



T.C.

SELÇUK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

EVREŞİMSEL SİNİR AĞLARI TABANLI OTOMATİK GÖRÜNTÜ RENKLENDİRME

Omar Abdulwahhab Othman OTHMAN

YÜKSEK LİSANS TEZİ

Bilgisayar Mühendisliği Anabilim Dalı

Nisan-2020

KONYA

Her Hakkı Saklıdır

ÖZET

YÜKSEK LİSANS TEZİ

EVREŞİMSEL SİNİR AĞLARI TABANLI OTOMATİK GÖRÜNTÜ RENKLENDİRME

Omar Abdulwahhab Othman OTHMAN

Selçuk Üniversitesi Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Betül UZBAŞ

2020, 57 Sayfa

Jüri
Dr. Öğr. Üyesi Betül UZBAŞ
Doç. Dr. Humar Kahramanlı ÖRNEK
Doç. Dr. Gülay TEZEL

Otomatik görüntü renklendirme, herhangi bir kullanıcı çabası olmadan girdi olarak alınan bir gri tonlamalı görüntüye renk ekleme görevidir. Gri tonda resimlerin gerçek renk değerlerinin bilinmiyor olması ve nesnelerin birçok renge ve renk tonuna sahip olması renklendirme işlemini zorlaştırmaktadır. Örneğin plastik nesneler birçok renge sahip olabilmekte ya da ağaçlar çeşitli tonlarda yeşil renge sahip olabilmektedir. Bu tez çalışmasında, gri tondaki görüntülerin Evrişimsel Sinir Ağı (ESA) kullanılarak otomatik renklendirilmesi önerilmiştir. ESA ve ImageNet veri setindeki resimler kullanılarak model geliştirilmiştir. Geliştirilen model kullanılarak girdi olarak alınan gri tondaki resmin rengi tahmin edilmektedir. Çalışmada renk uzayı olarak Lab Renk uzayı kullanılmaktadır. Model giriş olarak L kanalını almakta, çıkış olarak ab kanalını vermektedir. ImageNet veri setinden rassal seçilmiş 39,604 görüntü içeren veri seti ESA modelini eğitmek için % 80 eğitim ve % 20 validasyon olacak şekilde bölünmüştür. Ayrıca modeli test etmek için de iki veri seti(ctest10k ve Places205) kullanılmıştır. Önerilen yöntemin performansını değerlendirmek için Kök Ortalama Karesel Hatası(Root Mean Square Error ,RMSE), Ortalama Karesel Hatası (Mean-squared error, MSE), Ortalama Mutlak Hata(Mean Absolute Error ,MAE), Tepe Sinyal -Gürültü Oranı (Peak Signal to Noise Rratio ,PSNR), ve renklendirme süresi değerleri kullanılmıştır. Ctest10k veri setinden ortalama RMSE, MSE, MAE, PSNR ve renklendirme süresi değerleri sırasıyla 5.189889, 29.144883, 114.5571, 34.23678, 0.832537 olarak elde edilmiştir. Places205 veri setinden elde edilen ortalama RMSE, MSE, MAE, PSNR ve renklendirme süresi değerleri sırasıyla 7.098634, 52.15844, 120.4613, 31.28532, 0.840554 olarak bulunmuştur.

Anahtar Kelimeler: ESA, Görüntü renklendirme, Otomatik renklendirme, Lab renk uzayı, Derin öğrenme.

ABSTRACT

MS THESIS

AUTOMATIC IMAGE COLORIZATION BASED ON CONVOLUTIONAL NEURAL NETWORK

Omar Abdulwahhab Othman OTHMAN

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCE OF SELÇUK UNIVERSITY
THE DEGREE OF MASTER OF SCIENCE IN COMPUTER ENGINEERING

Advisor: Dr. Betül UZBAŞ

2020, 57 Pages

Jury
Assist.Prof.Dr. Betül UZBAŞ
Assoc.Prof.Dr. Humar Kahramanlı ÖRNEK
Assoc.Prof.Dr. Gülay TEZEL

