi otokodlayıcı.....	34
5. BULGULAR VE TARTIŞMA.....	36
5.1 Standart Otokodlayıcı Modellere Ait Sonuçlar.....	36
5.2 Kimlik Bağlantılı ResNet Tipi Otokodlayıcı Modellere Ait Sonuçlar.....	40
5.3 Evrişim Bağlantılı ResNet Tipi Otokodlayıcı Modellere Ait Sonuçlar.....	45
5.4 Genel Değerlendirme.....	49
6. SONUÇLAR VE ÖNERİLER.....	50
KAYNAKLAR	60
ÖZGEÇMİŞ	63

KISALTMALAR

2B	: 2 Boyutlu
3B	: 3 Boyutlu
AC	: Evrişim Bağlantılı Otokodlayıcı (Autoencoder with Convolution Block)
AE	: Otokodlayıcı (Autoencoder)
AI	: Kimlik Bağlantılı Otokodlayıcı (Autoencoder with Identity Block)
CNN	: Evrişimsel Sinir Ağları (Convolutional Neural Network)
PS	: Fotometrik Stereo (Photometric Stereo)
DPSN	: Derin Fotometrik Stereo Ağı (Deep Photometric Stereo Network)
ESA	: Evrişimli Sinir Ağları
FTİ	: Fiziksel Tabanlı İşleme
GAN	: Çekişmeli Üretici Ağlar (Generative Adversarial Network)
GNN	: Geometrik Sinir Ağları (Geometric Neural Network)
LSTM	: Uzun Kısa Süreli Bellek (Long Short-Term Memory)
MAE	: Ortalama Açısal Hata (Mean Angular Error)
MSE	: Ortalama Karesel Hata (Mean Squared Error)
FCN	: Tam Bağlantılı Ağ (Fully Connected Network)
PSNR	: Tepe Sinyal Gürültü Oranı (Peak Signal-to-Noise Ratio)
ReLU	: Düzeltilmiş Doğrusal Birim (Rectified Linear Unit)
ResNet	: Kalıntılı Sinir Ağları (Residual Neural Network)
RGB	: Kırmızı-Yeşil-Mavi (Red-Green-Blue)
RNN	: Yinelemeli Sinir Ağları (Recurrent Neural Network)
Tanh	: Hiperbolik Tanjant
YSA	: Yapay Sinir Ağları

SEMBOLLER

α	: Öğrenme katsayısı
ω	: Ağırlık
L	: Kayıp fonksiyonu
x	: Girdi verileri
y	: Çıktı verileri



ÇİZELGE LİSTESİ

Sayfa

Çizelge 4.1 : Farklı model tipleri için deneylenecek parametreler ve değerleri.....	27
Çizelge 4.2 : Deney kodlarına göre modellerin üretiminde kullanılan parametreler.	27
Çizelge 5.1 : Standart otokodlayıcı tipindeki modellerin eğitim sonuçları.....	37
Çizelge 5.2 : Kimlik bağlantılı ResNet tipi otokodlayıcı modellerin eğitim sonuçları.	41
Çizelge 5.3 : Evrişim bağlantılı ResNet tipi otokodlayıcı modellerin eğitim sonuçları.	46

ŞEKİL LİSTESİ

Sayfa

Şekil 1.1 : Santo ve diğ. (2017) tarafından sunulan yöntemde kullanılan yapay sinir ağı modelinin yapısı.....	6
Şekil 1.2 : DPSN modelinin ürettiği yüzey normalleri (Santo ve diğ, 2017).....	6
Şekil 1.3 : Görüntü girdisinden materyal parametrelerini ve dokularını tahminleyen ağların yapıları (Li ve diğ, 2017).....	7
Şekil 1.4 : Yüzey normallerinin derinlik bilgisiyle iyileştirilmesi aşamalarını gösteren akış (Qi ve diğ, 2018).....	8
Şekil 1.5 : PS-FCN modeline ait ağ yapısı (Chen ve diğ, 2018).....	8
Şekil 1.6 : Deschaintre (2019) tarafından geliştirilen modelin örnek çıktıları.....	9
Şekil 1.7 : Song ve Chang (2020) tarafından geliştirilen modelin yapısı.....	10
Şekil 2.1 : a) yükseklik dokusu, b) yükseklik dokusu kullanılarak işlenmiş 3B yüzey (Zhou ve diğ, 2017).....	12
Şekil 3.1 : Sıklıkla kullanılan aktivasyon fonksiyonlarından bazılarının kıyaslanması.....	17
Şekil 3.2 : LeNet modeline ait ağ yapısı (LeCun ve diğ, 1998).....	20
Şekil 3.3 : ResNet ağlarının yapı taşı olan bağlantı (He ve diğ, 2016).....	21
Şekil 4.1 : Birinci parametresi 2, ikinci parametresi 8 olan standart otokodlayıcı modeli.....	29
Şekil 4.2 : İlk parametresi 3, ikinci parametresi 8 olan standart otokodlayıcı tipindeki model.....	30
Şekil 4.3 : İlk parametresi 2, ikinci parametresi 16 olan standart otokodlayıcı modeli.....	31
Şekil 4.4 : Katman parametresi 2, filtre parametresi 8 olan modelin mimarisi.....	32
Şekil 4.5 : Katman parametresi 2, filtre parametresi 16 olan modelin mimarisi.....	33
Şekil 4.6 : Katman parametresi 3, filtre parametresi 8 olan modelin mimarisi.....	34
Şekil 5.1 : AE-C4 modelinin eğitim-mse(kayıp) grafiği.....	37
Şekil 5.2 : AE-C4 modeline ait eğitim-PSNR grafiği.....	38
Şekil 5.3 : AE-B4 modeline ait eğitim-mse(kayıp) grafiği.....	38
Şekil 5.4 : AE-B4 modeline ait eğitim-PSNR grafiği.....	39
Şekil 5.5 : AE-D1 modeline ait eğitim-mse(kayıp) grafiği.....	39
Şekil 5.6 : AE-D1 modeline ait eğitim-PSNR grafiği.....	40
Şekil 5.7 : AI-A1 deneyine ait modelin eğitim-mse(kayıp) grafiği.....	42
Şekil 5.8 : AI-A1 deneyine ait modelin eğitim-PSNR grafiği.....	42
Şekil 5.9 : AI-B4 modeline ait eğitim-mse(kayıp) grafiği.....	43
Şekil 5.10 : AI-B4 modeline ait eğitim-PSNR grafiği.....	44
Şekil 5.11 : AI-D4 deneyine ait modelin eğitim-mse(kayıp) grafiği.....	44
Şekil 5.12 : AI-D4 deneyine ait modelin eğitim-PSNR grafiği.....	45
Şekil 5.13 : AC-A1 modeline ait eğitim-mse(kayıp) grafiği.....	47

Şekil 5.14 : AC-A1 modeline ait eğitim-PSNR grafiği.....	47
Şekil 5.15 : AC-B4 modeline ait eğitim-mse(kayıp) grafiği.....	48
Şekil 5.16 : AC-B4 modeline ait eğitim-PSNR grafiği.....	48
Şekil 6.1 : En iyi standart otokodlayıcı modelin test seti örneklerinden bazıları için verdiği çıktılar.....	52
Şekil 6.2 : En iyi standart otokodlayıcı modelin test seti dışından rastgele görüntüler için verdiği çıktılar.....	53
Şekil 6.3 : En iyi kimlik bağlantılı ResNet tipi otokodlayıcı modelin test seti örneklerinden bazıları için verdiği çıktılar.....	54
Şekil 6.4 : En iyi kimlik bağlantılı ResNet tipi otokodlayıcı modelin test seti dışından rastgele görüntüler için verdiği çıktılar.....	55
Şekil 6.5 : En iyi evrişim katmanlı ResNet tipi otokodlayıcı modelin test seti örneklerinden bazıları için verdiği çıktılar.....	56
Şekil 6.6 : En iyi evrişim katmanlı ResNet tipi otokodlayıcı modelin test seti dışından rastgele görüntüler için verdiği çıktılar.....	57



FİZİKSEL TABANLI İŞLEME MATERYALLERİ İÇİN KABARTI DOKULARININ YAPAY SINIR AĞLARI İLE SENTEZİ

ÖZET

Prosedürel İçerik Üretimi (PIÜ), on yıllardır üzerinde çalışılan bir alandır. PIÜ, dijital ürünlerin maliyetlerini düşürmeyi mümkün kılabilir. Birçok dijital içerik, yaygın olarak kullanılan veya çok benzer olan varlıkları içerir. Bir varlık üretimi için uygun bir PIÜ fonksiyonu bulmak, üretim için gereken zamanı ve maliyeti azaltacaktır. Öte yandan, dijital varlıklar oluşturmak için işe yarar fonksiyonlar hazırlamak zordur. Bu tür fonksiyonlar, makine öğrenimi yöntemleri kullanılarak tahminlenebilir. Yeterli miktarda veri verildiğinde, belirli bir varlık türü üretmek için üretken bir yapay zeka modeli eğitilebilmektedir.

Grafik İşlem Birimlerinde (GPU) yapılan iyileştirmeler sayesinde günümüzde yapay zeka modelleri oluşturmak daha kolay hale gelmiştir. Özellikle Yapay Sinir Ağları (YSA) popüler hale gelmiş ve yıllardır aktif olarak üzerinde çalışılmaktadır. Sınıflandırabilen veya örnek üretebilen modellerin eğitebilmesi, YSA'nın özelliklerindedir. Otokodlayıcılar ve Çekişmeli Üretici Ağlar (ÇÜA), üretken modellere örnektir. Üretken modelleri kullanarak, stilleri transfer etmek veya diğer karmaşık işlevleri tahminlemek de mümkündür.

Fiziksel Tabanlı İşleme (FTİ) yöntemi, gerçekçi 3B modelleri ve sahneleri gerçek zamanlı olarak oluşturmanın bir yoludur. İyi sonuçlar elde edebilmek için çeşitli doku türlerinin kullanılması gerekmektedir. Kabartı dokuları, 3B model materyallerinde kullanılan doku türlerinden biridir. Yüzey normallerini ve yüksekliklerini simüle etmek için kullanılmaktadır. Normal dokuları, yüzey normal vektörlerini barındıran kabartı dokularıdır. Normal dokuları, otometrik stereo teknikleri kullanılarak veya yüksek detaylı modelin düşük detaylı versiyonuna yansıtılarak üretilebilmektedir. Söz konusu yöntemler zaman ve maliyet açısından pahalıdır. YSA ve PIÜ'yü birleştirmeye odaklanan birçok çalışma bulunmaktadır. Her ikisini de kullanarak, maliyetleri düşürürken makul kalitede malzeme dokuları üretmek mümkün olabilmektedir.

Bu tezde, birkaç üretici yapay sinir ağı modeli eğitilmiştir. Sonuçları, ağ karmaşıklığı ile birbirleriyle karşılaştırılmıştır. Artan katman ve filtre sayılarının her zaman çıktı kalitesini yükseltmediği tespit edilmiştir. Karmaşık modellerde, kaybolan gradyan sorunu gözlemlenmiştir. Kaybolan gradyan problemi bir noktadan sonra modelin daha fazla gelişmemesine ve eğitime devam etmenin faydasız olmasına sebep olmaktadır. Kalıntılı Sinir Ağları (ResNet) mimarisine sahip modellerin bu sorunu çözebileceği görülmüştür. Nihai sonuçlar ayrıca, veri artırmanın eğitim aşamasında büyük bir rol oynadığını göstermiştir.

Anahtar kelimeler: Derin Öğrenme, Evrimsel Sinir Ağları, Prosedürel İçerik Üretimi, Fiziksel Tabanlı İşleme, Doku Sentezi

SYNTHESIS OF NORMAL MAP TEXTURES FOR PHYSICALLY BASED RENDERING MATERIALS WITH ARTIFICIAL NEURAL NETWORKS

SUMMARY

Procedural Content Generation (PCG) is a field that has been studied on for decades. PCG may make it possible to lower costs of digital products. Many digital contents include commonly-used or very similar assets. Finding a proper function for PCG would decrease time and money required for the production. On the other hand, preparing useful functions to generate digital assets is difficult. This kind of functions can be approximated using machine learning methods. Given a sufficient amount of data, a generative artificial intelligence model can be trained to produce a specific type of asset.

Thank to the improvements on Graphics Processing Units (GPU), it is easier nowadays to build artificial intelligence models. Especially, Artificial Neural Networks (ANN) have become popular and been worked on actively for years. Training models that can classify or produce samples is of the capabilities of ANN's. Autoencoders and Generative Adversarial Networks (GAN) are examples of generative models. Using generative models, it is also possible to transfer styles or approximate other complex functions.

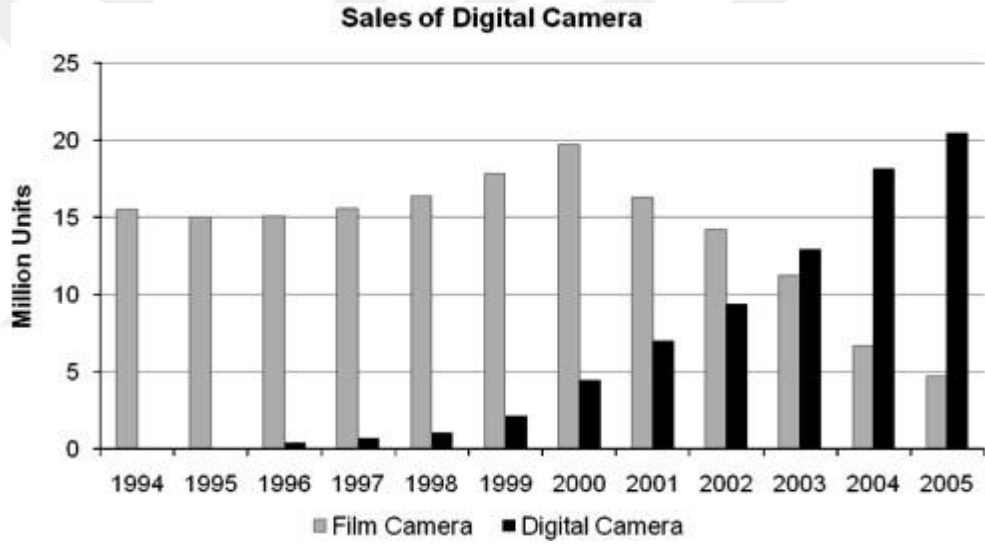
Physically Based Rendering (PBR) method is a way to render realistic 3D models and scenes in realtime. For good results, various type of textures are needed. Bump maps are a type of texture for 3D materials. They are used to simulate surface normals and heights. Normal maps are bump maps that store normal vectors of surface. They can be produced using photometric stereo techniques or projecting a high-poly model onto its low-poly version. These methods are time consuming and expensive. There are many studies that focus on combining ANN and PCG. Using both, it would be possible to produce material textures in reasonable quality while reducing the costs.

In this thesis, several generative artificial neural network models were trained. Their results were compared to eachother by the network complexity. It was determined that increasing layer and filter numbers doesn't always higher the output quality. The complex models caused vanishing gradient problem. Vanishing gradient problem, at some point, stops the model from getting more improved and makes it ineffective to continue the training phase. It was understood that the models with the architecture of Residual Neural Networks (ResNet) can handle this problem. Final results also showed that data augmentation plays a big role on the training phase.

Keywords: Deep Learning, Convolutional Neural Networks, Procedural Content Generation, Physically Based Shading, Texture Synthesis

1. GİRİŞ

Teknoloji geliştikçe üretim araçları da ucuzlamaktadır. Özellikle dijital ürün üretebilmek için gereken girdiler son yıllarda daha çok insan tarafından kullanılabilir hale gelmiştir. Giderek ucuzlayan ve gelişen teknolojiye ayak uyduramayan büyük firmalar ciddi zararlar etmektedir. Dijital fotoğrafçılığın yaygınlaşmaya başladığı ilk yıllarda bu yeni teknolojiye henüz adapte olamamış analog fotoğraf makinesi üreten köklü firmaların hisselerinde ciddi düşüşler gözlemlenmiştir (Lucas Jr ve diğ, 2009).



Şekil 1.1 : Analog ve Dijital Kamera Satışlarının Karşılaştırılması (Lucas Jr ve Goh, 2009).

Yapay Sinir Ağları'nın üretici yapay zeka modellerinde kullanımı giderek artmaktadır. Seslendirme, fotoğraf düzenleme, video montaj vb pek çok alanda iş gücü gereksinimlerini düşürmek amacıyla derin öğrenme yöntemleri kullanılmaktadır.

Derin öğrenme yöntemleri eğitim ve araştırma aşamasında yüklü bir maliyete sahip olsa da, yüksek tutarlılıkta çıktı verebilen ve ezber yapmamış olan modeller sağladıkları avantajlar sayesinde maliyetlerinden daha çok fayda sağlayabilmektedir.

Son zamanlarda üretilen pek çok otomobil, derin öğrenme yöntemleri ile eğitilmiş modelleri içerisinde barındıran cihazlara sahiptir. Pek çok sürücü yolculuk boyunca gideceği yolu bilmesine rağmen en az yoğunluktaki yolları görebilmek için dahi yapay zeka modellerinin saptamalarından faydalanmaktadır. Giderek kullanım alanları artan yapay zeka uygulamaları sadece tespit sistemlerinde değil aynı zamanda dijital içerik üretim sistemlerinde de kullanılmaktadır.

Üretici modellerin üretilbildiği derin öğrenme yöntemleriyle ses, görüntü gibi medya verileri sentezlenebilmektedir. Özellikle üç boyutlu model ve çevre, doku ve desen tasarımında tasarımcıların işlerini kolaylaştırabilmektedir. Kolaylık sağladığı kadar aynı zamanda tasarımcıların sadece anlatmak istediklerine odaklanabilmelerine de fırsat vermektedir.

2018 yılında geliştirilen Deepmimic adlı proje ile insan hareketleri simüle edilebilmiştir (Peng ve diğ., 2018). Pekiştirmeli öğrenme metodunun uygulandığı model üç boyutlu karakterler için gerçekçi animasyonlar üretebilmiştir. Fizik tabanlı animasyon üretebilen bu gibi yapay zeka modelleriyle, oldukça pahalı olan hareket yakalama (motion capture) teknikleri ucuzlatılabilmekte ve üretim süreci kısaltılabilmektedir.

Resimden-resime (pix2pix) dönüşüm yapan modeller de pek çok kompleks işlemi kolay hale getirebilmektedir. Su ve diğ. (2018) tarafından yapılan bir çalışma, eskiz çizimlerinden nesneyi algılayabilmekte ve nesneye ait yüzey normallerini hesaplayabilmektedir. Özellikle iki boyutlu (2B) grafiklerin farklı ışık kaynakları altında tekrar işlenmesinin gerektiği durumlarda bu yüzey normalleri kullanılabilir. Bu gibi çalışmalar, ışığın vurma açısının dinamik olarak değiştiği 2B video oyunlar için karakter üretiminde süreci hızlandırabilmektedir.

Pek çok çalışmada prosedürel içerik üretim (procedural content generation) teknikleri makine öğrenmesi ve derin öğrenme metotlarıyla birleştirilmiştir. Summerville ve Mateas (2016) bir Uzun Kısa Süreli Bellek Ağı (LSTM) modeli geliştirmişlerdir. Bu LSTM modeli ile bir platform oyunu olan Mario için bölümler üretmişlerdir. GAN yapısına sahip bir ağ ile de video oyunları için bölüm tasarımı yapılabilmektedir. Volz ve diğ. (2018) Mario oyunu için bölümler geliştiren Çekişmeli Üretici Ağ

(GAN) modeli eğitmişlerdir. Giacomello ve diğ. (2018) yaptıkları çalışmada eğittikleri GAN modeli ile Doom adlı oyun için bölümler üretmişlerdir. Torrado ve diğ. (2020) yeniden örnekleme yöntemini bölüm üretici GAN modelleri için kullanmışlardır.

Üç boyutlu grafiklerin tasarımında gerçekçi sonuçların elde edilebilmesini sağlayan en önemli etkenlerden biri materyallerdir. Materyaller; renk, kabartı/yükseklik, metaliklik, pürüzlülük gibi çeşitli özellikleri barındırabilmektedir. Her özellik için birer doku görüntüsü kullanılabilir.

Bir kabartı dokusu, bir modelin materyaline ait renk dokusu ile aynı şekilde yüzey üzerine kaplanmaktadır. Bu kaplama işlemine UV eşleme (UV mapping) adı verilir. X, Y ve Z karakterleri üç boyutlu uzaydaki eksenleri belirtmek için, W ise üç boyutlu nesnelerin döndürülmesinde kullanılan dördeylerde (quaternions) 4. eksenini belirtmek için kullanıldığından bu üç boyutlu uzaydaki nesnelerin iki boyutlu kaplamalarındaki koordinatları belirtmek amacıyla farklı karakterlerin kullanılmasına ihtiyaç duyulmuştur. Bu sebeple U ve V karakterleri tercih edilmiştir. Yöntemin ismindeki UV bu eksen harflerinden gelmektedir.

Blinn (1978) ve Woodham (1980) tarafından yapılan çalışmalarla 3B yüzeylerin üzerindeki noktaların yön vektörlerini içeren kabartı dokularının (normal map texture) üç boyutlu grafik işlemede kullanılması mümkün hale gelmiştir. Günümüze kadar olan süreçte kabartı dokuları genellikle iki yöntemle üretilmiştir; yüksek kaliteli üç boyutlu modelden (high-poly) düşük kaliteli üç boyutlu modele (low-poly) yüzey bilgilerini taşıma veya fotometrik metot.

İlk yöntemde, optimize edilmek istenen yüksek poligon içeren üç boyutlu modelin düşük poligon sayılı bir versiyonu üretilir. Ardından bu iki model iç içe geçirilerek yüksek poligonlu modelden düşük poligonlu modelin iki boyutlu kaplamasına denk gelen UV koordinatına yüzey normal değerleri transfer edilir. Kabartı dokuları üç kanallı görüntülerdir ve RGB formatında kaydedilir. Her bir renk kanalı bir eksen ifade eder. Bir pikseldeki renk değerleri birlikteyken bir vektör oluşturur. Bu vektör, pikselin UV eşlemesine göre denk geldiği üç boyutlu yüzeydeki noktanın normal vektörüdür. Üç boyutlu sahnenin işlenmesi esnasında (render) bu vektörler yüzeyin

ışığı nasıl yansıtacağını belirtir. Daha düşük poligon sayısı ile, yüksek poligon ile olduğundakiyle yaklaşık olarak aynı kalitede sonuç elde edilebilmesi ve düşük işlem gücü gereksinimi sayesinde gerçek zamanlı görüntü işlemede (realtime rendering) kullanılabilmesi sebebiyle kabartı dokuları üç boyutlu tasarım, oyun grafikleri, animasyonlar ve hatta filmlerde bir gereklilik haline gelmiştir.

Fotometrik metotta ise kabartı bilgileri alınmak istenen bir yüzeye farklı açılardan ışık yansıtılır. Her farklı ışık için birer görüntü alınır. Daha sonra bu görüntüler işlenerek yüzeyin kabartı dokusu çıkarılır. Günümüzde yüksek detaylı ve çözünürlüklü kaplamaların üretiminde kullanılmaktadır.

Hala kullanımda olan bu yöntemler zaman ve işlem gücü açısından maliyetlidir. Daha önce üretilmemiş benzersiz bir dokuya ihtiyaç duyulma ihtimali çok düşüktür. Her tasarım sürecinde taş, bitki, ahşap, metal, beton veya insan derisi gibi sıklıkla kullanılan yüzeylerin dokularının kaplamalarını tekrar tasarlamak yerine otomatize edilmiş bir sistemden elde etmek daha avantajlı olacaktır. Günümüzde özellikle prosedürel üretim teknikleriyle bu gibi sıkça ihtiyaç duyulan nesnelerin üretimi, parametrelendirilmiş fonksiyonlarla yapılabilmektedir.