The automatic colorization is the task of adding the color to a given gray image taken as input without any user effort. The fact that the true color values of gray scale images are unknown and the objects have many colors and shades makes coloring difficult For example, plastic objects can have many colors, or trees can have various shades of green. In this thesis, automatic colorization of gray-scale images using the Convolutional Neural Network has been proposed. The model was developed using the pictures in the ImageNet dataset and Convolutional Neural Network. The color of the gray image taken as an input is estimated using the developed model. Lab Color Space is used as color space in this study. The model takes the L channel as the input and gives the ab channel as the output. The data set containing 39,604 images randomly selected from the ImageNet data set was divided into 80% training and 20% validation to train the Convolutional Neural Network model. Also two data sets(ctest10k and Places205) are used for testing the performace of model. Root Mean Square Error (RMSE), Mean Square Error (MSE), Mean Absolute Error (MAE), Peak Signal-Noise Ratio (PSNR) and coloration time values were used to evaluate the performance of the proposed method. Average RMSE, MSE, MAE, PSNR and coloration time values were obtained from Ctest10k data set as 5.189889, 29.144883, 114.5571, 34.23678, 0.832537 respectively. Average RMSE, MSE, MAE, PSNR and coloration time values obtained from Places205 data set were found as 7.098634, 52.15844, 120.4613, 31.28532, 0.840554, respectively.

Keywords: CNN, Image colorization, automatic colorization, Lab color space, deep learning.

ÖNSÖZ

Bu çalışmanın yürütülmesi sırasında yardım ve desteklerini esirgemeyen danışmanım Dr. Öğr. Üyesı Betül UZBAŞ ve Dr. Öğr. Üyesı Sait Ali UYMAZ hocama saygı ve teşekkürlerimi sunarım. Yoğun çalışmalarım esnasında gösterdikleri sabır, anlayış ve desteklerinden dolayı aileme ve arkadaşlarıma teşekkür ederim.

Omar Abdulwahhab Othman OTHMAN
KONYA-2020

İÇİNDEKİLER

ÖZET ..... **iv**

ABSTRACT ..... **v**

ÖNSÖZ ..... **vi**

İÇİNDEKİLER ..... **vii**

SİMGELER VE KISALTMALAR ..... **ix**

1. GİRİŞ **1**

2. KAYNAK ARAŞTIRMASI **3**

3. MATERYAL VE YÖNTEM **9**

3.1. Görüntü Renklendirme Yöntemleri 9
3.1.1. Görüntü türleri 9
3.1.2. Kızılötesi Görüntü Renklendirme 10
3.1.3. Referans tabanlı renklendirme 11
3.1.4. Kullanıcı tarafından sağlanan renk ipuçları 14
3.1.5. Otomatik renk aktarımı 14
3.1.6. ESA kullanarak renklendirme 15
3.2 Makine Öğrenmesi 16
3.3. Derin Öğrenme 17
3.4. Evrişimsel Sinir Ağı (ESA) 18
3.4.1. Evrişim katmanı 19
3.4.2. Havuzlama katmanı 20
3.4.3. Tam bağlantılı katman 21
3.4.4. Dolgu ve Adım(Paddings and Stride) 23
3.4.5. Aktivasyon Fonksiyonları 24
3.4.6. Düzenleme(Regularization) 26
3.4.7. Toplu Normalleştirme (Batch Normalization) 27
3.4.8. Geri yayılım 28
3.4.9. Gradyan inişi 30
3.5. Veri Seti 31
3.6. Renklendirmenin Kalite Değerlendirmesi 32

4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA **35**

4.1. Referans Kullanarak Otomatik Renklendirme 35
4.2. Önerilen ESA Modeli Kullanarak Renklendirme 37

5. SONUÇLAR VE ÖNERİLER **50**

5.1 Sonuçlar 50
5.2 Öneriler 51

KAYNAKLAR **52**

ÖZGEÇMİŞ 57

SİMGELER VE KISALTMALAR

Kısaltmalar

CBIR : Content-Based Image Retrieval (İçerik Tabanlı Görüntü Erişimi)
ESA : Evrişimsel Sinir Ağı
HIS : Hue, Saturation and Intensity (Renk Özü, Doygunluk, Yoğunluk)
HSB : Hue Saturation and Brightness (Renk Özü, Doygunluk, Parlaklık)
HSV : Hue Saturation and Value(Renk Özü, Doygunluk, Değer)
HOG : Histogram of Oriented Gradients (Yönelimli Degradelerin Histogramı)
IR : Infrared (Kızılötesi)
LAB : luminance-red,green-coordinat- yellow, blue coordinate (L parlaklık koordinatı, a kırmızı/yeşil koordinatı ve b sarı/mavi koordinatı)
LSTM : Long Short-Term Memory (Uzun Kısa Süreli Bellek)
LMS : Long-Medium-Short wavelengths (Uzun-orta-kısa dalga boyları)
MAE : Mean Absolute Error(Ortalama Mutlak Hata)
MRI : Magnetic Resonance Imaging (Manyetik rezonans görüntüleme)
MSE : Mean-squared error(Ortalama karesel hata)
PSNR : Peak Signal to Noise Rratio(Tepe Sinyal-Gürültü Oranı)
ReLU : Rectified linear Units(Doğrultulmuş doğrusal birimler)
RGB : Red Green Blue (kırmızı yeşil mavi)
RMSE : Root Mean Square Error(Kök Ortalama KareselHata)
SGD : Stochastic Gradient Descent(Stokastik Degrade İniş)
TBA : Temel Bileşenler Analizi
YCbCr : luminance-blue minus luma-red minus luma (Y parlaklık, Cb mavi renk farklılığı, Cr kırmızı renk farklılığı)
YIQ : luminance information-orange-blue-purple-green (parlaklık bilgisi-turuncu-mavi-mor-yeşil)

1. GİRİŞ

Renkli görüntü, gri tonlamalı bir görüntüden daha fazla bilgi içerir ve çoğu senaryoda renkli görüntü gri tonlamalı görüntüden bilgi çıkarmada çok daha kullanışlıdır. Renkli görüntü, görüntünün rengi hakkında üç boyutlu bilgilerden oluşmaktadır. Renkli görüntüler canlı ve görsel olarak izleyicilere çekici gelmektedir.

Renkli görüntüler kırmızı, yeşil ve mavi ile tanımlanan üç boyutlu görüntülerdir. Gri tonlamalı görüntüler bunların sadece bir boyutuna sahip olan görüntülerdir (sadece parlaklıktan oluşmaktadır) ve bu nedenle tek boyutlu olduğu bilinmektedir. Renkli bir görüntüyü griye dönüştürme işlemi, renk hakkındaki bilgilerin kullanılmayacağı anlamına gelir. Renkli görüntüyü gri tonlamaya dönüştürmek oldukça kolaydır ancak bunun tersi o kadar kolay değildir. Renkli görüntüyü griye dönüştürme işlemini tersine çevirerek renkler geri alınabilir gibi görünebilir, ancak bu doğru değildir. Bunun nedeni, bir gri seviyeye götüren sayısız renk olabileceğidir, ancak bunun tersine döndüğümüzde hangi rengi hangi gri seviyeye karşılık geleceğine karar verilememektedir.

Renkli kameralar ortaya çıkmadan önce çekilmiş gri tonlamalı görüntülerin çoğuna görsel çekiciliğini büyük ölçüde artıran renkleri eklemek için büyük bir insan çabasına ihtiyacı vardır. Gri tonlamalı resim veya tıbbi görüntüleri renklendirmek için genellikle elle renklerin eklenerek çizilmesi yaygındır. Karikatür, sinema, gerçek zamanlı X-ışını tarayıcı görüntüleri, derin öğrenme kullanarak otomatik renklendirmek birçok durumda zaman ve maliyet tasarrufu açısından yardımcı olacaktır.

Gri seviyeli görüntülerin orijinal renkleri bilinmediği için renklendirme işlemi çözüm doğruluğu belli olmayan bir problemdir. Gri görüntünün parlaklığı ile aynı görüntünün tam renkleri arasındaki ilişki renkli görüntüdür. Renklendirme sorununa çözüm bulmak için manuel ve yarı otomatik yöntemler kullanılabilir. Manuel renklendirmede, görüntünün farklı kısımları algılanarak renklendirilir. Dijital görüntülerde, gri görüntüdeki her piksel için bir renk seçilmesi gerekir; böylece renklendirme yapılabilir fakat bu işlem zor bir işlemdir. Yarı Otomatik Renklendirme teknikleri bu işi kolaylaştırır ve gri görüntüyü bölümlere ayırır. Daha sonra bu segmentlere uygulanacak önceden tanımlanmış bazı renkleri kullanabilir. Yarı otomatik renklendirmeye örnek olarak “BlackMagic” yazılımı verilebilir. Bu uygulama gri görüntüyü bölümlere ayırır ve bu bölümlere kullanıcı tarafından önceden tanımlanmış renkli fırçalarla işlemi uygular. Bu yazılımın sonuçları tatmin edici ve gerçeğe yakın değildir. Görüntü renklendirme, elle veya referans bir görüntü gerektiren benzer

nesneler veya sahneler gerektiren yarı otomatik renklendirme yöntemleriyle yapıldığında zor bir iştir.

Bu tez çalışmasında derin öğrenme yaklaşımı kullanılarak gri tonlamalı resmi otomatik renklendirme için bir model sunulmuştur. Giriş olarak alınan gri tonlamalı görüntülerin segmentasyon ve renk aktarımı için ESA kullanılmıştır. Bu çalışmanın amacı bir referans görüntüye gerek kalmadan önerilen ESA modeli kullanılarak renklendirme işlemini otomatik olarak gerçekleştirmektir.