Prosedürel üretim teknikleri avantajlı olsa da her çeşit ürün için farklı bir prosedür/fonksiyon yazmak gerekmektedir. Kabartı dokularının prosedürel olarak üretimleri kaliteyi ve çeşitliliği kısıtlamaktadır.

Yapay sinir ağları elle öznitelik çıkarımına ihtiyaç duymamaktadır. Yapay sinir ağlarının doğası ve üç boyutlu nesnelerin materyal dokularının sonsuz varyasyona sahip olabilmesi, bu türdeki dokuların üretiminin otomatize edilebilmesi için en iyi yöntemlerden birinin derin öğrenme olduğu anlamına gelmektedir.

1.1 Tezin Amacı

Bu tez geliştirmekte olan yapay sinir ağlarının üç boyutlu grafik teknolojilerindeki kullanımını üzerinedir. Zaman ve maliyet açısından pahalı olan grafik materyallerinin yapay zeka modelleriyle ne kadar yakınsanabileceği incelenmektedir. 3B grafiklerin işlenmesi esnasında kullanılan kabartı dokularının renk dokularından

sentezlenebilmesi, işlem maliyetlerini düşürecektir. Fakat bu gibi dokular kolayca hesaplanamamaktadır.

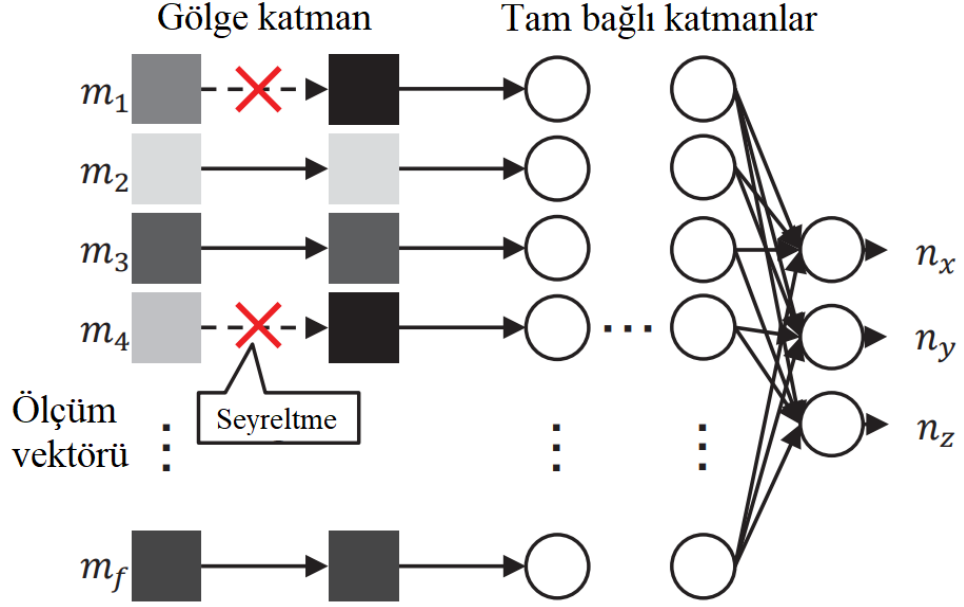
Kabartı dokuları gibi materyal parametrelerinin yakalanmasındaki (material capture) genel problemler; ışık kaynağının yönü, ışığın şiddeti, görüntülenen nesnenin konkav veya konveks olabilmesi gibi sebeplerden meydana gelmektedir. Bu yüzden sorunsuz materyal yakalama yöntemi henüz bulunmamaktadır. Yapılan çalışmalarla söz konusu problemlerin etkileri giderilmeye çalışılmıştır.

Yapay sinir ağları ile geliştirilen üretici modellerin ortak sorunlarından birisi ise kaybolan gradyan problemidir. Ağın daha kompleks hale getirilmesi daha verimsiz eğitilmiş modeller ortaya çıkarmaktadır. Bu tez çalışmasında, renk dokularından kabartı dokularının tahminlenmesi için üretilebilecek çeşitli tiplerde modeller farklı parametrelerle ele alınacaktır. Doku tahmini için en uygun model yapısı ve parametreleri tespit edilmeye çalışılacaktır. Kaybolan gradyan problemi için üretilen çözümler kabartı dokularının üretimi için tasarlanan ağlarda sınanacaktır.

1.2 Literatür Taraması

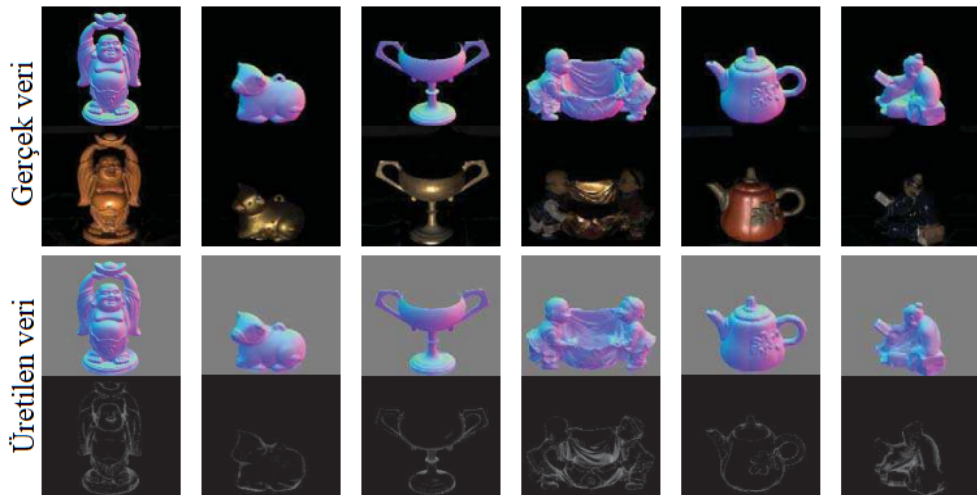
Parametrik doku üretimini daha kompleks durumlar için genişletmeyi amaçlayan bir çalışmada, tek görüntüden çeşitli fiziksel tabanlı işleme materyali dokuları üretilebilmiştir (Aittala ve diğ, 2016). Oluşturulan model ile materyallerin tahminlenen renk, parlaklık ve normal dokuları ile yapılan yeniden işleme (re-rendering) sonucunda elde edilen görüntü, ham görüntüye oldukça yakın olmaktadır.

2017'de yapılan bir çalışmada derin öğrenme tabanlı fotometrik stereo yöntemi sunulmuştur (Santo ve diğ, 2017). Geliştirilen yapay sinir ağı ile çeşitli açılardan ışıklandırılmış nesnelerin yüzeylerinin normal vektörleri hesaplanmıştır. İlgili modele ilişkin ağ yapısı Şekil 1.1'de verilmiştir.



Şekil 1.1 : Santo ve diğ. (2017) tarafından sunulan yöntemde kullanılan yapay sinir ağı modelinin yapısı.

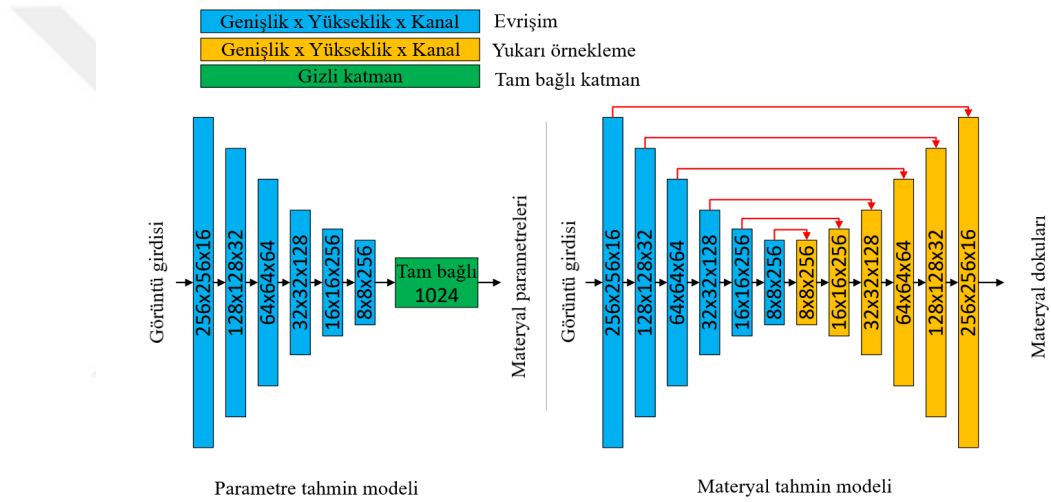
Modelin üretiminde Kabarcık Şekli Veri Seti (Blobby Shape Dataset) içerisinde bulunan 10 adet 3B modelden 8'i kullanılmıştır. Şekil veri setinden alınan 3B modeller çeşitli ışıklandırma koşullarında işlenmiştir ve yapay sinir ağının eğitimi için gerekli veriler üretilmiştir. Derin Fotometrik Stereo Ağı (DPSN) adı verilen model toplamda 76800 görüntü ile, her döngüde 1000 görüntü olacak şekilde 5000 döngü boyunca eğitilmiştir. Çalışmaya ait örnek çıktılar Şekil 1.2'de görülmektedir.



Şekil 1.2 : DPSN modelinin ürettiği yüzey normalleri (Santo ve diğ, 2017).

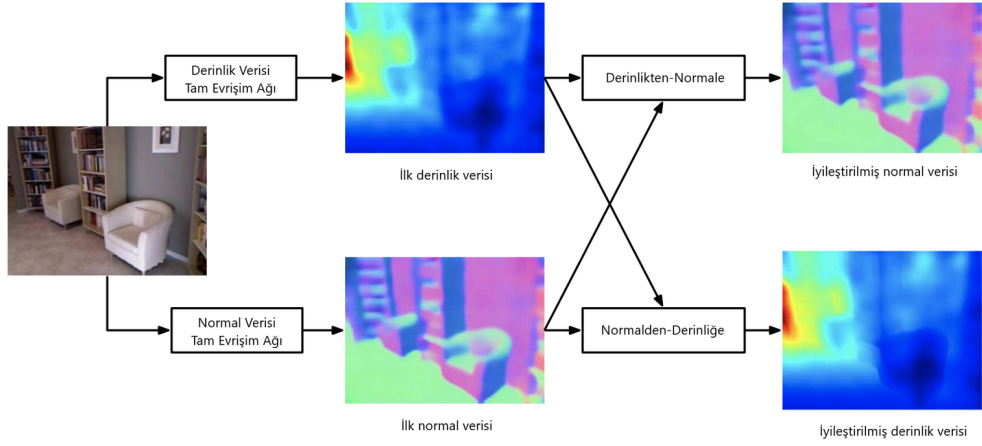
Li ve diğ. (2017) tarafından yapılan başka bir çalışmada yüzey materyallerini yakınsayan Evrişimsel Sinir Ağı geliştirilmiştir. Üretilen model bilinmeyen doğal ışık koşulları altında bulunan bir nesnenin çekilmiş tek fotoğrafını materyal bilgilerine çevirebilmektedir.

Materyal dokularını tahminleyen model bir çeşit otokodlayıcıdır. Standart bir otokodlayıcının kodlayıcı ve çözücü katmanları arasında bağlar kurulmuştur. Parametre tahmin modeli ise bir çeşit sınıflandırıcı ağıdır. Modelin son kısmında 1024 nörondan oluşan bir tam bağlı katman, girdi katmanından verilen görüntüden elde edilen materyal parametreleri tahminlerini çıkılmaktadır. Modellerin yapıları Şekil 1.3'te verilmiştir.



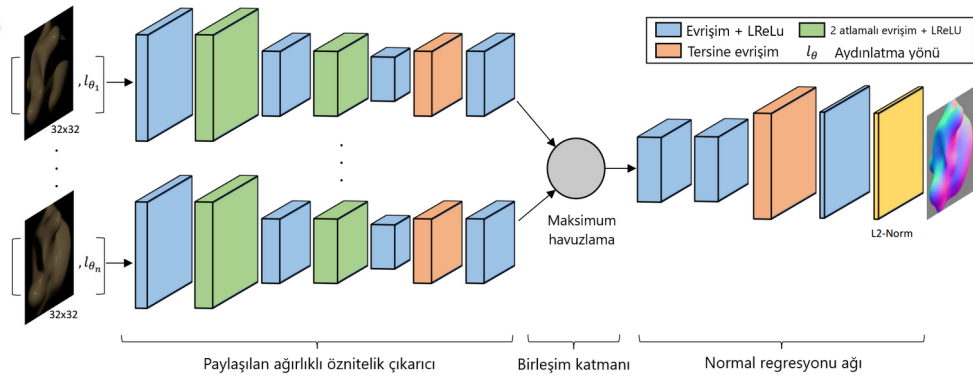
Şekil 1.3 : Görüntü girdisinden materyal parametrelerini ve dokularını tahminleyen ağların yapıları (Li ve diğ, 2017).

Qi ve diğ. (2018) yaptıkları çalışmada RGB görüntüden yüzey normallerini ve derinlik haritasını çıkaran Geometrik Sinir Ağları'nı (GNN) geliştirmişlerdir. Ağ yapısının temelini VGG-16 ve ResNet-50 modelleri oluşturmaktadır. Modelin derinlikten-normale ve normalden-derinliğe dönüşüm yapan kısımları bulunmaktadır. Algılanan yüzey normalleri, ağın derinlikten-normale dönüşüm yapan kısmı ile düzeltilmekte ve daha kaliteli çıktılar alınması sağlanmaktadır. Veri iyileştirme sistemi Şekil 1.4'te görüldüğü gibidir.



Şekil 1.4 : Yüzey normallerinin derinlik bilgisiyle iyileştirilmesi aşamalarını gösteren akış (Qi ve diğ., 2018).

Chen ve diğ. (2018) fotometrik stereoyu öğrenen bir ağ modeli geliştirmişlerdir. Esnek tam evrişimsel ağ (PS-FCN) olarak ifade ettikleri bu ağ, eğitim aşamasında ön tanımlı ışık vurma açılarında ihtiyaç duymamaktadır. Model birden fazla görüntüyü ve ışık yönünü sıraları farketmeksizin analiz edip çıkarımda bulunabilmektedir. Modele ait ağın katmanları Şekil 1.5'te görülmektedir.

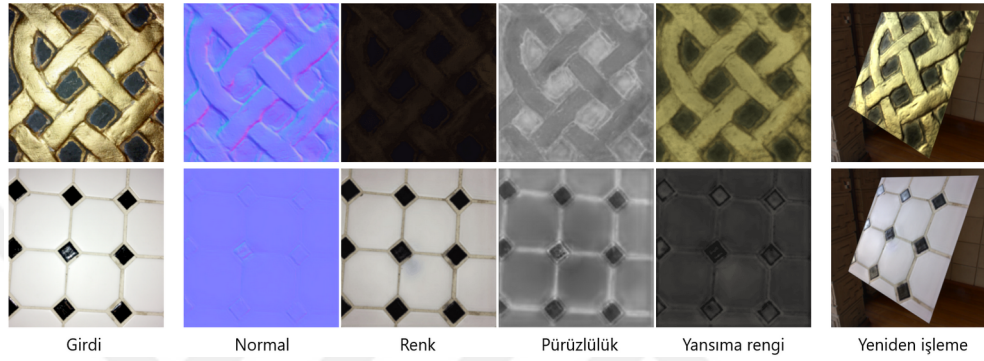


Şekil 1.5 : PS-FCN modeline ait ağ yapısı (Chen ve diğ., 2018).

Yüzey normali tahminleyen bir başka model Evrişimsel Sinir Ağı (ESA) Tabanlı Fotometrik Stereo'dur (CNN-PS) (Ikehata, 2018). Bu çalışmada isteğe bağlı miktarda girdi görüntüsü birleştirilip gözlem haritası olarak adlandırılan sabit boyutlu bir ara ifadeye çevrilmektedir. Daha sonra bu gözlem haritası bir ESA modeline girdi olarak verilmektedir. Modelin eğitiminde fiziksel tabanlı işleme ile üretilen sentetik görüntülerden oluşan veri seti kullanılmıştır. Model sentetik olmayan veri setlerinde

de başarılı sonuçlar vermiştir.

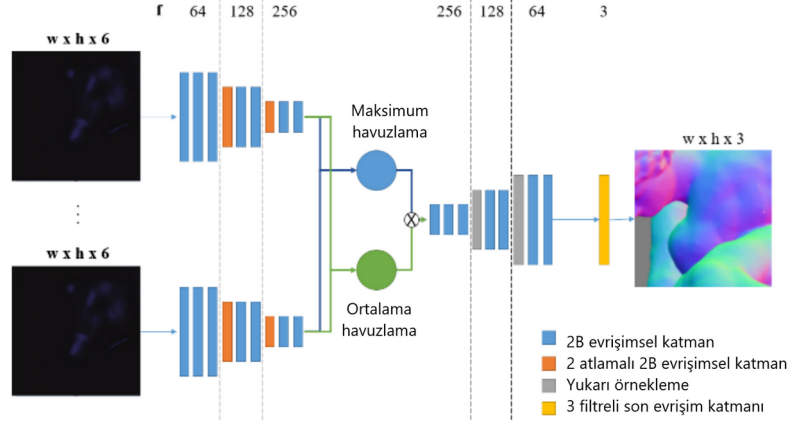
2019'da geliştirilmiş bir teknik tek bir ışık kaynağı ile çekilmiş tek bir fotoğraftan yüzeye ait materyalin normallerini, renk değerlerini, yansıma renklerini ve yansıma pürüzlülüğünü tahminleyebilmektedir (Deschaintre, 2019). Sistem kaynaklarını minimum düzeyde tüketerek çalışacak şekilde tasarlanmış model çoklu görüntü girdileri ile de çalışabilmektedir. Modelin tahminleyebildiği materyal çıktıları için bazı örnekler Şekil 1.6'da verilmiştir.



Şekil 1.6 : Deschaintre (2019) tarafından geliştirilen modelin örnek çıktıları.

Gao ve diğ. (2019) yaptıkları çalışma sonucunda fiziksel tabanlı işleme materyallerini tahminleyen otokodlayıcı tipinde bir model sunmuşlardır. Diğer çalışmalardaki modellerin aksine, çıktıların düzenlenmesi için sezgisel yöntemler kullanılmamıştır. Düzenleme işlemi doğrudan öğrenilmiş öznitelikler üzerinden gerçekleştirilmektedir. İsteğe bağlı miktarda görüntüleri girdi olarak alabilen model yüksek çözünürlüklü girdileri de desteklemektedir.

Aynı alanda yapılan başka bir çalışmada ise öznitelik birleştirme işlemi uygulanmıştır (Song ve Chang, 2020). Evrişim katmanlarından oluşan iki ayrı koldan geçerek boyutu indirgenmiş veri, maksimum havuzlama ve ortalama havuzlama işlemlerinden geçirilir. Ardından uç uca eklenen bu veriler yukarı örnekleme ve evrişim katmanlarından oluşan bir çözücüye iletilir. 3 filtreye sahip son evrişim katmanından çıkan görüntü, girdiye ait normal vektörlerini içermektedir. Modelin şeması Şekil 1.7'deki gibidir.



Şekil 1.7 : Song ve Chang (2020) tarafından geliştirilen modelin yapısı.

Son yıllarda yapılan çalışmalardan birine göre, sentetik veri setleri ile geliştirilen modeller gerçek verilerde yeterince iyi olmamaktadır (Asselin ve diğ., 2020). Adı geçen çalışma kapsamında gerçek veri setleri oluşturulmuştur. Materyallerin fotoğraflanması aşamasında eşit aralıklarla yerleştirilmiş 12 adet LED ampül ve ortalarında da Kinect Azure RGB-D kamera kullanılmıştır. Gelişmiş bir fotometrik stereo yöntemiyle elde edilen materyal bilgileriyle veri seti oluşturulmuştur. Çalışmadan kullanılan modeller bu veri seti ile eğitilmiştir. Model, yansılardan ve bazı oluşan gölgelerden etkilenmektedir fakat diğer çalışmalara kıyasla daha iyi sonuçlar verdiği gözlemlenmiştir.

2. FİZİKSEL TABANLI GRAFİK İŞLEME VE MATERYAL DOKULARI

2.1 Fiziksel Tabanlı Grafik İşleme

Fiziksel Tabanlı İşleme (FTI) bir üç boyutlu model grafiği işleme tekniğidir. Üç boyutlu görüntülerin olabildiğince gerçekçi ve yüksek performansla işlenebilmesi için geliştirilmiştir. Çoğunlukla bir renk dokusu (albedo), bir kabartı dokusu (normal map veya height map), bir metaliklik dokusu (metalness) ve pürüzlülük dokusu (roughness) ile kullanılır.

2.2 Fiziksel Tabanlı Grafik İşlemede Kullanılan Materyal Dokuları

2.2.1 Renk dokusu

3B bir modelin kaplamasında bulunan ve temel renk bilgilerini içeren dokudur. UV mapping adı verilen bir eşleme tekniği ile 2B bir görüntünün, 3B nesnenin hangi kısımlarını kaplayacağı belirtilir. Sahne işleme (render) aşamasında ise bu renk dokusundaki piksel değerlerinin, ışığın şiddeti ve açısına göre hangi tonda ekrana yansıtılacağı hesaplanır. En eski ve en temel kaplama (texture) türüdür.

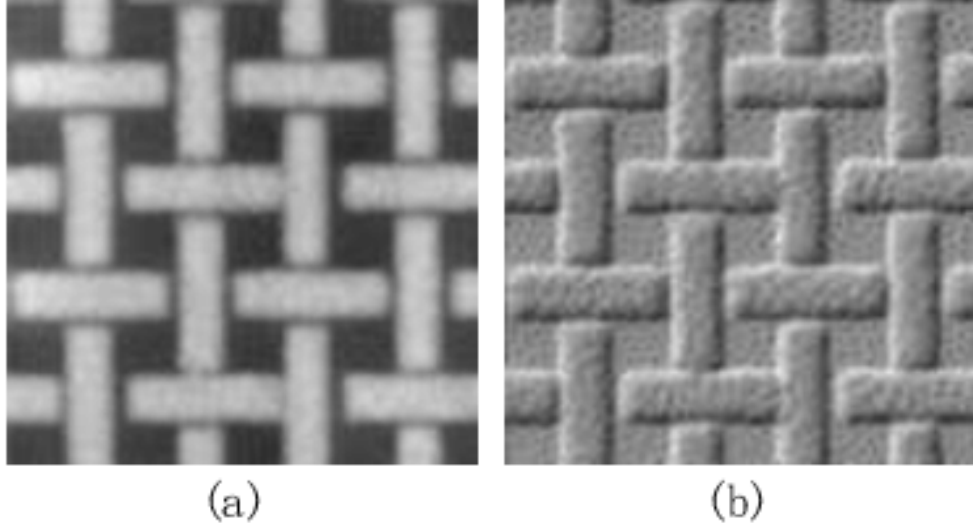
2.2.2 Kabartı dokuları (Bump Map Texture)

3B görüntülerin işlenmesinde yüzey detayları farklı doku çeşitleriyle artırılabilir. Bu amaç için kullanılan dokulardan bazıları; yükseklik dokusu ve normal dokusudur.

2.2.2.1 Yükseklik dokusu (Height map texture)

Yükseklik dokuları üç boyutlu bir modelin üzerinde bulunmayan ayrıntıları taklit

ederek modellerin poligon sayılarından tasarruf edilmesini ve işlem performansının artırılmasını sağlar. Gri tonlamalı tek kanal bir görüntü olan yükseklik dokusunda beyaz renge yaklaştıkça yükselti artarken, siyah renge yaklaştıkça ise yükselti düşer. Yükseklik dokuları, sahne ışıklarının yoğunlukları ve açıları ile birlikte işlenir. Böylece 3B yüzey üzerinde yükseltiye göre parlaklık hesaplaması yapılmış olur.



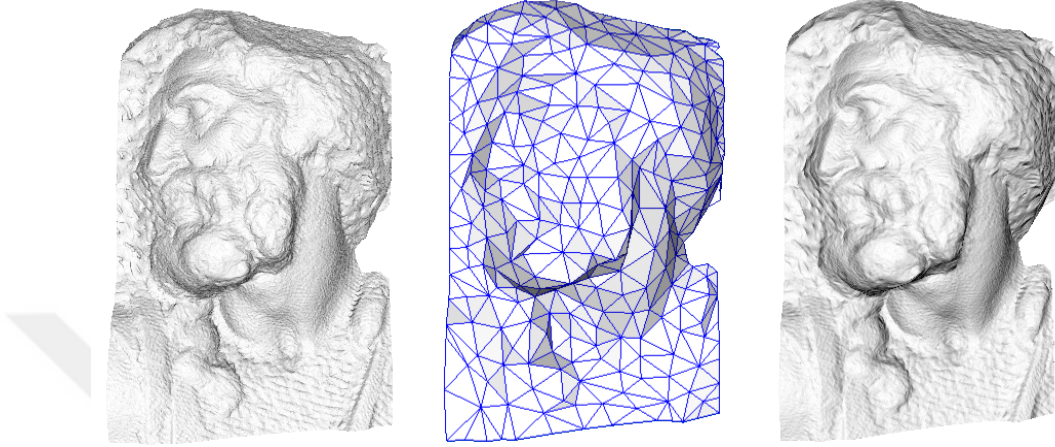
Şekil 2.1 : a) yükseklik dokusu, b) yükseklik dokusu kullanılarak işlenmiş 3B yüzey (Zhou ve diğ, 2017).

2.2.2.2 Normal dokusu (Normal map texture)

Normal dokuları, FTİ materyallerinde kullanılan, 3 boyutlu bir nesnede aslında bulunmayan ayrıntıları sahne ışığı yardımıyla varmış gibi göstermek için kullanılan dokulardır. Yüzeyin her bir noktasının normal vektörünü piksel değerleri olarak içerdiği için bu şekilde adlandırılmışlardır. Eski yöntemlerde kullanılan yükseklik dokuları (height map) sadece gri tonlardan oluşmaktaydı. Beyaza yakın pikseller yüksek kısımları belirtirken siyaha yakın kısımlar alçak kısımları belirtmekteydi. Fakat performans açısından efektif olmayan bu yöntem gittikçe yerini normal dokularına bırakmıştır. Normal dokusu ayrıca Dot3 bump map olarak da adlandırılmaktadır.

Blinn (1978) bir yüzeyin üzerindeki noktaların normal vektörlerinin, düz bir başka yüzeyin asıl yüzeydeki detaylara sahipmiş gibi davranış sergileyebilmesi için kullanılabileceğini orataya koymuştur. Bu teknik ile yüksek kaliteli üç boyutlu bir

modelin yüzey normalleri daha düşük kaliteli haline uygulanarak yüksek performans ve yüksek görüntü kalitesi elde edilebilmektedir (Krishnamurthy ve Levoy, 1996). İlerleyen zamanlarda bu alandaki çalışmalar devam etmiştir (Cohen ve diğ, 1998) (Cignoni ve diğ, 1998). Cignoni ve diğ. (1998) yaptıkları çalışmada yüzey normallerini iki boyutlu bir görüntü (texture) olarak saklamışlardır.



Şekil 2.1 : Soldan sağa; Orijinal model (4 milyon üçgen), sadeleştirilmiş model (500 üçgen), normal dokusu ile kaplanmış sadeleştirilmiş model (500 üçgen) (Cignoni ve diğ, 2005).

Kabartı dokuları ilk yıllarda sadece 3 boyutlu grafiklerin render aşamasında kullanılmaktaydı. Performans açısından avantajlı olduğu anlaşıldıkça grafik kartları özellikle kabartı dokularını gerçek zamanlı işleyebilecek şekilde geliştirildiğinden günümüzde kabartı dokuları oyunlardan video işleme yazılımlarına kadar pek çok alanda kullanılmaktadır.

Yer değiştirme dokuları (displacement map) alternatif olarak kullanılsa da yine yüksek işlem gücü gerektirmektedir. Bu doku türü 3 boyutlu nesnede bulunmayan ayrıntıları grafik kartlarında oluşturarak yüksek seviyede ayrıntılı görüntü işlemeyi sağlamaktadır. Halen yüksek kaliteli görüntü işleme tekniklerinde kullanılsa da gerçek zamanlı uygulamalarda displacement map kullanımından düşük performans sebebiyle kaçınılmaktadır.

Kabartı dokuları fotometrik yöntemler ile birden fazla görüntü kullanılarak oluşturulabilmektedir (Woodham, 1980). Üretilen bu doku düşük detaylı nesnelere

uygulandığında yüksek detaylı nesne taklit edilmiş olur. Böylece yüksek detaylı grafikler düşük sistem gereksinimleriyle erişilebilir hale getirilmiş olur. Ayrıca, yüksek seviyede poligon (high-poly) içeren 3 boyutlu nesnelerin optimizasyon amacıyla daha az seviyede poligonlu (low-poly) nesnelere çevrilmesi aşamasında da kabartı dokuları üretilebilmektedir (Krishnamurthy ve Levoy, 1996).

YSA'nın doğasındaki çok boyutlu problem çözme kapasitesi ve üretken YSA modelleri ile, bu ve buna benzer işlemlerin otomatize edilmesi sağlanabilir ve daha kısa sürede yeterli seviyede detaya sahip ürünler ortaya çıkarılabilir.

2.2.3 Metaliklik dokusu (Metalness map texture)

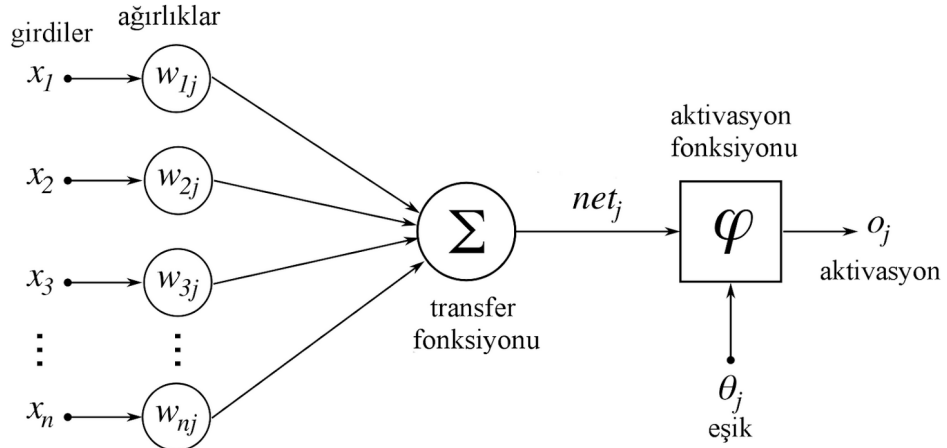
3B sahnede görüntülenecek olan bir nesnenin hangi kısımlarının metalik özelliklere sahip olacağını belirten gri tonlamalı dokulardır. Dokunun beyaza (1.0 veya 255) yakın kısımları metalik ışıklandırma ve gölgelenme özellikleri gösterirken siyaha (0) yakın kısımlarda bu özellikler daha azdır veya yoktur.

2.2.4 Pürüzlülük dokusu (Roughness map texture)

3B nesnenin ışığı ne kadar soğuracağını belirleyen dokulardır. Fiziksel pürüzlülük ile bir ilişkisi bulunmamaktadır. Pürüzlü belirtilen yüzeyler daha mat görüntüye sahip olurken, pürüzsüz yüzeyler daha iyi yansıtıcı özelliğe sahip olurlar.

3. YAPAY SİNİR AĞLARI (YSA)

Yapay Sinir Ağları (YSA), insan beyninin çalışma yapısını taklit eden bir yapay zeka uygulamasıdır. İnsan veya istatistik standartları ile çözümü imkansız veya zor olan problemlerin çözümünde YSA kullanılabilir. Sınıflandırma yapabilen YSA modelleri olduğu gibi üretken YSA modelleri de mevcuttur. Çeşitli YSA modelleri ile görüntü işleme, ses işleme, doğal dil işleme gibi pek çok işlem yapılabilmektedir. YSA ile diğer makine öğrenmesi yöntemleri arasındaki en önemli fark, YSA modellerinin öznelik çıkarımını kendiliğinden yapmasıdır. Bu önemli fark sayesinde çok boyutlu bir uzayda veriler arasındaki korelasyonlar ve bağlantılar kolayca bulunabilmekte ve el ile öznelik çıkarma gerekliliği ortadan kalkmaktadır.



Şekil 3.1 : Yapay Sinir Ağı Nöronu (Chrislb, 2005).

YSA belirli bir sayıda katmanlardan oluşur ve bu katmanlar yine belirli bir sayıda nöron içermektedir. Her bir nöron kendisine gelen giriş verilerini alır ve bir aktivasyon işleminden geçirdikten sonra kendisinden sonra gelen nörona veya nöronlara iletir. İki nöron arasında bulunan bağlar birer ağırlık (w) değerine sahiptir. Bir nöronun çıkış değeri o nörona giren veriye ve verinin geldiği bağın " w " değerine bağlıdır. Backpropagation algoritması sayesinde bir YSA sondan başa doğru geri besleme ile eğitilir ve her bir bağdaki " w " değerleri güncellenir. Bir süre sonra

yeterince eğitilmiş olan YSA modeli daha doğru kararlar almaya başlar. YSA kullanım alanlarına göre çeşitli şekillerde tasarlanabilir. Örneğin görüntü işleme için Evrişimsel Sinir Ağları (ESA; CNN, Convolutional Neural Networks) kullanılırken zamana bağlı verileri işlemek için Yinelemeli Sinir Ağı (RNN, Recurrent Neural Networks) ve Uzun Kısa Süreli Bellek (LSTM, Long Short-Term Memory) gibi YSA modelleri kullanılmaktadır. Reinforcement Learning yöntemi ile YSA modelleri, canlı olarak girdi ve çıktıları tecrübe edip öğrenebilmektedir.

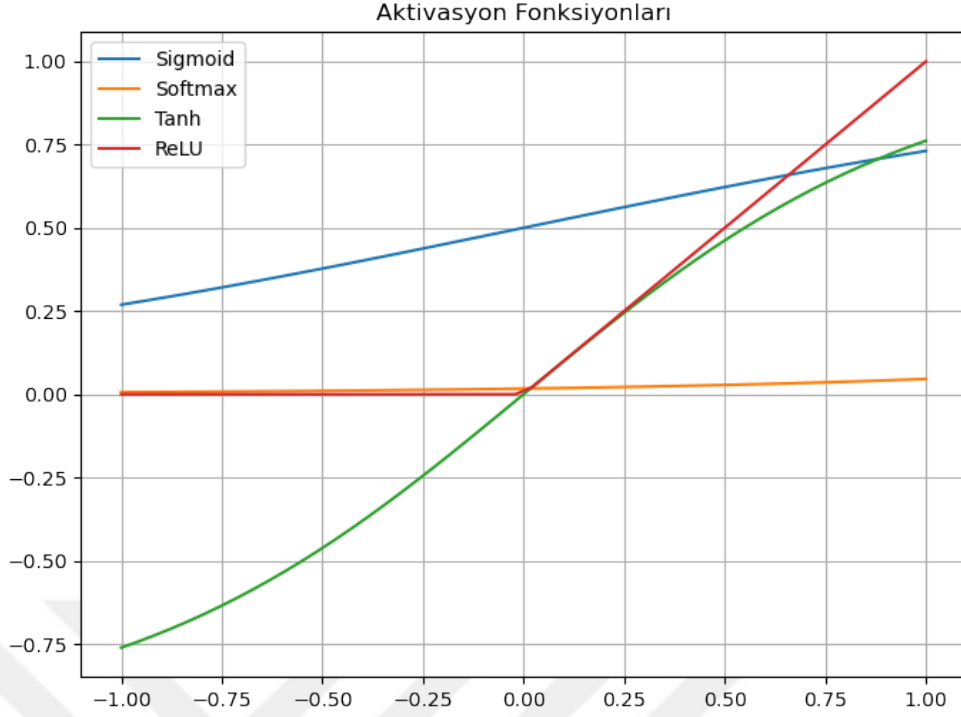
3.1 Yapay Sinir Ağları'nın Aşamaları

Yapay sinir ağları ile yapılan iki tür işlem vardır: ileri besleme (feedforward) ve geri yayılım (backpropagation). Bir yapay sinir ağı geri yayılım aşaması ile eğitilir. Eğitilmiş bir ağ ise ileri besleme yöntemiyle kullanılır.

3.1.1 İleri besleme

Yapay sinir ağlarında nöronlar eğitilebilir olan parametrelerdir. Her iki nöron bağlantısı arasında öğrenilen bir ağırlık değeri bulunur. Girişten gelen bir girdi verisi geçtiği bağlantının ağırlığıyla çarpılır. Ardından elde edilen değer bir aktivasyon fonksiyonuna parametre olarak verilir. Aktivasyon fonksiyonları verilerin skalasını limitlemektedir ve sadece belirli bir eşiği aşabilen; yani nöronun aktivasyonunu sağlayabilecek seviyede olan değerlerin geçişine izin verir. Geri kalan veriler sıfırlanmaktadır veya çok küçük değerler haline gelmektedir.

Sigmoid, Softmax, Hiperbolik Tanjant (tanh), Düzeltilmiş Doğrusal Birim (ReLU) en çok kullanılan aktivasyon fonksiyonlarından. Bu fonksiyonların $[-1, 1]$ aralığındaki kısımları Şekil 3.1'de karşılaştırılmıştır.



Şekil 3.1 : Sıklıkla kullanılan aktivasyon fonksiyonlarından bazılarının kıyaslanması.

3.1.2 Geri yayılım

Geri yayılım sadece eğitim aşamasında gerçekleşen bir süreçtir. Girdi verileri giriş katmanından çıkış katmanına kadar iletilir. Çıktı değerleri ile gerçek değerler arasındaki fark kadar hata veya kayıp değeri olarak isimlendirilir. Son katmandaki ağırlıklar, kayıp değerinin öğrenme oranı (α) ile çarpımı kadar kaydırılır. Bu şekilde optimum ağırlık değerlerini hesaplamaya yarayan yöntem gradyan iniş (gradient descent) algoritmasıdır (Robbins ve Monro, 1951). Gradyan iniş formülü 3.1 numaralı eşitlikte verilmiştir. Aynı işlem yapay sinir ağının başına kadar her katman için yapılır. Bir ağı son dan başa tüm ağırlıklarının güncellenmesi geri yayılım (back propagation) olarak adlandırılır. Bir geri yayılım sonunda bir yığınlık veri (batch) eğitim için kullanılmış olur. Eğitim setindeki tüm verilerin birer defa eğitimde kullanılması ise 1 eğitim döngüsünü (epoch) oluşturur.

$$w_{yeni} = w_{eski} - \alpha \frac{\partial L}{\partial w} \quad (3.1)$$

Yüksek öğrenme oranı, ağırlıkların optimum noktalarının yakalanmasını zorlaştırmaktadır. Eğitimin yüksek kayıp değeri ile bitmesine sebep olabilmektedir. Hatta yüksek ve sabit öğrenme oranları kayıp değerinin eğitim ilerledikçe düşmemesine neden olabilmektedir. Düşük öğrenme oranı ise ağırlıkların optimum değerlerine çok yavaş bir hızda gidilmesine yol açmaktadır. Bu da istenen eğitim kalitesine ulaşmak için fazla zaman harcanması anlamına gelmektedir. En iyi öğrenme oranı için modelin kayıp değerine, öğrenme hızına ve momentumuna bakılması gerekmektedir. Yüksek öğrenme hızı ile eğitime başladıktan sonra giderek öğrenme oranını düşürmek optimum noktaya en hızlı şekilde ulaşmanın yolu olabilmektedir. Bu işlem için optimize ediciler kullanılmaktadır (Ruder, 2016). Adam (Kingma ve Ba, 2014), Adagrad (Duchi ve diğ, 2011), Adadelata (Zeiler, 2012), Rmsprop en yaygın optimize edicilerdendir.

Öğrenme aşamasında hata miktarını ve yapılması gereken ağırlık düzeltmesinin yönünü tespit etmek için kayıp fonksiyonları kullanılmaktadır. Kayıp fonksiyonları aynı zamanda bir objektif fonksiyonudur. Bu fonksiyon ile, amaçlanan hedefe ulaşabilmek için maksimizasyon veya minimizasyon işlemi yapılır. Basit ve sıkça kullanılan bir kayıp fonksiyonu olan Ortalama Karesel Hata (MSE) yöntemine göre; gerçek değer ile tahmin edilmiş olan değer arasındaki farkın karesi ölçülür ve tüm bu ölçümlerin ortalaması hesaplanır.

3.2 Yapay Sinir Ağları Modelleri

Yapay Sinir Ağları (YSA), makine öğrenmesinden farklı olarak öznitelik çıkarımını otomatik olarak yapabilmektedir. Bu avantajı sayesinde doğal dil işlemeden görüntü işleme kadar pek çok alanda kullanılmaktadır.

Yapay zeka modelleri ayırıcı modeller ve üretici modeller olarak ikiye ayrılır. Yapay Sinir Ağları ile hem ayırıcı modeller hem de üretici modeller oluşturulabilmektedir.

3.2.1 Ayırıcı modeller (Discriminators)

Ayırıcı modeller, eğitim aşamasında verilerin ait oldukları sınıfların arasındaki sınırları belirler. Tahmin aşamasında ise tahmin edilecek verilerin özniteliklerine

göre hangi sınırların arkasında kaldığını hesaplar. Sınıflandırma yapılan modeller ayırıcı modeller grubuna dahildir.

Yapay Sinir Ağları, ayırıcı model olarak da kullanılabilir. Çoğunlukla sınıf sayısı kadar nöron içeren bir çıkış katmanı hazırlanır. Tüm çıkış nöronları 0 veya 1 verecek şekilde, tüm model eğitilir. Bir veri yalnızca bir sınıfa ait ise çıkışlardan sadece biri 1 değerini alır, diğerleri ise 0 olmalıdır. Bunun için one-hot kodlaması kullanılır. Her sınıfa sadece kendisine ait olan basamaktaki rakamın 1, diğerlerinin 0 olduğu benzersiz bir kod verilir. Bu kodların çıkış verisi olarak kullanılması ile yeterince eğitilen ağ, tahmin aşamasında girdideki verinin ait olabileceği sınıfların çıktı değerini yüksek verir. Her çıkış nöronu, ait olduğu sınıfın yüzdelik olasılığını gösterir.

3.2.2 Üretici modeller (Generators)

Üretici modeller, girdi olarak verilen verilerin dağılım fonksiyonunu yakınsayan modellerdir. Yeterince veri olduğunda bu verilerin dağılımı hakkında çıkarımlar yapılabilmektedir. Geleneksel makine öğrenmesi yöntemleriyle ortalaması ve standart sapması hesaplanan verilerin sınıfına ait yeni bir veri, sınıfın dağılım fonksiyonunun bu parametrelerle hesaplanmasıyla üretilebilmektedir.

Yapay Sinir Ağları ile üretici modeller de tasarlanabilmektedir. Çıktı olarak, üretilmesi istenen verilerin uzayından örnekler verilmektedir. Yeterince eğitilen ve ezberleme yapmamış olan yapay sinir ağı, eğitim verilerinin uzayından yeni ve benzersiz resimler üretebilir hale gelmektedir. Otokodlayıcılar, Çekişmeli Üretici Ağlar (GANs) sıkça kullanılan üretici yapay sinir ağı modeli örnekleridir.

3.3 Yapay Sinir Ağları ile Görüntü İşleme

3.3.1 Evrişimsel sinir ağları (ESA, CNN)

Konvolüsyon (evrişim) işlemi, görüntü işlemede bir filtrenin bir girdiye uygulanması aşamasıdır. Bir filtrenin bir görüntüye uygulanması sonrasında öznitelik haritası (feature map) elde edilir. Çeşitli filtreler yeterli sayıda girdi görüntülerine

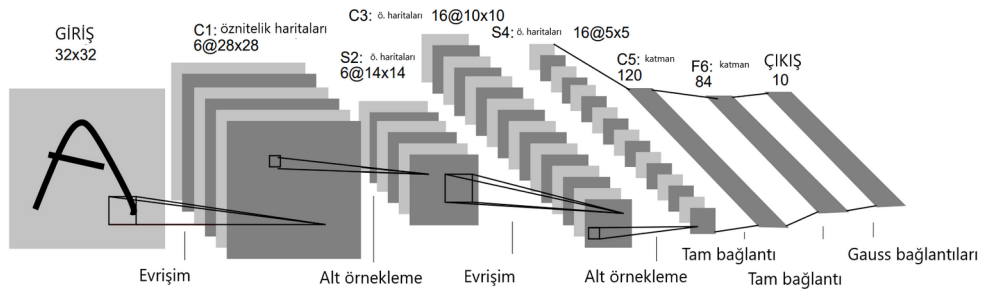
uygulanarak ortak ve bariz öznitelikler çıkarılır ve bir YSA'ya iletilerek hangi özniteliklerin ne kadar ağırlığa sahip olması gerektiği bulunabilir.

ESA'lar, evrişim işlemleri içeren YSA modellerine verilen genel bir isimdir. Bu yöntem ile 2 boyutlu görüntülerde bulunan verilerin oluşturduğu örüntüler tespit edilebilmektedir. ESA, evrişim aşaması uygun filtre boyutları kullanıldığı zaman 2. boyutun dışındaki uzaylarda da kullanılabilir.

3.3.2 Evrişimsel sinir ağları modelleri

3.3.2.1 LeNet

LeCun ve diğ. (1998) tarafından el yazılarını ve bilgisayar çıktısı yazıları tanımlama problemine çözüm geliştirmek amacıyla LeNet modeli sunulmuştur. Önerilen yöntem geleneksel evrişimsel sinir ağlarının temellerini içermektedir. Tek kanallı görüntüyü girdi olarak alan model, 10 adet çıktı nöronlarından birinden 1 diğerlerinden 0 değeri verecek şekilde eğitilmiştir. Genel olarak 1'e en yakın değeri veren nöronun denk geldiği rakam, giriş katmanından verilen görüntüdeki rakamdır. Geliştirilen modelin şeması Şekil 3.2'de görülmektedir.



Şekil 3.2 : LeNet modeline ait ağ yapısı (LeCun ve diğ, 1998).

3.3.2.2 AlexNet

AlexNet, 2012 yılında ImageNet veri seti ile gerçekleştirilen bir yarışma etkinliği için üretilmiştir (Krizhevsky ve diğ, 2012). LeNet'in daha gelişmiş bir versiyonudur. 5 evrişimsel katmandan ve 3 tam bağlı katmandan oluşmaktadır. Yarışmayı kazanan model yaklaşık 60 milyon parametre içermektedir.

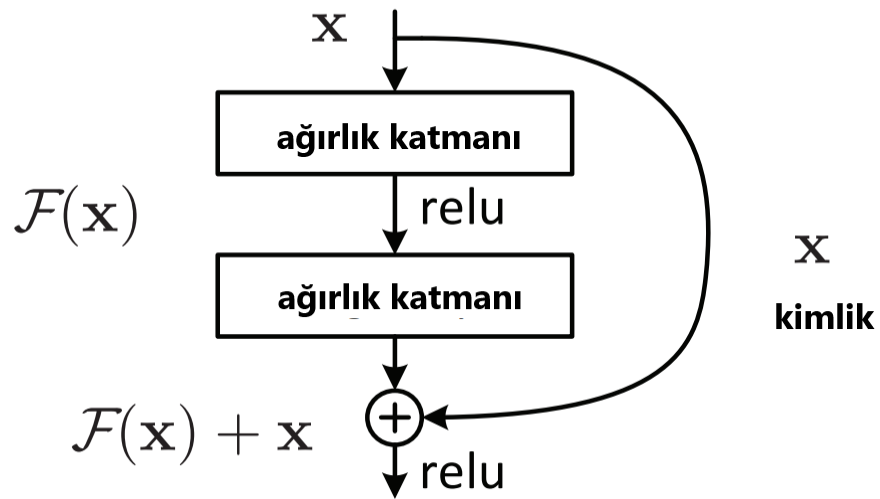
3.3.2.3 VGG Net

Evrişimsel sinir ağlarında yüksek hiper parametre sayılarını düşürebilmek için yüksek boyutlu filtreler yerine daha düşük boyutlu filtrelerin uygulandığı evrişim katmanları kullanılmıştır (Simonyan ve Zisserman, 2014). Bu yöntemin sunulduğu VGG Net modeli, hiper parametreleri azaltsa da uzun sürede eğitilebilmektedir ve kaybolan gradyan problemini açığa çıkarmaktadır.