Bu tez çalışması beş bölümden oluşmaktadır. Birinci bölümde renklendirme, renklendirmenin faydası, kullanım alanları ve yöntemleri anlatılmaktadır. İkinci bölümde tez çalışma konusu olan gri tonlamalı görüntü renklendirme ile ilgili yapılan literatürdeki çalışmalar bulunmaktadır.

Üçüncü bölümde tezle ilgili olan uygulamada kullandığımız materyaller, derin öğrenme, ESA, ESA katmanları, aktivasyon fonksiyonları ve veri setinden bahsedilmektedir. Bu bölümde ayrıca renklendirmede kullanılan yöntemler ve renklendirmenin temel çerçevesi anlatılmaktadır.

Dördüncü bölümde iki farklı renklendirme yaklaşım verilmiştir. Bunlardan biri referans resim kullanarak otomatik renklendirme yöntemi, diğeriyse ESA modeli kullanarak renklendirme yöntemidir.

Son bölümde önerilen model kullanılarak elde edilen sonuçlar ve gelecekte yapılacak çalışmalar için öneriler ve tavsiyeler bulunulmaktadır.

2. KAYNAK ARAŞTIRMASI

Renklendirme problemi, pikselleri sadece parlaklığa sahip olan gri ölçekli bir görüntüye üç boyutlu (RGB) piksel değerlerini atamayı içermektedir. Farklı renkler, farklı canlılık veya tona karşın aynı parlaklığa sahip olabileceğinden dolayı renklendirme probleminin özü itibariyle doğru bir çözümü bulunmamaktadır. Birçok yazar bu alanda öncü çalışmalar yapmıştır.

Welsh ve ark. (2002), renk aktarma tekniğine dayanan renklendirme algoritmasını önermiştir. Bu algoritma, referans bir renkli görüntü gerektirmektedir. Parlaklık değerine dayanan bu yöntemle renkler gri ölçekli görüntüye aktarılmaktadır. Sonuçlar iyi olmasına rağmen uzamsal tutarlılık sorunu göstermektedir. Bu problemi ortadan kaldırmak için "kartelalar" kullanmıştır. Renklerin nasıl aktarılması gerektiğini belirtmek için kullanıcılardan bu renk kartelalarını hem kaynak hem de hedef görüntülerde tanımlamaları ve ilişkilendirmeleri istenmiştir.(Welsh ve ark., 2002)

Ironi ve ark. (2005), renklerin referans görüntüden gri ölçekli görüntüye aktarılması için ayrıca bir örnek temelli renklendirme algoritması kullanmaktadır. Öznitelik uzayını sınıflandırmak için danışmanlı öğrenme tekniğini kullanmışlardır. Bu düzenin temel dezavantajı, referans görüntünün otomatik olarak değil de manuel şekilde bölütlenmesini gerektirmesidir.(Ironi ve ark., 2005)

Tai ve ark. (2005), çalışmalarında genel bir renk aktarma algoritması sunmuşlardır. Hem referans görüntünün hem de gri ölçekli görüntünün olasılıksal bölütlemesi için Gauss karışım modelini kullanmışlardır. Gri ölçekli görüntünün bölütlemesi için sadece parlaklık değerini kullanmışlardır. Yaklaşımlarının temel kısıtlılığı, kaynak görüntünün kullanıcı tarafından seçilmesi gerekliliğidir (Tai ve ark., 2005).

Vieira ve ark. (2003), referans görüntünün otomatik olarak seçilmesi için renkli görüntü veritabanı kullanarak kaynak görüntünün kullanıcı tarafından seçilmesi kısıtlılığını ortadan kaldırmışlardır (Vieira ve ark., 2003)

Gupta ve ark. (2012), otomatik olarak referans görüntünün ve gri ölçekli görüntünün süper pikselleri arasındaki uyuşmaları bulmak için hızlı bir art arda öznitelik eşleştirme düzenini benimsemişlerdir (Gupta ve ark., 2012).

Levin ve ark. (2004), optimizasyona dayalı bir yarı otomatik renklendirme düzeni önermişlerdir. Başlangıçta kullanıcı gri ölçekli görüntüde renkleri karalamakta

ve sonrasında renkleri otomatik olarak tüm görüntüye yaymaktadır (Levin ve ark., 2004).