3.3.2.4 ResNet

Bir başka müsabaka için geliştirilmiş olan ResNet, 100'den fazla katmana sahiptir (He ve diğ., 2016). Klasik evrişimsel katmanlara ek olarak atlama bağlantılarına sahiptir. Bu atlama bağlantıları, yapay sinir ağlarının eğitiminde yaşanan kaybolan gradyan problemi için geliştirilmiştir. Geri yayılım sürecinde gittikçe küçülen türevler sebebiyle belirli sayıda katmandan sonra eğitim kalitesinin düşmesi bu problem sebebiyledir. Modelin ilk katmanlarından sonraki katmanlarına doğru yapılan bağlantılar bu problemin etkisini azaltmaktadır. Ayrıca ResNet ağı, yığın normalizasyonunun kullanıldığı ilk modeldir. Normalizasyonun yığın boyunca yapılması sayesinde evrişimsel sinir ağlarının eğitim süresi düşürülebilmektedir.

ResNet tipi ağlarda kullanılan bağlantı yöntemi Şekil 3.3'te verilen şemada gösterilmiştir.



Şekil 3.3 : ResNet ağlarının yapı taşı olan bağlantı (He ve diğ., 2016).

3.3.2.5 Inception Net (GoogLeNet)

Google geliřtiricileri tarafından sunulan Inception Net modeli, farklı filtre boyutlarını aynı anda kullanan bir ađ yapısına sahiptir (Szegedy ve diđ, 2015). Kùçük detaylar için küçük boyutlu filtre boyutları kullanılırken, büyük ve kapsamlı öznitelikler için büyük filtre boyutları kullanılmaktadır. Inception Net, yapısı sayesinde çeřitli boyutlardaki öznitelikleri bir arada tespit edebilmektedir. 1x1, 3x3 ve 5x5 filtre boyutlarına sahip olan modelin dezavantajı yüksek hiper parametre sayısına sahip olmasıdır.

3.3.2.6 DenseNet

ResNet ađlarındaki yapının bir benzeri DenseNet modelinde kullanılmıřtır (Huang ve diđ, 2017). Evriřim katmanına sahip her bir blok, kendisinden sonraki bloklara bir geçiř katmanı ile bađlıdır. ResNet ađlarında önceki katmanlardan gelen veriler sıradaki katmanın deđerleriyle toplanırken, DenseNet modelinde uç uca ekleme yapılarak bu verilerin arasındaki bađlantılar kurulur.

3.3.2.7 Otokodlayıcılar (Autoencoders)

Otokodlayıcılar (Rumelhart ve diđ, 1985), YSA kullanarak veri çıktıyı alınabilen bir üretici modeldir. Bir otokodlayıcı, girdilerin önemli özelliklerini tespit edip bunlarla verilerin gizli uzayını (latent space) oluşturup tekrar bu uzayı kullanarak girdilere benzer çıktılar üretebilir. Temelde iki kısımdan oluşur; kodlayıcı (encoder) aşaması ve çözücü (decoder) aşaması. Kodlayıcı aşamasında veriler YSA içerisinde işlenerek gizli uzaya aktarılır. Çözücü aşaması ise kodlayıcı aşamasının simetrik hali olup gizli uzaydaki matrisi kullanarak veriler üretir. Bir YSA, evriřim aşaması içeren bir kodlayıcıya ve bu aşamanın tam tersinden oluşan bir çözücüye sahipse, girdi-çıkıtı ilişkisini öğrenip girdileri taklit etmeye çalışabilir. Ayrıca girdilere, doğru çıktı alabilmek için, uygulanması gereken görüntü işleme fonksiyonlarını öğrenip bunları uygulayabilir.

4. MATERYAL VE YÖNTEM

4.1 Veri Seti Büyütme (Data Augmentation)

Yapay zeka uygulamaları geliştirirken yeterli sayıda ve çeşitlilikte veri bulunmadığı zaman eldeki verilerin kullanılmasıyla veri seti genişletmesi yapılabilir. Görüntü, ses, zaman serisi veya metin gibi çeşitli türlerdeki veriler, taşıdıkları anlamlar bozulmayacak şekilde biçimlendirilip çeşitlendirilebilir ve var olan veri setine ekleme yapılabilir.

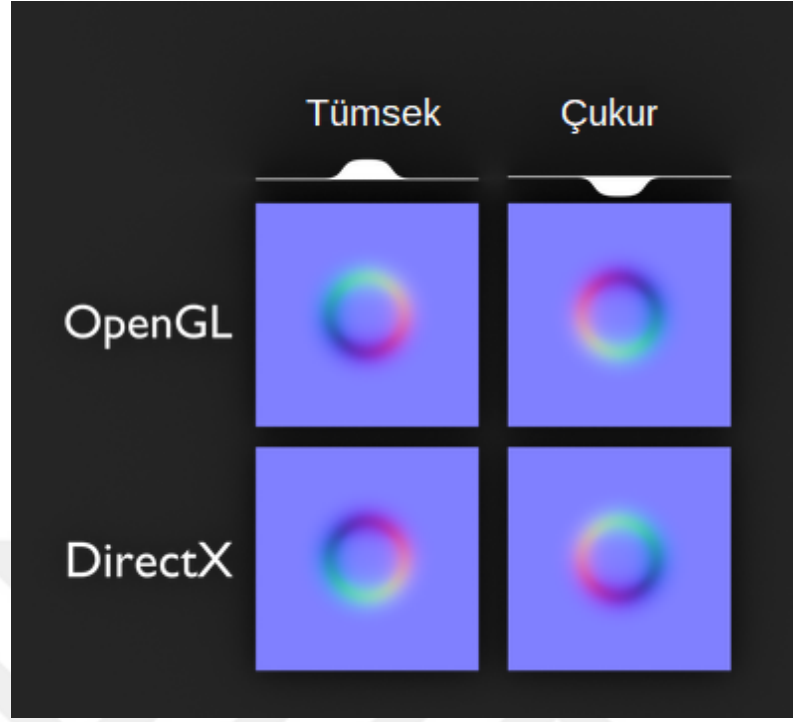
4.1.1 Görüntü veri setlerinin büyütülmesi

Görüntü verilerinin çoğaltılması işlemlerinde çeşitli morfolojik görüntü işleme teknikleri kullanılmaktadır. Sınıflandırıcı yapay zeka modellerinin eğitimlerinde genellikle bu işlemlerin girdi verilerine uygulanması yeterlidir. Bölütleme, stil transferi, görüntüden görüntüye çeviri (pix2pix) yapan ve giriş verilerine uygulanacak olan işlemlerin çıkış verilerini etkilediği tüm modellerde çıkış verilerini biçimlendirmek de gerekmektedir.

Bu tez çalışmasında eğitilen yapay sinir ağı modelleri girdi olarak renkli (RGB) yüzey kaplamaları almakta ve çıktı olarak ise yüzeyin normal vektörlerini tahminleyerek kabartı haritası vermektedir. Bu sebeple girdi verilerine uygulanan tüm morfolojik işlemler tekabül eden çıktı verilerine de aynı parametrelerle uygulanmıştır.

Kabartı haritalarında yüzeyin üç boyutlu uzaydaki konumuna ve yönüne göre değişkenlik gösteren bilgiler mevcuttur. Daha alçak piksellerden daha yüksek piksellere geçiş sırasında uygulanan renk, geçişin yönüne göre değişmektedir. Bu durum grafik motoruna göre de değişiklik gösterebilmektedir. OpenGL ve DirectX motorlarının yüzey normallerini işlerken kullandığı Y eksenini doğrultusu birbirine

zıttır.



Şekil 4.1 : Yüzey kabartmalarının grafik motorlarında zıt Y eksenini doğrultularında işlenmesi (Duha, 2021).

4.1.2 Döndürme

Kabartı haritalarına döndürme işlemlerinin uygulanması yüzeylerin yükselti geçiş yönünü değiştirmektedir. Yüzeylerin üç boyutlu ortamdaki uzaysal kabartı bilgileri hata içereceği için kabartı haritalarına döndürme vb. morfolojik işlemlerin uygulanmaması veya çok küçük açılarla uygulanması gerekmektedir.

Bu tez çalışmasında veriler saat yönünde ve tersinde en fazla onar derece döndürülerek artırılmıştır.

4.1.3 Bükme (Shearing)

Görüntülerin yukarı ve aşağı kenarlarının zıt yönlerde kaydırılmasıyla gerçekleştirilen morfolojik işlemdir. Bu yöntem kullanılarak yapılan veri çoğaltımı sayesinde eğitilen modelin perspektif, kamera bakış açısı ve nesne yönü gibi etmenlerden daha az olumsuz etkilenmesi beklenir.

Bu tez çalışmasında iki yönde de 0.25 değerinde bükme işlemi yapılarak verilerin çeşitliliği artırılmıştır.

4.1.4 Yakınlaştırma

Görüntülerin büyütülüp orijinal çözünürlüğü koruyacak şekilde kırılması amacıyla uygulanan görüntü çoğaltma tekniğidir. Bu çalışmada verilere 0.5 oranında yakınlaştırma uygulanıp veri büyütme işlemi uygulanmıştır.

4.1.5 Yatay çevirme

Soldaki piksellerin sağa denk gelecek şekilde görüntü aynalama işlemidir. Bu çalışmada çıktı verilerinin hatalı olmasına sebep olduğu için kullanılmamıştır.

4.1.6 Dikey çevirme

Yukarıdaki piksellerin aşağıya denk gelecek şekilde görüntü aynalama işlemidir. Bu tez çalışmasındaki veriler için uygun olmadığından veri çoğaltma esnasında bu özellik kapalı tutulmuştur.

4.1.7 Doldurma modu

Döndürme ve bükme gibi morfolojik işlemlerin sonucunda nihai görüntüde boşluklar oluşur. Bu boşluğu oluşturan piksellerin varsayılan renk olan siyah kalmaması için çeşitli yöntemlerle doldurma işlemi yapılır. Sabit değer verme, en yakın bilinen pikselin rengini kopyalama, yansımaları alma ve tekrar ettirme gibi yöntemlerle boş pikseller doldurulabilmektedir. Bu çalışmada çıktı verilerinin hatalı olmaması için veri çoğaltma esnasında boş pikseller en yakın piksellerin rengiyle doldurulmuştur.

4.1.8 Parlaklık aralığı

Görüntüdeki piksellerin parlaklığının artırılması veya azaltılması için kullanılmaktadır. İki parametre arasındaki değerler rastgele bir şekilde kullanılarak yeni üretilen görüntülere farklı parlaklık değerleri verilmektedir. Değerin 1.0 olması parlaklığın aynı kalacağı anlamına gelmektedir.

Kabartı dokuları yön bilgisi içerdiği için bu dokuların parlaklık veya zıtlık gibi değerlerinin değiştirilmesi, normal vektörlerinin değerlerini de değiştirir. Bu da materyalin karşılık geldiği yüzeye ait bilgilerin hatalı hesaplanmasına sebep olur. Bu nedenle bu tez çalışmasında parlaklık değişimi sadece renk dokularına uygulanmıştır. Kaplama dokularının parlaklık değerleri değiştirilmemiştir.

Renk dokularının parlaklığında 0.75 – 1.25 arası rastgele değerler kullanılmıştır. Parlaklığı fazla artırmak veya azaltmak doku verilerindeki bilgilerin azalmasına veya kaybolmasına sebep olacağından bu aralıktaki değerler veri artırımı işlemleri için uygun görülmüştür.

Kabartı dokuları için parlaklık parametrelerinin her ikisi de 1.0 olarak ayarlanmıştır. Bu sayede Keras kütüphanesi, iki farklı parametre seti ve iki aynı tohum (seed) değeri ile veri çoğaltırken sadece renk dokularının parlaklığını değiştirmiştir.

4.2 Deneysel Çalışmalar

Bu tez çalışmasında 3 farklı model, farklı evrişim katmanı sayıları veya farklı filtre sayıları ile deneylenmiştir. Veri seti için Pixar'ın Pixar One Twenty Eight ve Pixar One Thirty verileri kullanılmıştır (Sisson, 2018) (Pedersen, 2018). Veri sayısı başarılı bir model eğitebilmek için yeterli olmadığından veri seti büyütme işlemleri uygulanmıştır. Veri seti büyütme işlemleri için Python'da Keras kütüphanesinin içinde bulunan ImageDataGenerator fonksiyonu çalıştırılmıştır. İşlem sonucunda 924 adet renk dokusu ve 924 adet kabartı dokusu elde edilmiştir. Verilerin %70'i (646 tanesi) verilerin eğitimi, %30'u (278 tanesi) ise eğitilmiş modellerin test edilmesi için kullanılmıştır. Eğitim öncesinde veriler rastgele bir şekilde karıştırılmıştır. Karıştırma işlemi için farklı tohum (seed) değerleri kullanıldığında belirgin bir fark gözlemlenmediğinden çapraz doğrulama işlemi uygulanmamıştır.

Bu çalışmada her model tipi için 4 farklı katman sayısı ve 4 farklı filtre sayısı için yapay sinir ağı yapıları oluşturulmuştur. Katman ve filtre sayıları için kullanılan parametreler Çizelge 4.1'de verilmiştir. Çizelge 4.2'de her deney kodu için hangi parametrelerin verildiği listelenmiştir. Listelenen deney kodları, tez çalışması kapsamında incelenen 3 tip sinir ağı mimarisi için de kullanılmıştır. Oluşturulan tüm

modeller 200 adet döngü ile eğitilmiştir. Tüm modellerin eğitimlerinde yığın boyutu (batch size) 32 olarak seçilmiştir. Toplamda 48 adet olan deneylerin sonuçları birbirleriyle karşılaştırılmıştır. Katman sayılarının, filtre sayılarının ve ResNet ağlarında olduğu gibi hatırlatıcı bağlantıları otokodlayıcılarda kullanmanın kabartı dokuları oluştururken sonuca etkisi incelenmiştir.

Çizelge 4.1 : Farklı model tipleri için deneylenecek parametreler ve değerleri.

Katman Sayısı	Başlangıç Filtre Sayıları
2	8, 16, 32, 64
3	8, 16, 32, 64
4	8, 16, 32, 64
5	8, 16, 32, 64

Çizelge 4.2 : Deney kodlarına göre modellerin üretiminde kullanılan parametreler.

Deney Kodu	Katman Sayısı	Başlangıç Filtre Sayısı
Deney A1	2	8
Deney A2	2	16
Deney A3	2	32
Deney A4	2	64
Deney B1	3	8
Deney B2	3	16
Deney B3	3	32
Deney B4	3	64
Deney C1	4	8
Deney C2	4	16
Deney C3	4	32
Deney C4	4	64
Deney D1	5	8
Deney D2	5	16
Deney D3	5	32
Deney D4	5	64

Çıktı olarak alınan görüntülerin karşılaştırılabilmesi için PSNR metriği seçilmiştir. PSNR metriği Ortalama Karese Hata (MSE) fonksiyonu ile ters orantılıdır (4.2). MSE, iki sayı dizisinin elemanları arasındaki farkların karelerinin toplanıp tüm dizi boyutuna bölünmesiyle elde edilir. Denklem 4.1’de I ve K boyutları aynı olan iki farklı matris, m ve n değerleri ise bu matrislerin yatayda ve dikeyde uzunluklarıdır. Denklem 4.2’de MAX_I değeri, verinin alabileceği en yüksek değerdir. Bu değer, tam sayı formatındaki görüntülerde 255, ondalık sayı formatındaki görüntülerde ise 1.0’dır. MSE kullanılarak, iki görüntü arasındaki farkın yüksek olduğu piksellerden

çok yüksek, düşük olduğu piksellerden çok düşük sayılar elde edilebilir. Bu sayede hatanın yüksek olduğu nöronlara daha yüksek miktarda düzeltme işlemi uygulanabilir. MSE, Yapay Sinir Ağları'nda kayıp fonksiyonu olarak sıkça kullanılmaktadır. Bu tez çalışmasındaki modellerin eğitiminde kayıp fonksiyonu MSE olarak seçilmiştir.

Yapılan bazı çalışmalar (Thomos ve diğ., 2005) (Li ve Cai, 2007), kablosuz olarak aktarılan bir verinin PSNR değerinin 20 - 25 dB arasında olması durumunda kalite kaybının kabul edilebilir seviyede olduğunu ortaya koymuştur. Bu sebeple bu tez çalışmasında 20 dB ve üzeri PSNR değerlerine sahip çıktılar veren modeller üretmek amaçlanmıştır.

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (4.1)$$

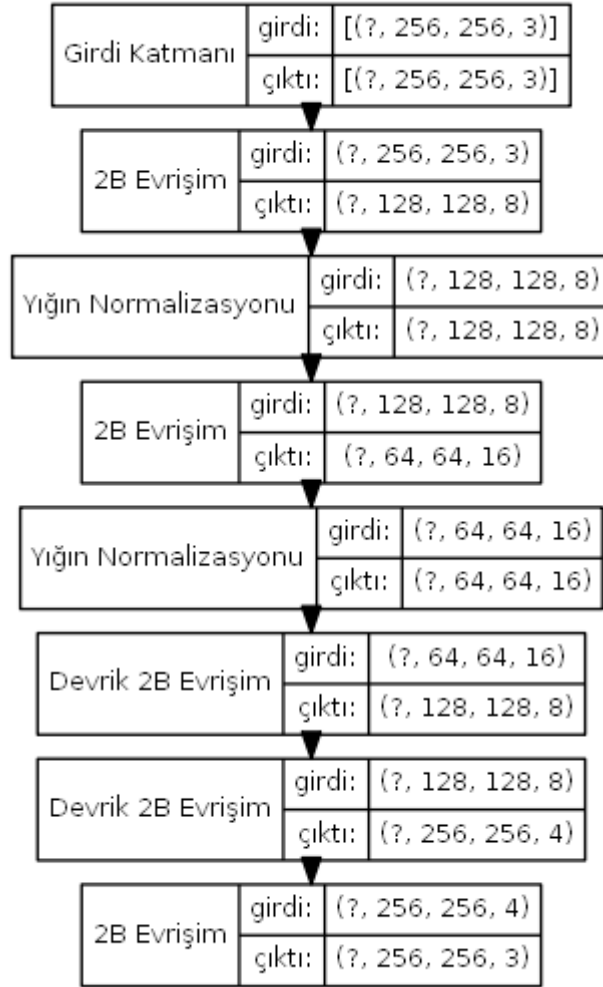
$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \quad (4.2)$$

4.2.1 Standart otokodlayıcı

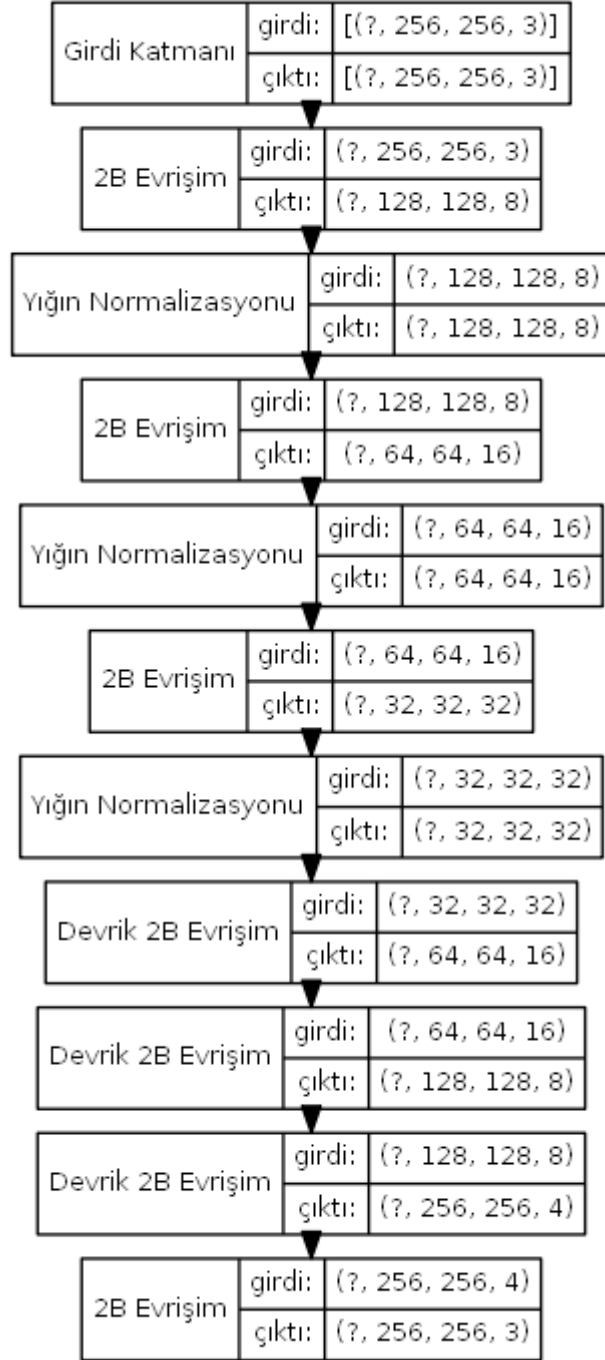
Bu model tipinde geleneksel otokodlayıcı modeli kullanılmıştır. Girdi verileri modelin kodlayıcı kısmında belirli sayıda katmandan geçerek önce küçültülüp derinleştirilmiştir. Ardından bu veriler modelin çözücü kısmında tekrar büyütülüp derinlikleri azaltılmıştır.

Deneysel çalışmanın bu kısmında katman sayısı için 4 farklı değer ve başlangıç filtre sayısı için 4 farklı değer kullanılarak tüm kombinasyonlarla 16 adet standart otokodlayıcı model oluşturulmuştur. Katman sayısı parametresi 2, başlangıç filtre sayısı 8 olan modelin şeması Şekil 4.1'de verilmiştir. Aynı tip modelin ilk parametresi 3 yapıldığında oluşan hiyerarşik yapısı Şekil 4.2'de gösterilmiştir. Katman sayısı parametresi 2, başlangıç filtre sayısı 16 olan model ise Şekil 4.3'teki

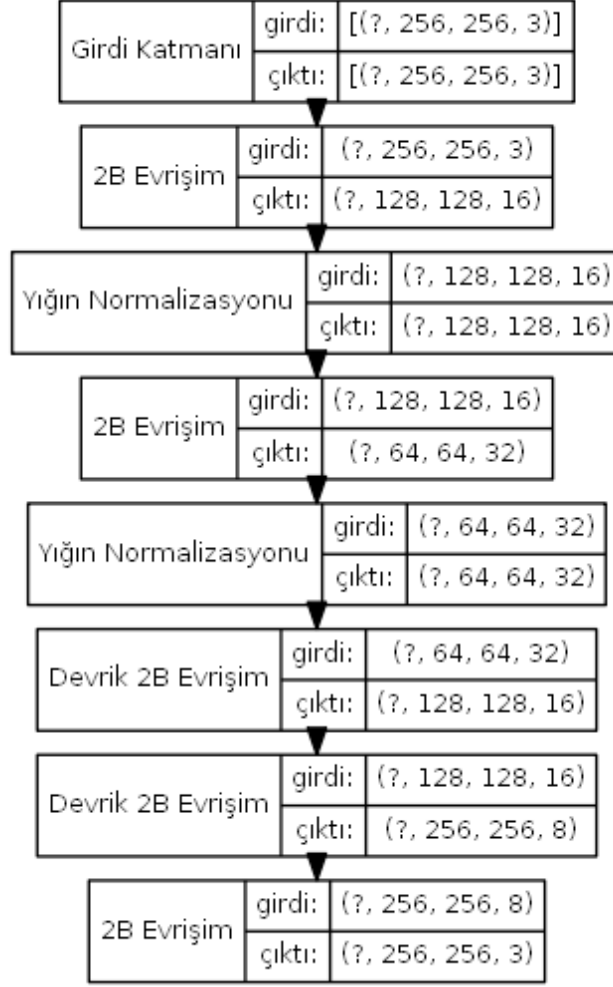
gibidir.



Şekil 4.1 : İlk parametresi 2, ikinci parametresi 8 olan standart otokodlayıcı modeli.



Şekil 4.2 : İlk parametresi 3, ikinci parametresi 8 olan standart otokodlayıcı tipindeki model.

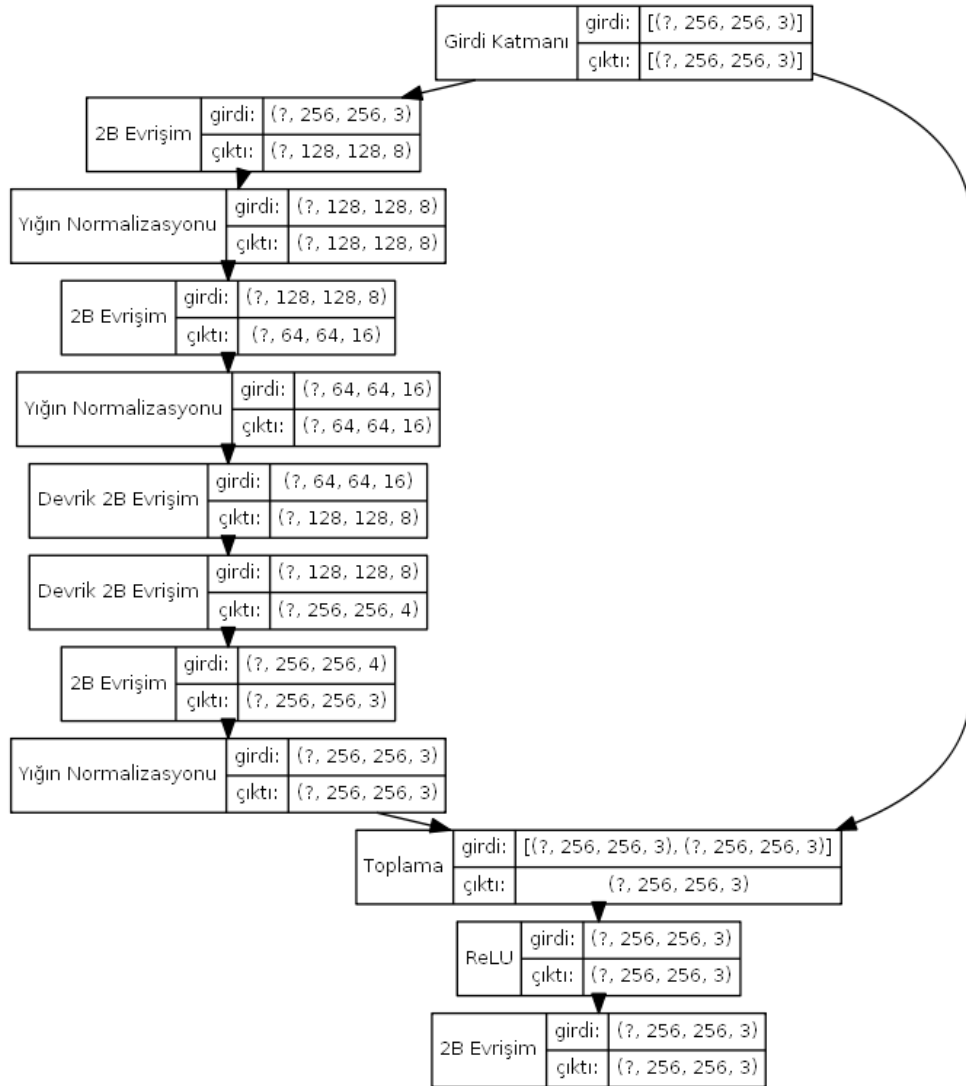


Şekil 4.3 : İlk parametresi 2, ikinci parametresi 16 olan standart otokodlayıcı modeli.

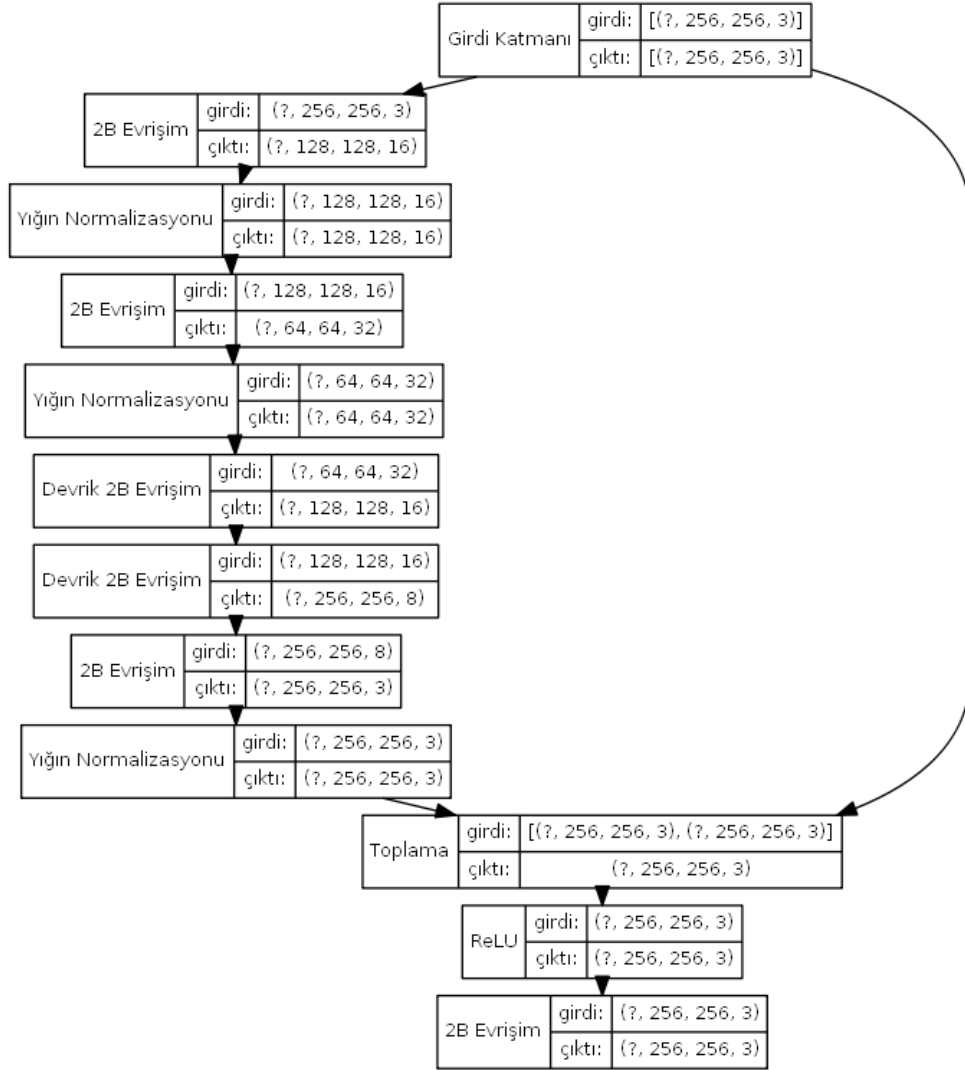
4.2.2 Kimlik eşlemeli ResNet tipi otokodlayıcı

Bu tipteki modellerde, ResNet ağlarındaki girdiyi hatırlatma tekniği kullanılmıştır. Standart bir otokodlayıcının son katmanında sadece evrişim işlemi uygulanmıştır. Bu katmanda aktivasyon fonksiyonu kullanılmamıştır. Evrişim katmanından çıkan verilerin doğrudan yığın normalizasyonu katmanına geçmesi sağlanmıştır. Bu katmanın sonucunda ortaya çıkan verinin boyutları girdinin boyutlarıyla aynıdır. Son yığın normalizasyonu katmanı çıktısının sayıları ile modelin girdisindeki sayılar karşılıklı olarak toplanmıştır. Toplama işleminin sonucunda ortaya çıkan aynı boyuttaki 3 boyutlu matris ReLU aktivasyon katmanına iletilmiştir. Aktivasyon fonksiyonundan çıkan veriler nihai halini alması için Sigmoid aktivasyon fonksiyonunun kullanıldığı son evrişim katmanına aktarılmıştır.

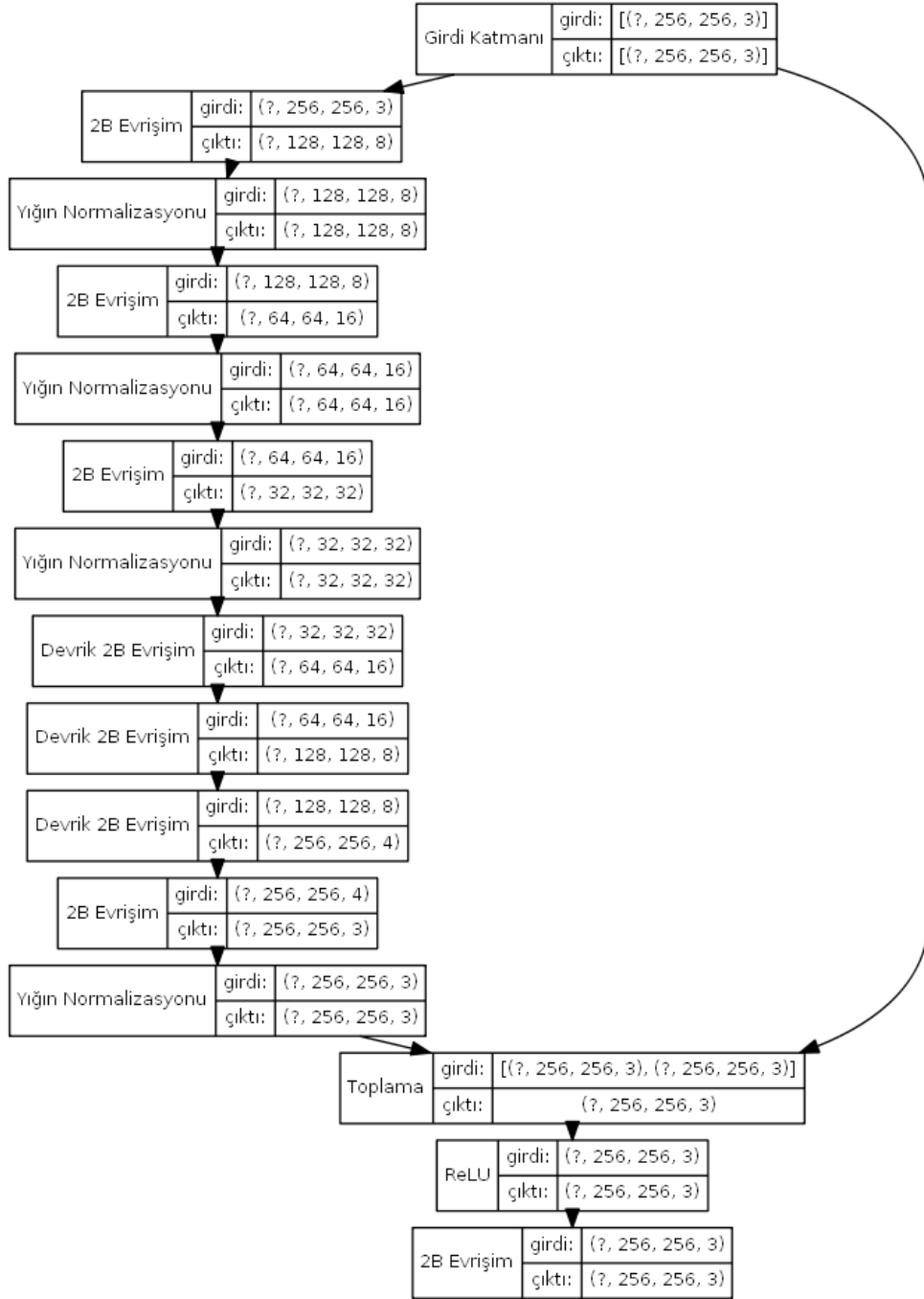
Bu tipteki modellerin örnek mimari yapıları Şekil 4.4, Şekil 4.5 ve Şekil 4.6’da verilmiştir.



Şekil 4.4 : Katman parametresi 2, filtre parametresi 8 olan modelin mimarisi.



Şekil 4.5 : Katman parametresi 2, filtre parametresi 16 olan modelin mimarisi.



Şekil 4.6 : Katman parametresi 3, filtre parametresi 8 olan modelin mimarisi.

4.2.3 Evrişim katmanlı ResNet tipi otokodlayıcı

Bu model tipinde girdinin bir evrişim katmanından geçirilmesinin ardından çıktıya yakın katmanlara bağlanması gerçekleştirilmiştir. Böylece her bir pikselin en az bir ağırlık değeri ile çarpma işleminden geçtikten sonra çıkışa gelebilmesi sağlanmıştır.

Kimlik eşlemeli ResNet tipi modellerde girdi verileri herhangi bir matematik işleminden geçirilmeden doğrudan toplama işlemine dahil edilir. Bir kabartı dokusu, kendisine karşılık gelen renk dokusundaki piksel değerlerini içermez. Bu sebeple girdinin bir defalık 1x1 boyutlarında filtrelerden oluşan evrişim katmanından geçirilmesinden sonra toplama işlemine alınacağı bu tip bir model ile deneysel çalışmalara devam edilmiştir.



5. BULGULAR VE TARTIŞMA

Bu tez çalışmasında, deney sonuçlarının kolay bir şekilde karşılaştırılabilmesi için deneylere modelleri tanımlayan kodlar atanmıştır. Standart otokodlayıcı modeller AE (autoencoder), kimlik eşlemeli ResNet tipi otokodlayıcı modeller AI (autoencoder + identity), evrişim katmanlı ResNet tipi otokodlayıcı modeller AC (autoencoder + convolution) olarak belirtilmiştir. Aynı parametreleri alan her tipteki modellere aynı harf ve rakamdan oluşan kodlar verilmiştir. Katman sayısı 2, 3, 4 ve 5 olan modellerin deneyleri sırasıyla A, B, C ve D harfleriyle kodlanmıştır. Filtre sayıları 8, 16, 32 ve 64 ile başlayan modellerin deney kodlarının sonuna ise sırasıyla 1, 2, 3 ve 4 eklenmiştir.

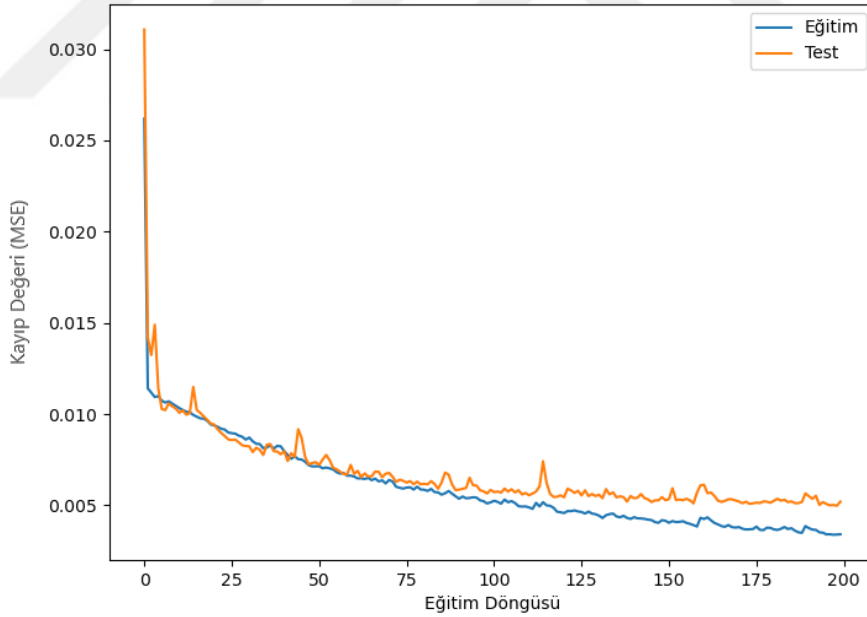
Her biri 200 döngü ile eğitilmiş olan modellerin toplam parametre sayıları, eğitilebilir parametre sayıları, eğitim ve test setleri için en düşük kayıp değerleri ve en yüksek PSNR değerleri karşılaştırılmıştır.

5.1 Standart Otokodlayıcı Modellere Ait Sonuçlar

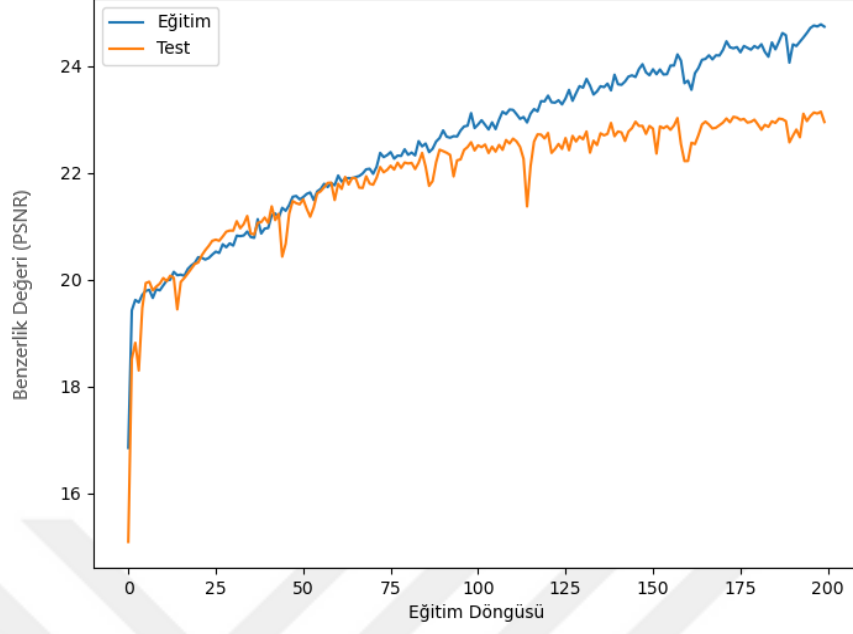
AE kodlu modellere ait deney sonuçları Çizelge 5.1’de gösterilmiştir. Standart otokodlayıcı modellerle yapılan deneyler arasından eğitim setinde en iyi kayıp ve PSNR değerlerini veren model AE-C4’te tespit edilmiştir. Şekil 5.1’de bu modele ait eğitim-kayıp grafiği verilmiştir. Şekil 5.2’de ise modelin eğitim-PSNR grafiği verilmiştir. Test setinde en iyi kayıp ve PSNR değerlerine ulaşılan model AE-B4’tedir. AE-B4 kodlu deneyin modeline ait eğitim-kayıp grafiği Şekil 5.3’te, eğitim-PSNR grafiği ise Şekil 5.4’te verilmiştir. AE-D1’in ise hem eğitim setinde hem de test setinde en kötü değerleri vermiş olduğu görülmektedir. AE-D1 kodlu deneydeki modele ait eğitim-kayıp grafiği Şekil 5.5’teki, eğitim-PSNR grafiği ise Şekil 5.6’daki gibidir.

Çizelge 5.1 : Standart otokodlayıcı tipindeki modellerin eğitim sonuçları.

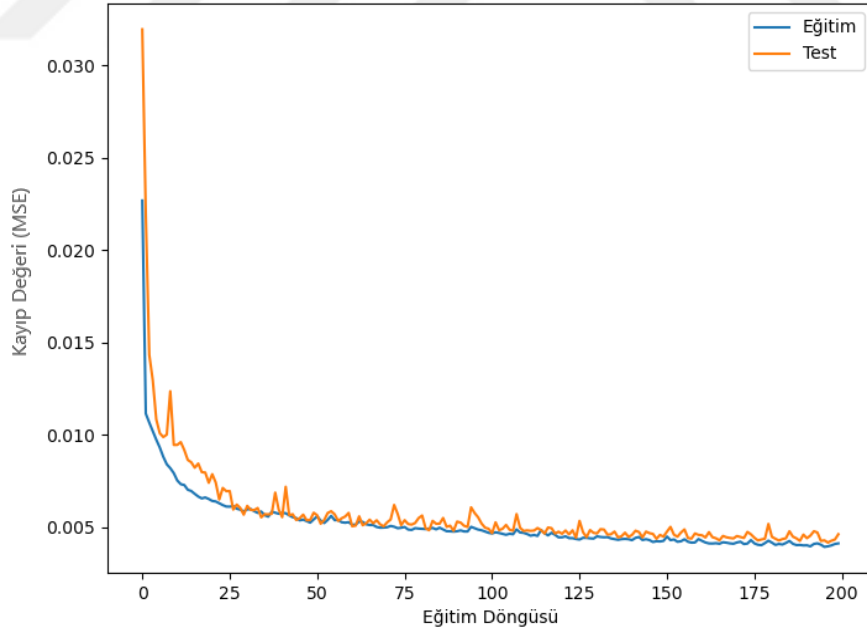
Deney Kodu	Toplam Parametre	Eğitilebilir Parametreler	En Düşük Kayıp (Eğitim Seti)	En Düşük Kayıp (Test Seti)	En Yüksek PSNR (Eğitim Seti)	En Yüksek PSNR (Test Seti)
AE-A1	3015	2967	0.0066	0.0061	21.9662	22.2556
AE-A2	11211	11115	0.0057	0.0054	22.6017	22.8021
AE-A3	43155	42963	0.0050	0.0048	23.1569	23.2905
AE-A4	169251	168867	0.0044	0.0045	23.6656	23.6177
AE-B1	12359	12247	0.0073	0.0069	21.4957	21.7022
AE-B2	48331	48107	0.0059	0.0057	22.4210	22.5334
AE-B3	191123	190675	0.0047	0.0047	23.4044	23.3510
AE-B4	760099	759203	0.0039	0.0042	24.1115	23.9265
AE-C1	49479	49239	0.0083	0.0081	20.9097	20.9701
AE-C2	196299	195819	0.0062	0.0064	22.1768	22.0370
AE-C3	781971	781011	0.0047	0.0100	23.3440	22.7462
AE-C4	3121443	3119523	0.0034	0.0050	24.7799	23.1452
AE-D1	197447	196951	0.0098	0.0095	20.2805	20.3110
AE-D2	787147	786155	0.0071	0.0081	21.5223	21.0208
AE-D3	3143315	3141331	0.0050	0.0074	23.0691	21.3903
AE-D4	12562723	12558755	0.0045	0.0074	23.5438	21.4166



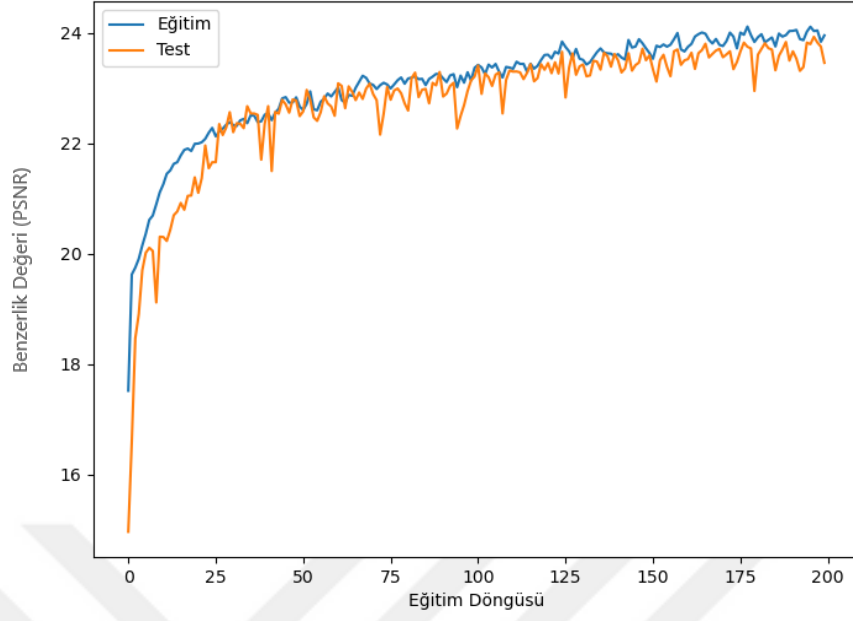
Şekil 5.1 : AE-C4 modelinin eğitim-mse(kayıp) grafiği.



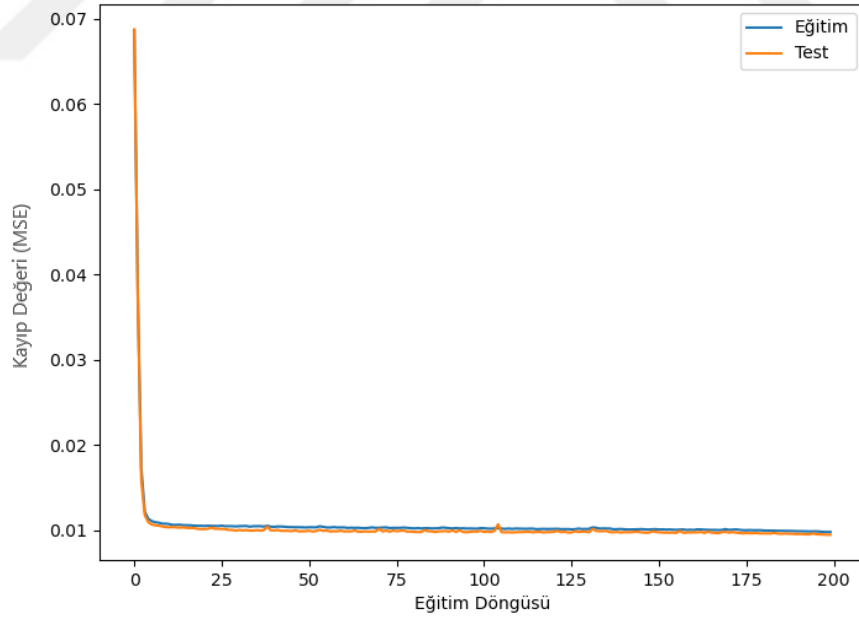
Şekil 5.2 : AE-C4 modeline ait eğitim-PSNR grafiği.



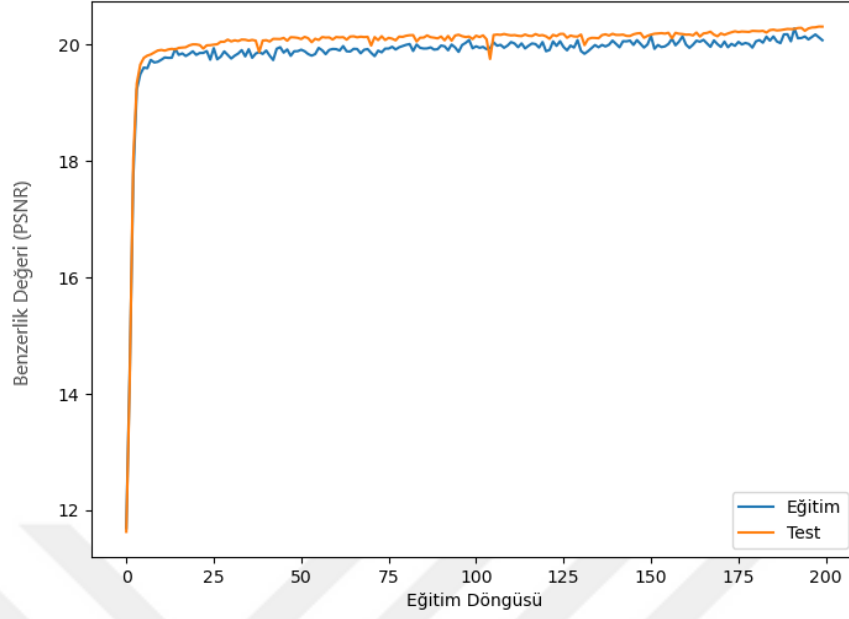
Şekil 5.3 : AE-B4 modeline ait eğitim-mse(kayıp) grafiği.



Şekil 5.4 : AE-B4 modeline ait eğitim-PSNR grafiği.



Şekil 5.5 : AE-D1 modeline ait eğitim-mse(kayıp) grafiği.



Şekil 5.6 : AE-D1 modeline ait eğitim-PSNR grafiği.

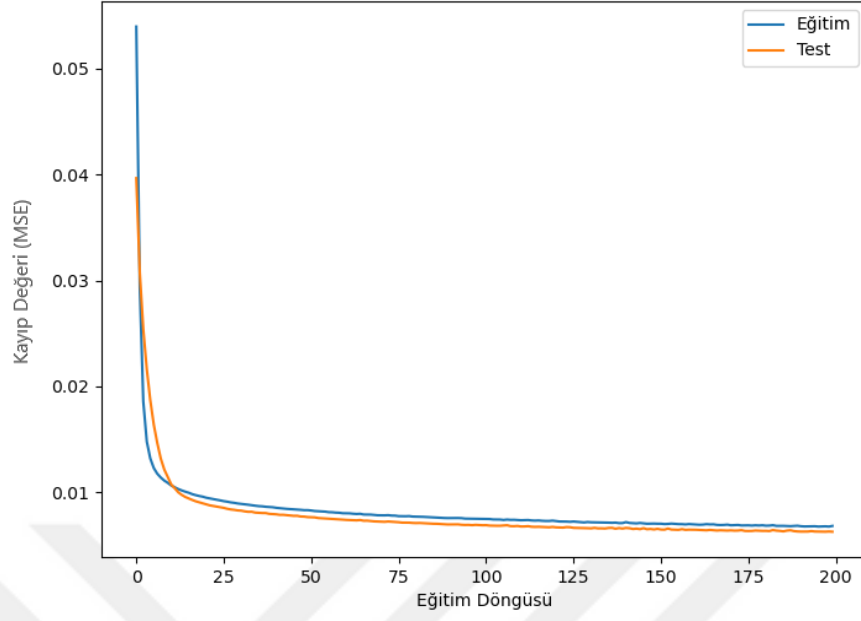
5.2 Kimlik Bağlantılı ResNet Tipi Otokodlayıcı Modellere Ait Sonuçlar

Deneysel çalışmalar Kalıntılı Sinir Ağı (ResNet, Residual Neural Network) tiplerindeki modellerle devam ettirilmiştir. Bir otokodlayıcı ve bu otokodlayıcıyı ilk katmandan son katmana bağlayan bir kimlik bağlantısından oluşan modellerin parametre sayılarıyla birlikte en düşük kayıp ve en yüksek PSNR değerleri Çizelge 5.2’de verilmiştir.

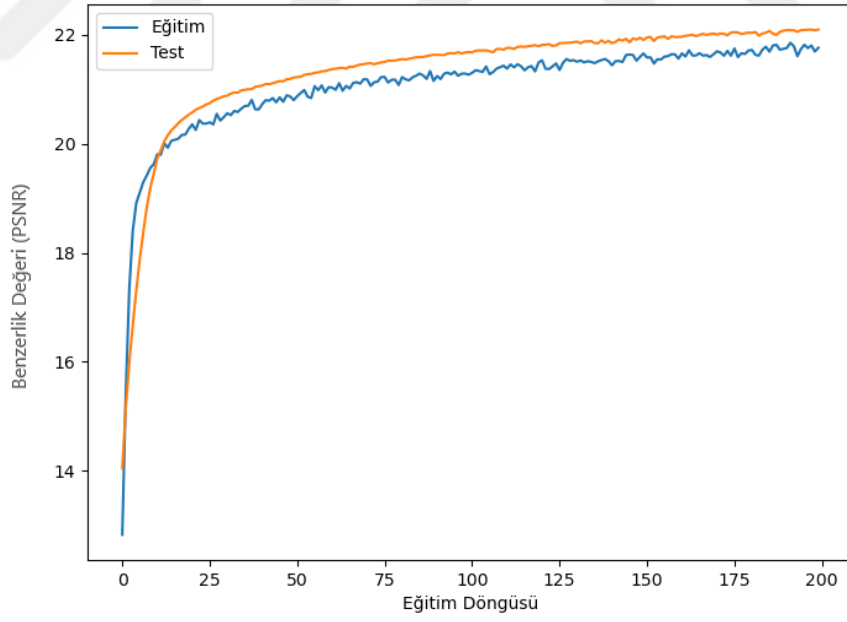
Çizelge 5.2 : Kimlik bağlantılı ResNet tipi otokodlayıcı modellerin eğitim sonuçları.

Deney Kodu	Toplam Parametre	Eğitilebilir Parametreler	En Düşük Kayıp (Eğitim Seti)	En Düşük Kayıp (Test Seti)	En Yüksek PSNR (Eğitim Seti)	En Yüksek PSNR (Test Seti)
AI-A1	3111	3057	0.0068	0.0063	21.8575	22.0992
AI-A2	11307	11205	0.0062	0.0057	22.2538	22.5052
AI-A3	43251	43053	0.0053	0.0051	22.8445	23.0058
AI-A4	169347	168957	0.0049	0.0048	23.1272	23.2971
AI-B1	12455	12337	0.0066	0.0062	21.9638	22.1661
AI-B2	48427	48197	0.0063	0.0060	22.1412	22.2847
AI-B3	191219	190765	0.0056	0.0055	22.6019	22.7260
AI-B4	760195	759293	0.0046	0.0047	23.5335	23.3818
AI-C1	49575	49329	0.0065	0.0062	21.9940	22.1788
AI-C2	196395	195909	0.0062	0.0060	22.2240	22.2729
AI-C3	782067	781101	0.0057	0.0055	22.5870	22.7249
AI-C4	3121539	3119613	0.0041	0.0051	23.9426	23.0527
AI-D1	197543	197041	0.0058	0.0061	22.4147	22.2293
AI-D2	787243	786245	0.0056	0.0059	22.6802	22.3739
AI-D3	3143411	3141421	0.0063	0.0060	22.0728	22.3307
AI-D4	12562819	12558845	0.0035	0.0054	24.6237	22.7364

Kimlik eşlemeli ResNet tipi modeller arasında en kötü değerler AI-A1 deneyinde elde edilmiştir. Buna rağmen test seti değerleri, eğitim seti değerlerinden daha iyidir. Şekil 5.7’de test setine ait kayıp değerinin eğitim setine göre daha düşük olduğu görülmektedir. Test setinden elde edilen PSNR değerinin eğitim setinden yüksek olduğu ise Şekil 5.8’de gösterilmektedir.

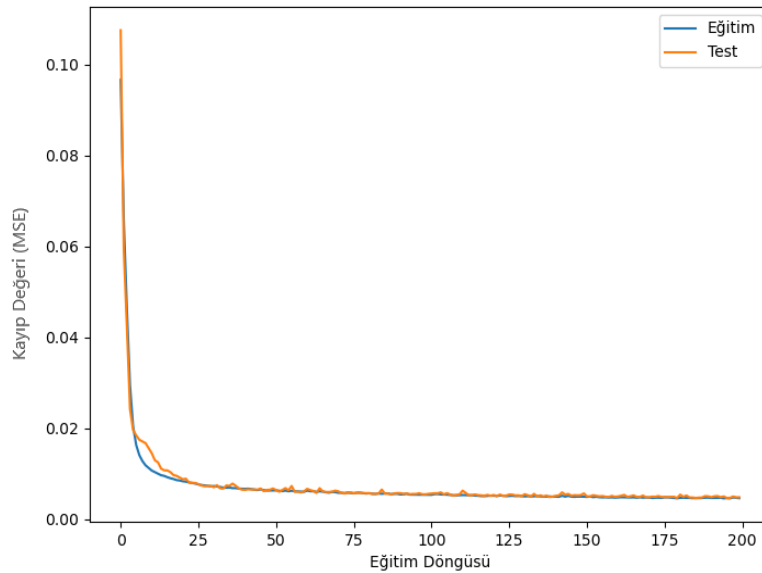


Şekil 5.7 : AI-A1 deneyine ait modelin eğitim-mse(kayıp) grafiği.

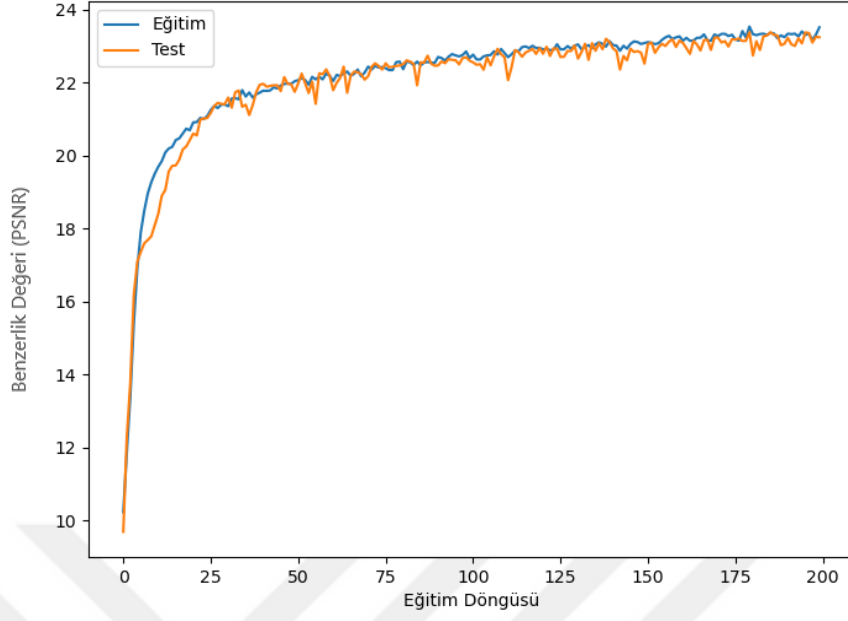


Şekil 5.8 : AI-A1 deneyine ait modelin eğitim-PSNR grafiği.

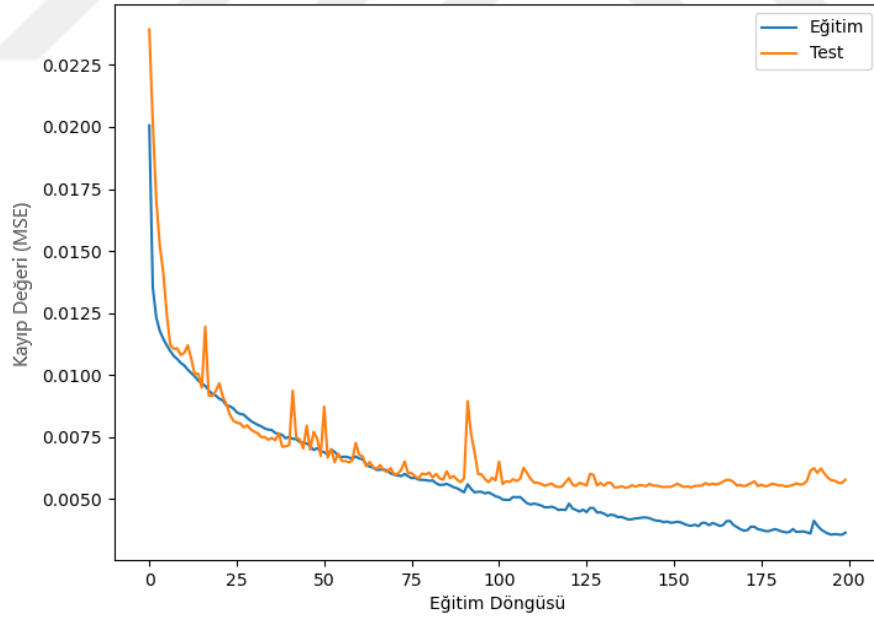
Eđitim seti iin en iyi deęerler AI-D4'te grlmŖtr. Test seti iin en iyi deęerler ise AI-B4'te elde edilmiŖtir. AI-B4 deneyindeki modele ait eđitim-mse(kayıp) grafięi Ŗekil 5.9'da, eđitim-PSNR grafięi Ŗekil 5.10'da grlmektedir. Eđitim ve test setleri iin en kt deęerleri AI-A1 modeli vermektedir. Toplam ve eđitilebilir parametre sayıları yaklaŖık aynı seviyelerde olan farklı tipteki modellerin en kt ve en iyi deęerleri verme sıralarının deęiŖtięi gzlemlenmiŖtir. Standart otokodlayıcı modellerde katman sayısı arttıka modelin eđitim seviyesi dŖerken, kimlik eŖlemeli ResNet tipi aę modelinde en kt model en az katman sayısına sahip olandır. Bu durum kalıntı aęlarda (residual networks) katman sayısının artmasının sonuca olumsuz bir etkisinin bulunmadıęını gstermiŖtir. Fakat eđitim ve test setleri iin iyi sonular veren modellerin farklı olması bir uyumsuzluk gstergesidir. Standart otokodlayıcı tipteki modellerle yapılan deneylerde eđitim ve test setleri iin en iyi sonucu veren modeller arasında 1 katman fark varken; kimlik eŖlemeli ResNet tipi modellerle yapılan deneylerde katman farkının 2 olduęu grlmŖtir. Bu yapıdaki modellerde katman sayısı arttıka eđitim setinin daha iyi sonular verdięi fakat test setinin sabit kaldıęı, yani ezberleme (overfit) yapıldıęı tespit edilmiŖtir. Ŗekil 5.11'de AI-D4 modeline ait eđitim-mse(kayıp) grafięi grlmektedir. Ŗekil 5.12'de aynı modele ait eđitim-PSNR grafięi verilmiŖtir.



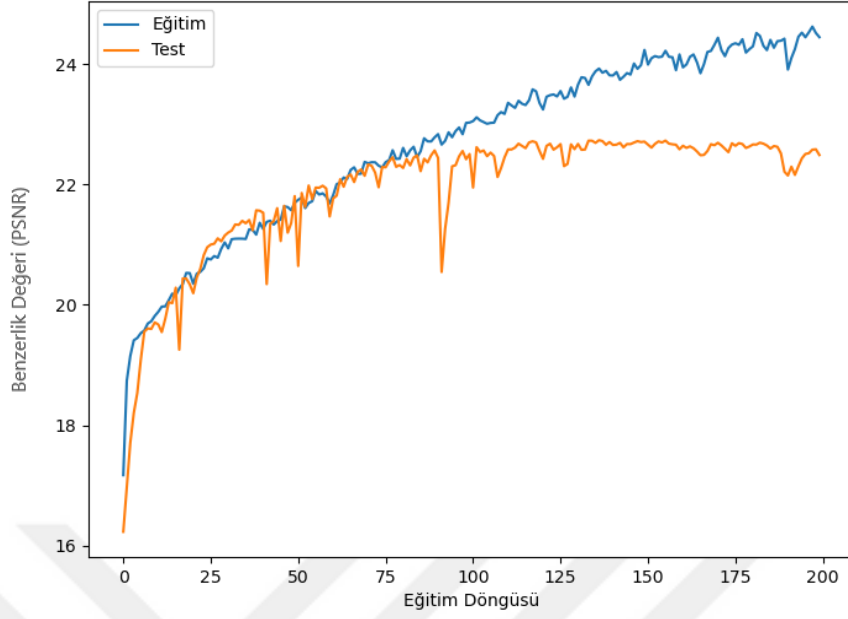
Ŗekil 5.9 : AI-B4 modeline ait eđitim-mse(kayıp) grafięi.



Şekil 5.10 : AI-B4 modeline ait eğitim-PSNR grafiği.



Şekil 5.11 : AI-D4 deneyine ait modelin eğitim-mse(kayıp) grafiği.



Şekil 5.12 : AI-D4 deneyine ait modelin eğitim-PSNR grafiği.

5.3 Evrişim Bağlantılı ResNet Tipi Otokodlayıcı Modellere Ait Sonuçlar

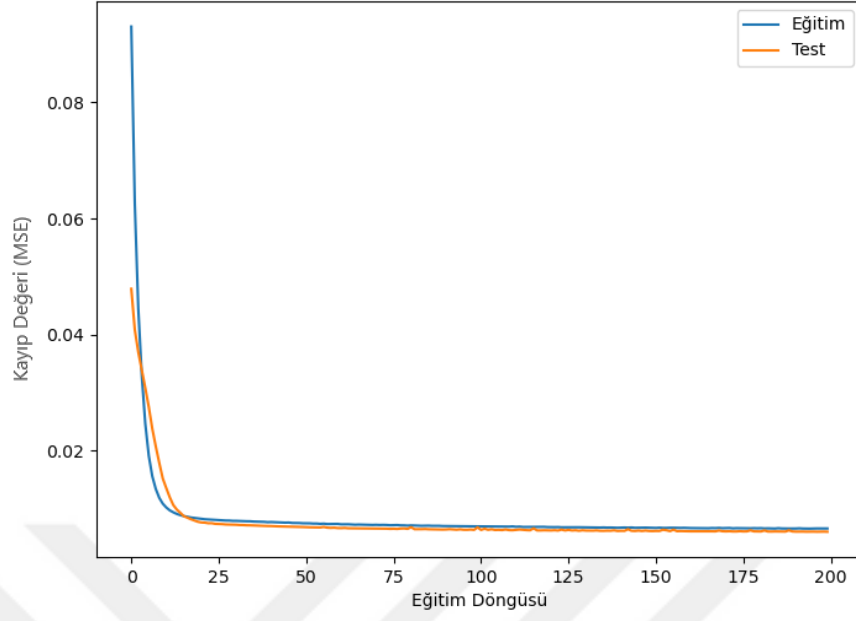
Bir otokodlayıcıyı baştan sona bağlayan bağlantının bir evrişim katmanından geçirildiği modellerle yapılan çalışmaların sonuçları, modellerin parametre sayılarıyla birlikte Çizelge 5.3'te gösterilmiştir.

Çizelge 5.3 : Evrişim bağlantılı ResNet tipi otokodlayıcı modellerin eğitim sonuçları.

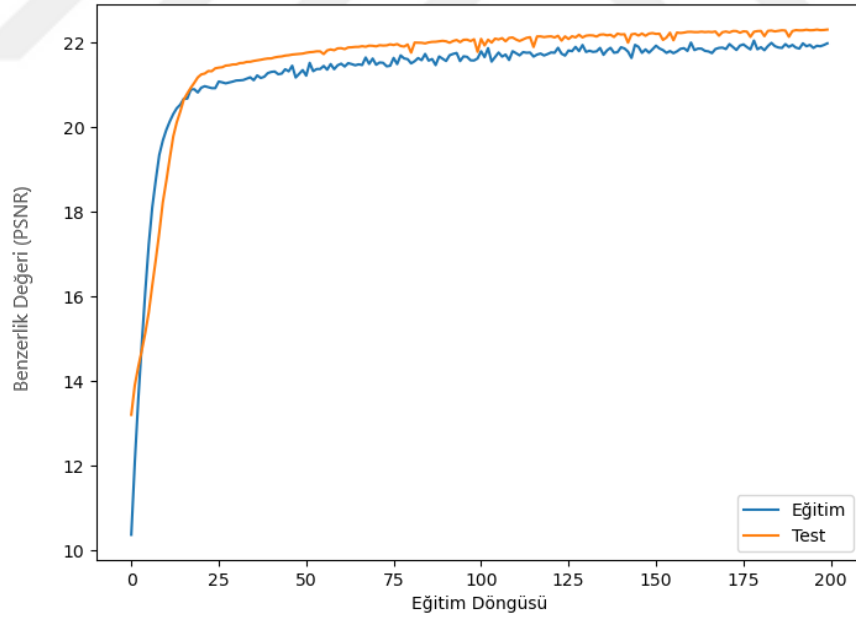
Deney Kodu	Toplam Parametre	Eğitilebilir Parametreler	En Düşük Kayıp (Eğitim Seti)	En Düşük Kayıp (Test Seti)	En Yüksek PSNR (Eğitim Seti)	En Yüksek PSNR (Test Seti)
AC-A1	3135	3075	0.0100	0.0060	22.0419	22.3019
AC-A2	11331	11223	0.0059	0.0056	22.4590	22.6134
AC-A3	43275	43071	0.0055	0.0053	22.7803	22.8843
AC-A4	169371	168975	0.0049	0.0048	23.1501	23.2685
AC-B1	12479	12355	0.0063	0.0058	22.1584	22.4646
AC-B2	48451	48215	0.0055	0.0054	22.5892	22.7931
AC-B3	191243	190783	0.0051	0.0049	23.0970	23.1488
AC-B4	760219	759311	0.0048	0.0047	23.3471	23.3522
AC-C1	49599	49347	0.0061	0.0057	22.2654	22.5160
AC-C2	196419	195927	0.0059	0.0056	22.3785	22.5970
AC-C3	782091	781119	0.0050	0.0049	23.1595	23.1863
AC-C4	3121563	3119631	0.0053	0.0052	22.9286	22.8970
AC-D1	197567	197059	0.0059	0.0057	22.4209	22.5540
AC-D2	787267	786263	0.0056	0.0054	22.7760	22.7870
AC-D3	3143435	3141439	0.0054	0.0054	22.7439	22.7949
AC-D4	12562843	12558863	0.0057	0.0055	22.5676	22.6701

Evrişim bağlantılı ResNet tipi otokodlayıcı modeller arasında AC-A1 kodlu deneyde kullanılan model en kötü sonuçları vermiştir. En düşük kayıp ve en yüksek PSNR değerleri ise AC-B4 deneyinden elde edilmiştir. AC kodlu deneylerin sonuçları incelendiğinde modelin otokodlayıcı kısmına ait evrişim katmanlarının filtre sayılarının artırılmasının sonuca olumlu etkisi olduğu gözlemlenmiştir. Katman sayılarının artışının sonuca etkisi ise 3 katman parametresine kadar olumlu iken daha sonrasında olumsuz dönmüştür. 3 katmanlı kodlayıcı ve 3 katmanlı çözücüye sahip 64 adet filtreden başlayan otokodlayıcı ve evrişim bağlantılı model hem eğitim setinde hem de test setinde en iyi sonucu vermiştir. Modelin, test setine ait kayıp değeri eğitim setine ait kayıp değerinden düşüktür. Aynı modelin, test setine ait PSNR değeri ise eğitim setine ait PSNR değerinden yüksektir. Bu, modelin ezberleme yapmadığının göstergesidir.

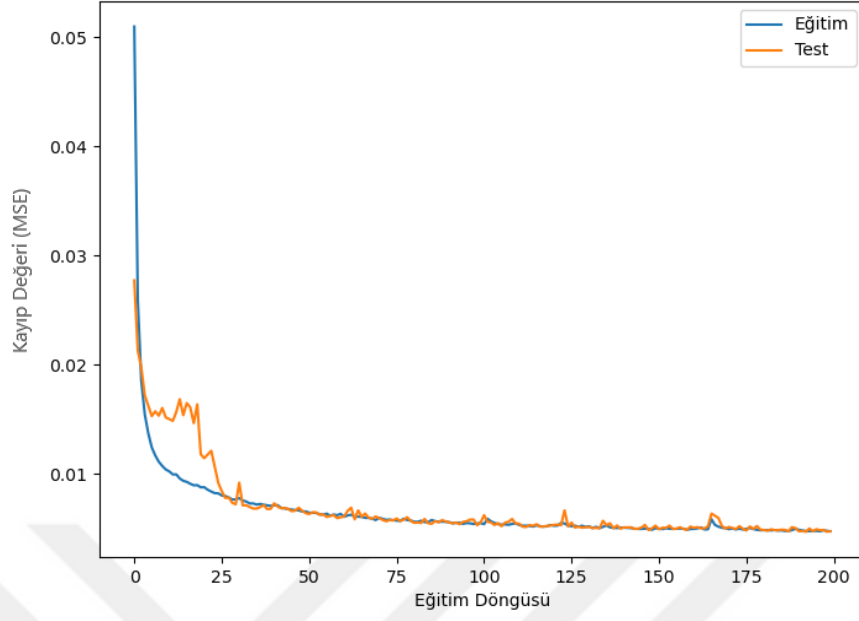
AC-A1 kodlu deneye ait modelin eğitim-mse(kayıp) grafiği Şekil 5.13'tedir. Şekil 5.14'te de aynı modele ait eğitim-PSNR grafiği bulunmaktadır. AC-B4 modeline ait eğitim grafikleri ise Şekil 5.15'te ve Şekil 5.16'da verilmiştir.



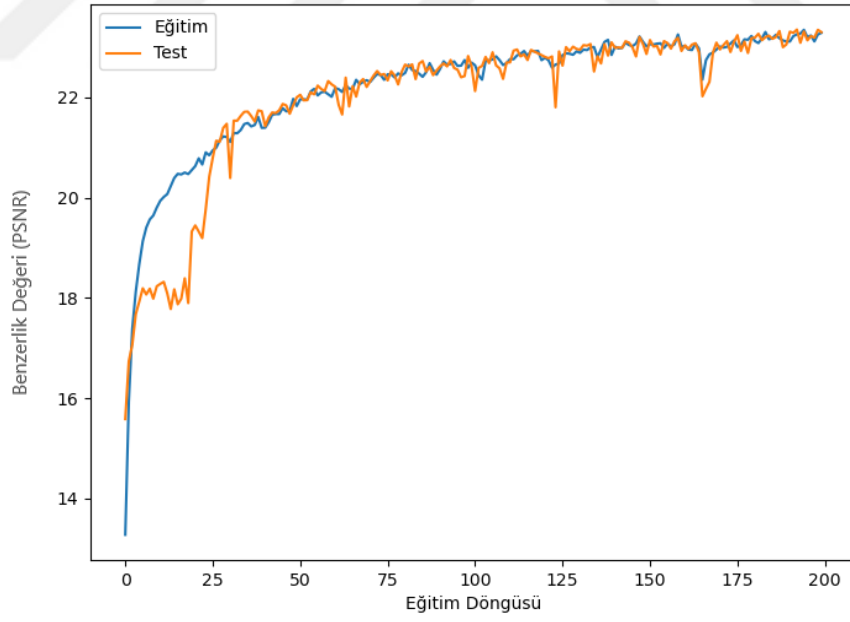
Şekil 5.13 : AC-A1 modeline ait eğitim-mse(kayıp) grafiği.



Şekil 5.14 : AC-A1 modeline ait eğitim-PSNR grafiği.



Şekil 5.15 : AC-B4 modeline ait eğitim-mse(kayıp) grafiği.



Şekil 5.16 : AC-B4 modeline ait eğitim-PSNR grafiği.

5.4 Genel Deęerlendirme

Bütün deneyler arasından test setinde en iyi kayıp ve PSNR deęerlerine sahip olan model AE-B4 kodlu deneyde incelenen modeldir. Standart otokodlayıcı tipindeki bu modelin katman parametresi 3'tür ve filtre sayısı 64'ten başlamaktadır. Model, test setinde en iyi deęerleri vermiş olsa da eğitim setinde kayıp deęeri daha düşük ve PSNR deęeri ise daha yüksektir. Bu, modelin yeterince iyi olmadığını, girdi verilerini ezberlemek üzere olduğunu göstermektedir. AE kodlu deneyler arasında eğitim setinde en iyi deęerleri veren AE-C4 kodlu modelde ise test setine ait kayıp deęerinde artış ve PSNR deęerinde düşüş gözlemlenmiştir. Model, katman sayısı arttıkça, eğitilebilir parametrelerindeki artışa rağmen ezberleme eğilimi göstermektedir.

Bütün deneyler arasında en kötü sonucun gözlemlendięi deney ise yine standart otokodlayıcı tipinde olan AE-D1 kodlu modele aittir. Model ezberleme yapmış olmasa da PSNR deęeri ortalamanın çok altındadır.

6. SONUÇLAR VE ÖNERİLER

3B grafiklerin üretiminde ve işlenmesinde prosedürel üretim teknikleri gittikçe yaygınlaşmaktadır. Prosedürel üretim sayesinde üretimi yapılacak olan desenlerin, materyallerin veya 3B modellerin genel bir formu matematik fonksiyonlarına dönüştürülür. İsteğe bağlı ya da rastgele parametreler ile bu fonksiyonlar çalıştırılarak, tasarımcıların ve geliştiricilerin günlerini harcayabileceği grafik varlıkları birkaç saniye içerisinde ortaya çıkarılabilmektedir.

3B grafiklerin kullanıldığı oyun, animasyon ve film endüstrilerinde maliyetleri çokça düşürse de, prosedürel üretim fonksiyonlarının geliştirilme aşaması meşakkatli olabilmektedir. Her çeşit nesne için farklı fonksiyonların tasarlanması veya yakınsanması gerekmektedir.

Evrensel yakınsama teoremine göre, yapay sinir ağları ile belirli girdilere (x) belirli çıktılar (y) veren fonksiyonun ($y = f(x)$) yakınsanabilmesi mümkün olmaktadır. Yapay sinir ağları kullanılarak prosedürel üretim tekniklerinde kullanılabilecek fonksiyonlar hesaplanabilir.

3B grafiklerde yoğunlukla kullanılan kabartı dokuları gelişen teknoloji ile giderek daha optimize edilmiş formlara dönüşmektedir. İlk 3B grafik işleyicilerinde gri tonlamalı yükselti dokuları (height map) kullanılırken, 3B modellerin yüzeylerindeki ayrıntıların işlenebilmesini sağlayan kaplamalar zamanla bu ayrıntıların 3B uzaydaki vektörlerini içeren dokular (normal map) haline gelmiştir.

Bir 3B modelin renk kaplaması (albedo/color map) ile kabartı dokusu (height map/normal map) arasında doğrudan bir ilişki bulunmamaktadır. Yani renk değerleri ile yüzey ayrıntısı bilgileri arasında basit bir matematik fonksiyonu ile ifade edilebilecek bir bağlantı yoktur. Fakat yüzeyin tipi, 3B modelin geometrisi veya dokuların biçimi ile kabartı dokusu arasında ilişki kurulabilir.

Bu tez çalışmasında, renk dokusu bilinen 3B modellerin kabartı dokularını tahmin edilebilecek yapay sinir ağı modelleri önerilmiş ve bu modeller birbirleriyle karşılaştırılmıştır. Deneysel çalışmalar kapsamında, çeşitli filtre ve katman sayıları ile farklı tiplerde otokodlayıcı modeller üretilmiştir.

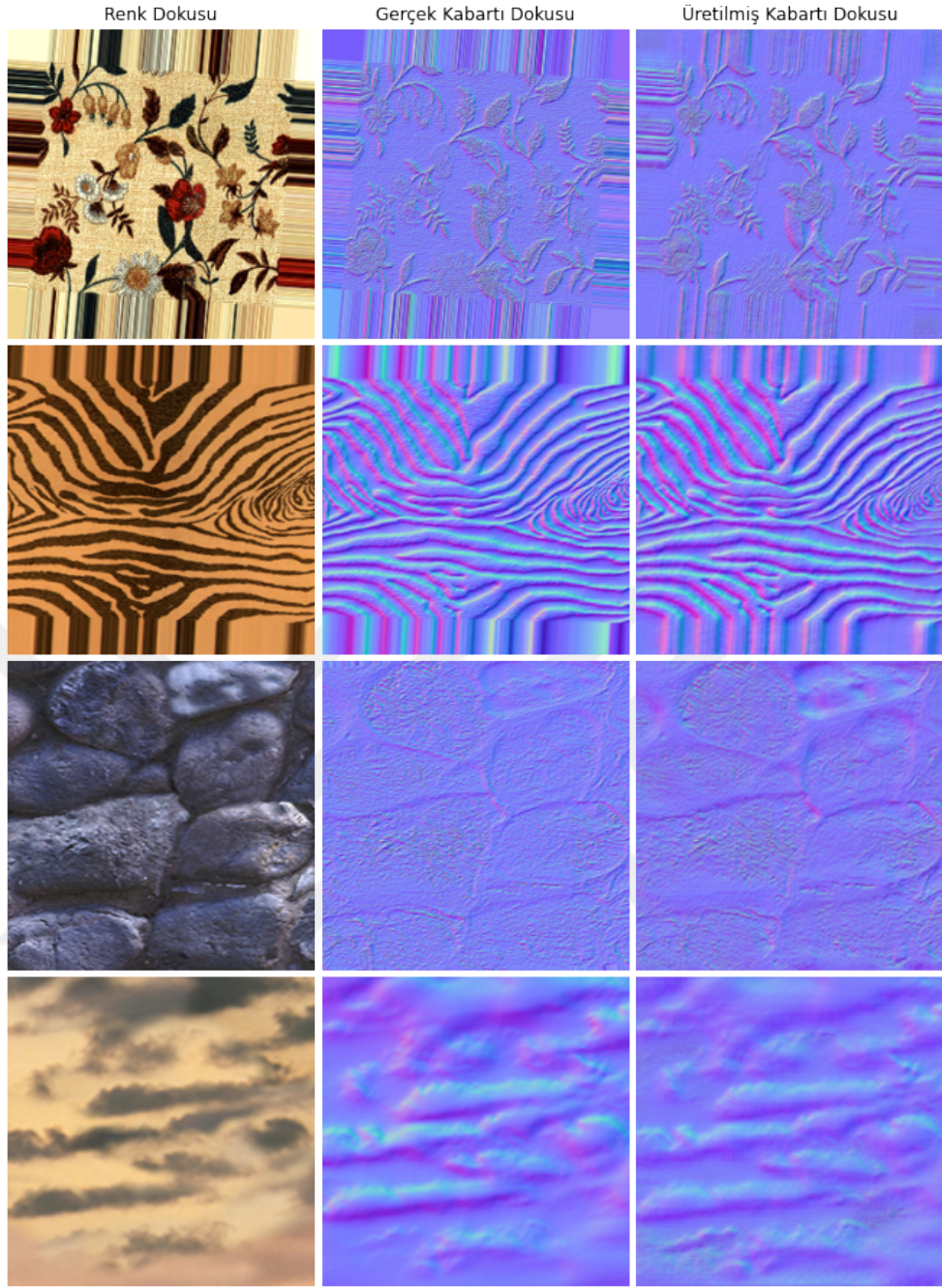
Yapılan deneylerin sonucunda, standart otokodlayıcılarda katman ve filtre sayıları arttıkça eğitim kalitesinde düşüş meydana geldiği görülmüştür. Kaybolan gradyan problemi sebebiyle belirli bir seviyeden sonra girdileri çıktılara dönüştürecek olan fonksiyon yakınsanamamaktadır.

Katman sayısı arttıkça, bir yapay sinir ağındaki nöron sayısı da artmaktadır. Teorik olarak; nöron sayılarındaki artışın sonuca olumlu etkisinin olması beklenmektedir. Fakat bu tez çalışmasında standart otokodlayıcı modellerle kaplama dokularının üretiminde, katman sayısındaki artışın sonuca etkisinin olumsuz yönde olduğu görülmüştür. Oluşturulan yapay sinir ağı modelleri, kaybolan gradyan problemi sebebiyle kompleks fonksiyonları yakınsamakta güçlük çekmektedir. Fazla katman kullanımı, modelin eğitimi esnasında ağıın sonundan (çıkıtı katmanına yakın nöronlardan) başına (girdi katmanına yakın nöronlara) doğru geri yayılım sırasında yaptığı işlemlerin etkisinin giderek azalmasına sebep olmaktadır.

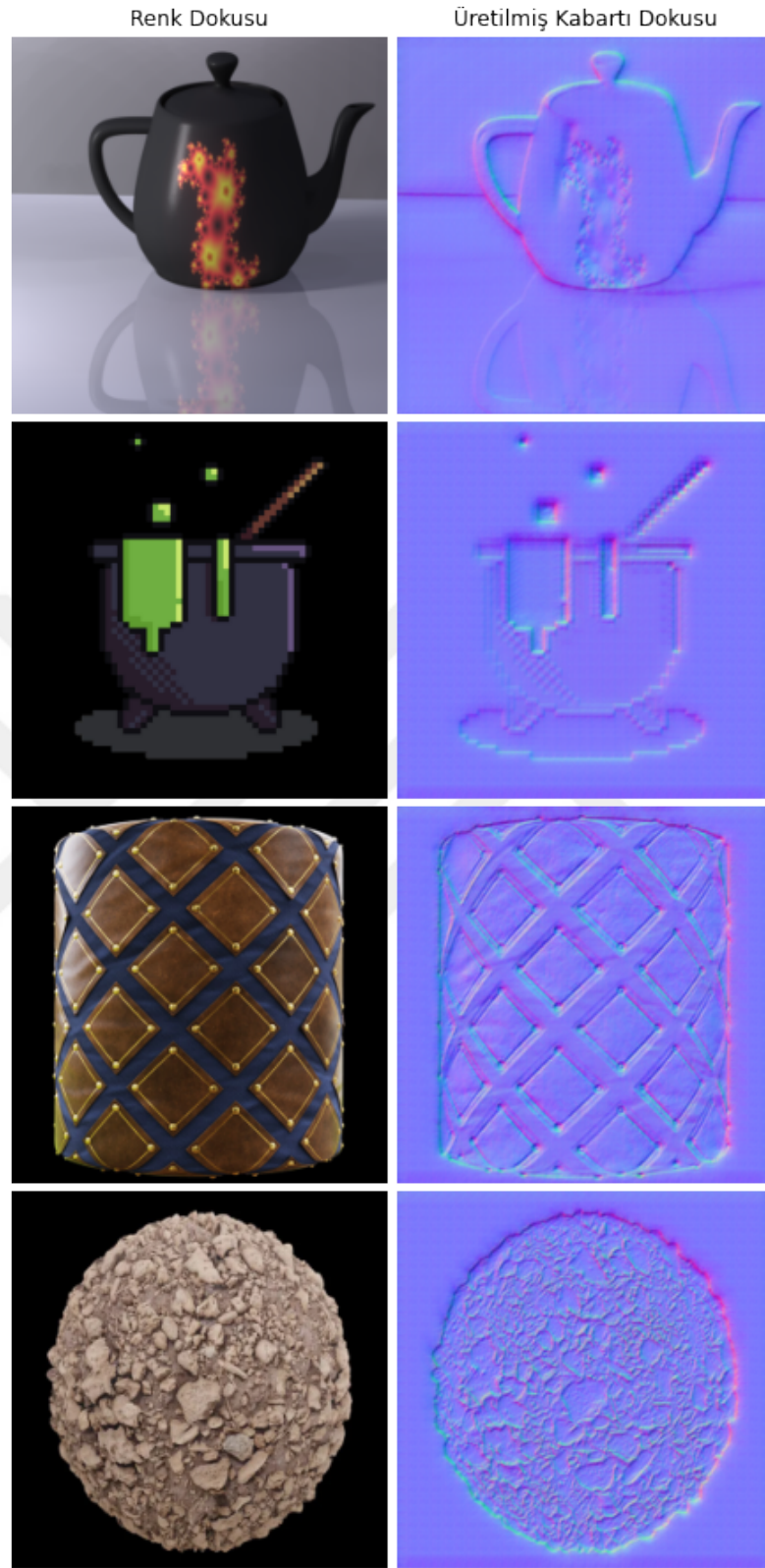
Otokodlayıcı modellerde kalıntılı modellerde olduğu gibi hatırlatıcı bir bağ kullanmanın olumlu etkisi olduğu gözlemlenmiştir. Hatırlatıcı bağlar, kalıntılı modellerde kaybolan gradyan problemine karşı çözüm olarak kullanılabilir. Hatırlatıcı bağlar, kalıntılı modellerde kaybolan gradyan problemine karşı çözüm olarak kullanılabilir.

Basit bir evrişim katmanından geçirilen hatırlatıcı bağın etkisi, evrişim katmanı olmayan bağlara göre daha fazladır. Bu evrişim katmanı eğitim kalitesini artırmakla beraber ezberlemenin önüne de geçmektedir.

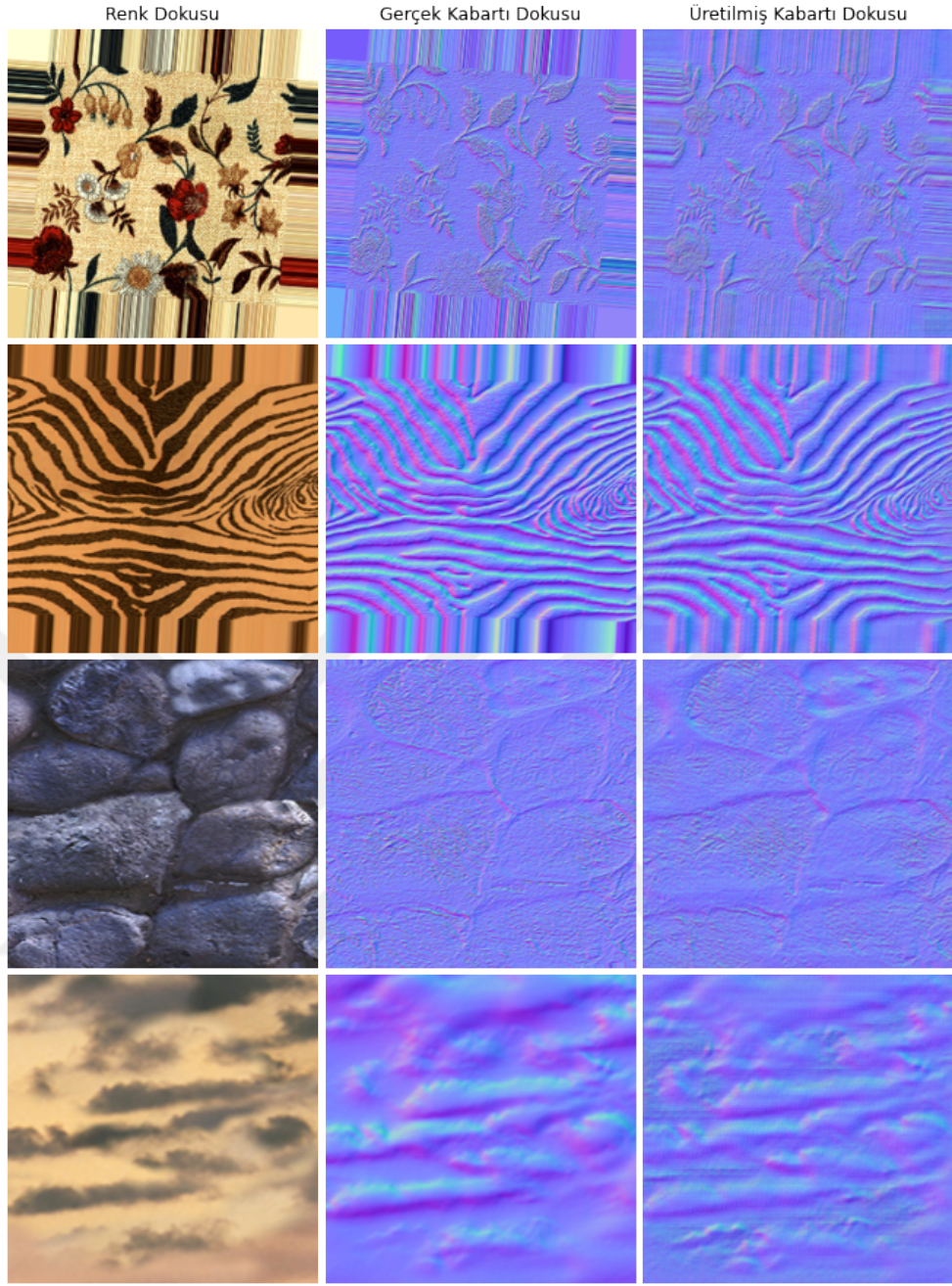
Tez çalışması kapsamında yapılan deneylerde üç farklı tipte üretilen modellerden her tip için test setinde en iyi değerleri verenler çeşitli görüntülerle çalıştırılmıştır. Modellerin çıktı kaliteleri algısal olarak da karşılaştırılmıştır. Modellerin sadece benzer kategorilerdeki görüntülerde iyi sonuç verme ihtimaline karşı test seti dışından, yani bir materyal kaplaması olmayan görüntüler de bu karşılaştırma aşamasına dahil edilmiştir.



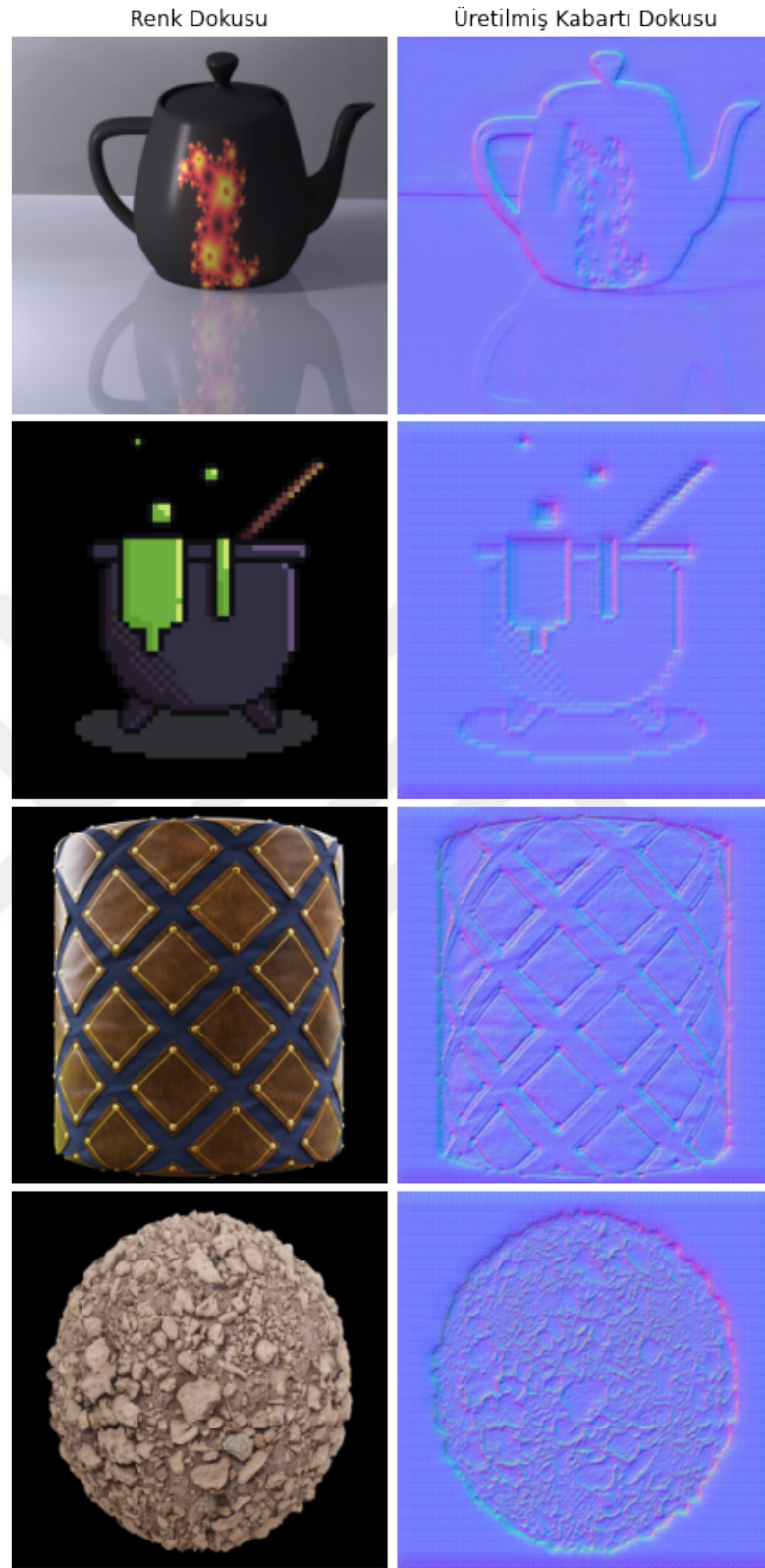
Şekil 6.1 : En iyi standart otokodlayıcı modelin test seti örneklerinden bazıları için verdiği çıktılar.



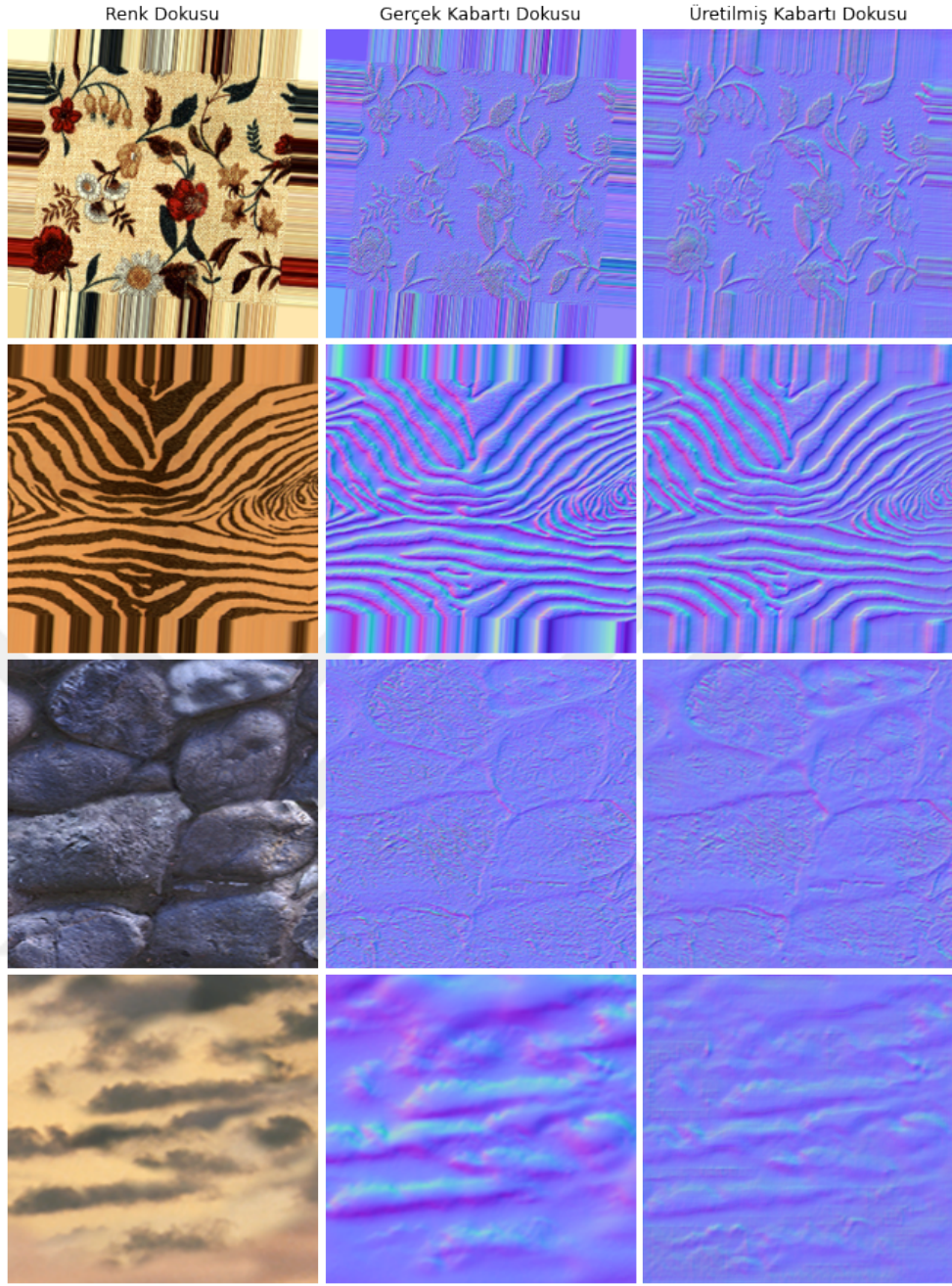
Şekil 6.2 : En iyi standart otokodlayıcı modelin test seti dışından rastgele görüntüler için verdiği çıktılar.



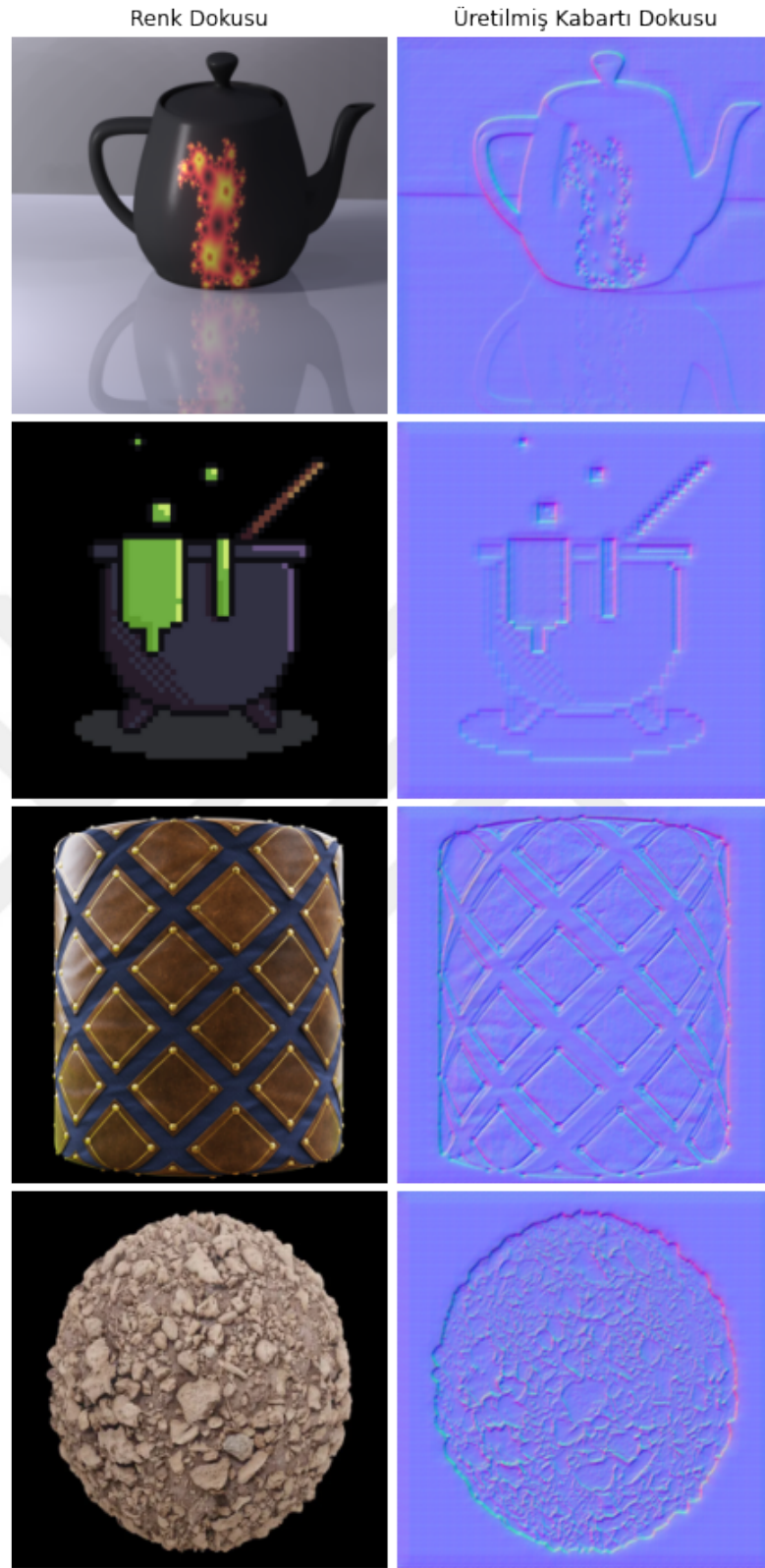
Şekil 6.3 : En iyi kimlik bağlantılı ResNet tipi otokodlayıcı modelin test seti örneklerinden bazıları için verdiği çıktılar.



Şekil 6.4 : En iyi kimlik bağlantılı ResNet tipi otkodlayıcı modelin test seti dışından rastgele görüntüler için verdiği çıktılar.



Şekil 6.5 : En iyi evrişim katmanlı ResNet tipi otkodlayıcı modelin test seti örneklerinden bazıları için verdiği çıktıları.



Şekil 6.6 : En iyi evrişim katmanlı ResNet tipi otokodlayıcı modelin test seti dışından rastgele görüntüler için verdiği çıktılar.

Bu tez çalışmasında bahsedilmiş olan deneyler, veri setinin daha fazla çeşitlendirilmesiyle tekrar sınanabilir. Bu tez çalışması sırasında; morfolojik döndürme işlemi görüntülere kısıtlı miktarda uygulanmıştır. Yatay çevirme, dikey çevirme gibi işlemler ise, veri setindeki bilgileri bozacağı için uygulanamamıştır. Veri çeşitlendirme çalışmalarında uygulanan morfolojik işlemlerin, kabartı dokularında da kullanabileceği versiyonları üretilebilirse çeşitlendirme aşamasındaki kısıtlamalar kalkabilir ve modeller daha iyi eğitilebilir. Alternatif olarak; 3B modelleme ve tasarım uygulamaları kullanılarak düz bir yüzeye (plane) rastgele materyaller uygulanıp, yüzeyin farklı kamera açılarıyla renk ve kabartı dokularının işlenmesi (render) ve kaydedilmesi ile bu alanda yapılan çalışmalar için veri çeşitliliği sağlanabilir.

Bu çalışmada, sentezlenmiş görüntülerin karşılaştırılması için PSNR ölçümü kullanılmıştır. PSNR metriği, MSE ile ters orantılıdır. PSNR ve MSE arasında bulunan bu ilişkiden dolayı, modeller inşa edilirken kayıp fonksiyonu olarak MSE seçilmiştir. Model eğitilirken, MSE değeri minimize edilmeye çalışılmaktadır. MSE küçüldükçe de PSNR değeri yükselmektedir, yani çıktı görüntülerinin kalitesi artmaktadır. Çalışmalar sırasında yararlanılmış olan Python kütüphanelerinde (Tensorflow/Keras) MSE'nin kullanıma hazır halde olması deneyleri kolaylaştırmıştır. Bu çalışma kapsamında geliştirilen ve denenilen modeller, kabartı dokularını, normal dokuları (normal map) olarak üretmektedir. Normal dokularında piksel değerlerinin asıl işlevi renk belirtmek değil, görüntü kanallarının (örneğin; RGB) üç boyutlu uzayda birer eksene (örneğin; XYZ) tekabül edeceği şekilde yön vektörleri oluşturmalarıdır. Bu yön vektörleri, üç boyutlu modelde denk geldikleri yüzeyin yönünü (normal vektörü) belirtmektedir. Yön vektörlerinin karşılaştırılması için doğrudan değer farklarına bakmak modelin yeterince iyi eğitilememesine sebep olabilir. Yön belirten vektör, açı gibi değerler için açısal hata kullanmak modelin daha iyi eğitilmesini sağlayabilir. Daha iyi sonuçlar elde edebilmek için Ortalama Açısal Hata (MAE), kayıp fonksiyonu olarak kullanılabilir. Sentez görüntü ve gerçek görüntü arasındaki kalite farkını incelemek için ise MSE yerine MAE kullanılarak PSNR benzeri bir metrik oluşturulabilir.

Ek olarak; kabartı dokularındaki büyük kabartılarla beraber küçük kabartıların da

yüksek kalitede tahminlenebilmesi için farklı çerçeve boyutundaki filtrelere sahip evrişim katmanlarının bir arada kullanılması denenebilir. Girdi katmanının farklı parametrelere sahip evrişim katmanlarına paralel olarak bağlandığı Inception tipinde modeller üretilebilir.



KAYNAKLAR

- Aittala, M., Aila, T. & Lehtinen, J.** (2016). Reflectance modeling by neural texture synthesis. *ACM Transactions on Graphics (ToG)*, 35(4), 1-13.
- Asselin, L.-P., Laurendeau, D. & Lalonde, J.-F.** (2020). Deep SVBRDF estimation on real materials. In *2020 International Conference on 3D Vision (3DV)*, p. 1157-1166.
- Blinn, J. F.** (1978). Simulation of wrinkled surfaces. *ACM SIGGRAPH computer graphics*, 12(3), 286-292.
- Chen, G., Han, K. & Wong, K.-Y. K.** (2018). PS-FCN: A flexible learning framework for photometric stereo. In *Proceedings of the European conference on computer vision (ECCV)*, p. 3-18.
- Chrislb** (2005). Diagram of an artificial neuron. https://commons.wikimedia.org/wiki/Artificial_neural_network
- Cignoni, P., Ganovelli, F., Gobbetti, E., Marton, F., Ponchio, F. & Scopigno, R.** (2005). Batched multi triangulation. In *VIS 05. IEEE Visualization, 2005.*, p. 207-214.
- Cignoni, P., Montani, C., Rocchini, C. & Scopigno, R.** (1998). A general method for preserving attribute values on simplified meshes. In *Proceedings Visualization'98 (Cat. No. 98CB36276)*, p. 59-66.
- Cohen, J., Olano, M. & Manocha, D.** (1998). Appearance-preserving simplification. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, p. 115-122.
- Deschaintre, V.** (2019). Material acquisition using deep learning. In *SIGGRAPH Asia 2019 Doctoral Consortium* (pp. 1-4).
- Duchi, J., Hazan, E., & Singer, Y.** (2011). Adaptive subgradient methods for online learning and stochastic optimization. *Journal of machine learning research*, 12(7).
- Duha, V.** (2021). Blender 3D normal maps standard is essential if you're just getting started. <https://medium.com/blender-3d/blender-3d-normal-maps-standard-is-essential-if-youre-just-getting-started-fa8ec6dd5856>
- Gao, D., Li, X., Dong, Y., Peers, P., Xu, K. & Tong, X.** (2019). Deep inverse

rendering for high-resolution SVBRDF estimation from an arbitrary number of images.. *ACM Trans. Graph.*, 38(4), 134-1.

- Giacomello, E., Lanzi, P. L. & Loiacono, D.** (2018). Doom level generation using generative adversarial networks. In *2018 IEEE Games, Entertainment, Media Conference (GEM)*, p. 316-323.
- He, K., Zhang, X., Ren, S. & Sun, J.** (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, p. 770-778.
- Huang, G., Liu, Z., Van Der Maaten, L. & Weinberger, K. Q.** (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, p. 4700-4708.
- Ikehata, S.** (2018). CNN-PS: CNN-based photometric stereo for general non-convex surfaces. In *Proceedings of the European conference on computer vision (ECCV)*, p. 3-18.
- Kingma, D. P., & Ba, J.** (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Krishnamurthy, V. & Levoy, M.** (1996). Fitting smooth surfaces to dense polygon meshes. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, p. 313-324.
- Krizhevsky, A., Sutskever, I. & Hinton, G. E.** (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
- 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. & Cai, J.** (2007). Robust transmission of JPEG2000 encoded images over packet loss channels. In *2007 IEEE International Conference on Multimedia and Expo*, p. 947-950.
- Li, X., Dong, Y., Peers, P. & Tong, X.** (2017). Modeling surface appearance from a single photograph using self-augmented convolutional neural networks. *ACM Transactions on Graphics (ToG)*, 36(4), 1-11.
- Lucas Jr, H. C. & Goh, J. M.** (2009). Disruptive technology: How Kodak missed the digital photography revolution. *The Journal of Strategic Information Systems*, 18(1), 46-55.
- Pedersen, L.** (2018). Pixar One Thirty. <https://renderman.pixar.com/pixar-one-thirty>
- Peng, X. B., Abbeel, P., Levine, S. & van de Panne, M.** (2018). Deepmimic: Example-guided deep reinforcement learning of physics-based

- character skills. *ACM Transactions on Graphics (TOG)*, 37(4), 1-14.
- Qi, X., Liao, R., Liu, Z., Urtasun, R. & Jia, J.** (2018). Geonet: Geometric neural network for joint depth and surface normal estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, p. 283-291.
- Robbins, H., & Monro, S.** (1951). A stochastic approximation method. *The annals of mathematical statistics*, 400-407.
- Ruder, S.** (2016). An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*.
- Rumelhart, D. E., Hinton, G. E. & Williams, R. J.** (1985). Learning internal representations by error propagation. California Univ San Diego La Jolla Inst for Cognitive Science.
- Santo, H., Samejima, M., Sugano, Y., Shi, B. & Matsushita, Y.** (2017). Deep photometric stereo network. In *Proceedings of the IEEE international conference on computer vision workshops*, p. 501-509.
- Simonyan, K. & Zisserman, A.** (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Sisson, D.** (2018). Pixar One Twenty Eight. <https://renderman.pixar.com/pixar-one-twenty-eight>
- Song, E. & Chang, M.** (2020). Photometric Stereo Using CNN-based Feature-Merging Network. In *2020 20th International Conference on Control, Automation and Systems (ICCAS)*, p. 865-868.
- Su, W., Du, D., Yang, X., Zhou, S. & Fu, H.** (2018). Interactive sketch-based normal map generation with deep neural networks. *Proceedings of the ACM on Computer Graphics and Interactive Techniques*, 1(1), 1-17.
- Summerville, A. & Mateas, M.** (2016). Super mario as a string: Platformer level generation via lstms. *arXiv preprint arXiv:1603.00930*.
- Szegedy, C. et al.** (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, p. 1-9.
- Thomos, N., Boulgouris, N. V. & Strintzis, M. G.** (2005). Optimized transmission of JPEG2000 streams over wireless channels. *IEEE Transactions on image processing*, 15(1), 54-67.
- Torrado, R. R., Khalifa, A., Green, M. C., Justesen, N., Risi, S. & Togelius, J.** (2020). Bootstrapping conditional gans for video game level generation. In *2020 IEEE Conference on Games (CoG)*, p. 41-48.
- Volz, V., Schrum, J., Liu, J., Lucas, S. M., Smith, A. & Risi, S.** (2018). Evolving

mario levels in the latent space of a deep convolutional generative adversarial network. In *Proceedings of the genetic and evolutionary computation conference*, p. 221-228.

Woodham, R. J. (1980). Photometric method for determining surface orientation from multiple images. *Optical engineering*, 19(1), 139-144.

Zeiler, M. D. (2012). Adadelta: an adaptive learning rate method. *arXiv preprint arXiv:1212.5701*.

Zhou, X., Zhong, G., Qi, L., Dong, J., Pham, T. D. & Mao, J. (2017). Surface height map estimation from a single image using convolutional neural networks. In *Eighth International Conference on Graphic and Image Processing (ICGIP 2016)* 10225, p. 1022524.



ÖZGEÇMİŞ

Ad-Soyad : Muhammed Ömer Faruk SELVİ

Doğum Tarihi ve Doğum Yeri :

E-posta :

ÖĞRENİM DURUMU:

- **Lisans** : 2017, Anadolu Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği

MESLEKİ DENEYİM:

- Bursa Teknik Üniversitesi (Araştırma Görevlisi), 19.11.2019-halen çalışıyor