Premoˇvze ve Thompson (2002), bölgeleri kullanıcı tarafından seçilen renk paleti ile renklendirilmiş sınıflardaki pankromatik ortogonal görüntüye ayırmaktadırlar. Bu renklendirme çeşidi yüksek kaliteli renklendirme yapmasına rağmen zaman alıcıdır. Dahası, başlangıçta işaretlenmiş piksellerde en ufak bir değişiklik varsa renklendirmenin tamamen yeniden hesaplanması gerekmektedir (Premoˇvze ve Thompson, 2002).

Renklendirme için renk aktarma yöntemi, tam otomatik renklendirmenin tipik bir çeşididir. Birçok yazar, renklendirme işlemi için Lab (L parlaklık koordinatı, a kırmızı/yeşil koordinatı ve b sarı/mavi koordinatı) renk uzayı kullanımını önermişlerdir.

Ruderman ve ark. (1998), Lab renk uzayını önermektedir. L'nin parlaklık değerini ve ab kromatik değerini sağladığı yerlerde ilk olarak RGB'den LMS'ye sonrasında LMS'den Lαβ'ya dönüşüm gerçekleştirilmektedir. YCbCr (Y parlaklık, Cb mavi renk farklılığı, Cr kırmızı renk farklılığı) ve HSV/HSB (Hue (Renk Özü), Saturation (Doygunluk), Value(Değer))/ (Hue (Renk Özü), Saturation (Doygunluk), Brightness (Parlaklık)) gibi renk uzayları da renk aktarma yönteminde kullanılabilmektedir. Gri ölçekli görüntünün yalnızca parlaklık değerine sahip olmasından dolayı birçok yazar referans görüntü ve gri ölçekli görüntü arasındaki piksellerin daha iyi eşleşmesi için enerji, entropi, kontrast vb. gibi görüntünün doku özelliklerine dayanan düzenler önermiştir. Karşılaştırma yalnızca parlaklık değerine dayalıysa her zaman iyi renklendirmeyle sonuçlanmaz. Farklı piksel renkleri aynı parlaklık değerine sahip olabileceğinden dolayı bazen yapay renklendirme efekti vermektedir. Bu problem ile başa çıkmak için yazarlar farklı düzenler önermişlerdir (Ruderman ve ark., 1998).

Kuzovkin ve ark. (2015), açıklayıcı tabanlı bir görüntü renklendirme önermişlerdir. Bir dizi istatistiksel açıklayıcı ile yama eşleşme algoritmasını kullanmaktadırlar. Ayrıca yapay olguları ortadan kaldırmak için bir düzenlileştirme yöntemi vermektedirler. Hem referans görüntü hem de gri ölçekli görüntü için doku gradyanlarını, gradyanların histogramını (Histogram of Oriented Gradients, HOG) ve gri ölçeğin yoğunluğunu sıralayarak hesaplamak için 38 boyutlu açıklayıcıları kullanılmaktadır (Kuzovkin ve ark., 2015).

Ji,Chen (2008), Kok ve ark (1996), tıbbi görüntüleri renklendirmek için Pseudo renklendirme tekniğini kullanmışlardır. Renkli görüntü, örnek bir görüntü olarak renkli

görüntünün siyah beyaz görüntüye eşlenmesini sağlamak için kullanılır ve sonrasında yazarlar, Pseudo teknikleri kullanarak siyah beyaz görüntünün gri düzeylerini renklendirmek için uygun eşlemeyi kullanmışlardır(Kok ve ark., 1996; Ji ve Chen, 2008) .

Hsieh ve ark. (2009), görüntü histogramının çalışmasını ve farklı renkli görüntü seviyelerinde farklı ışık yoğunluğunu seçmek için bir çakışma tanımı ifade etmişlerdir. Önerdikleri başka bir teknikte, görüntü detayları ile ilgili ve görüntü renk kontrastının bazı durumlarını geliştirmişlerdir, ancak bu teknik çok fazla ayrıntıya sahip olmayan görüntülerde geçerlidir. Daha sonra, endüstriyel görüntülerinin renklendirilmesi için harekete geçmişlerdir, ve görüntüleri daha küçük parçalara ayırarak renklendirme sonuçları elde etmişlerdir (Hsieh ve ark., 2009).

Başka bir teknikte Rujuta (2013), renkli referans görüntüden siyah beyaz gri tonlamalı görüntüye renk aktararak gri tonlamalı görüntüleri "renklendirmek" için genel bir teknik tanıtmışlardır. Parlaklık ve doku bilgilerini kendi görüntüleri arasında eşleştirerek kaynağa ait tüm renk halini hedef görüntüye aktarmaktadırlar (Rujuta, 2013).

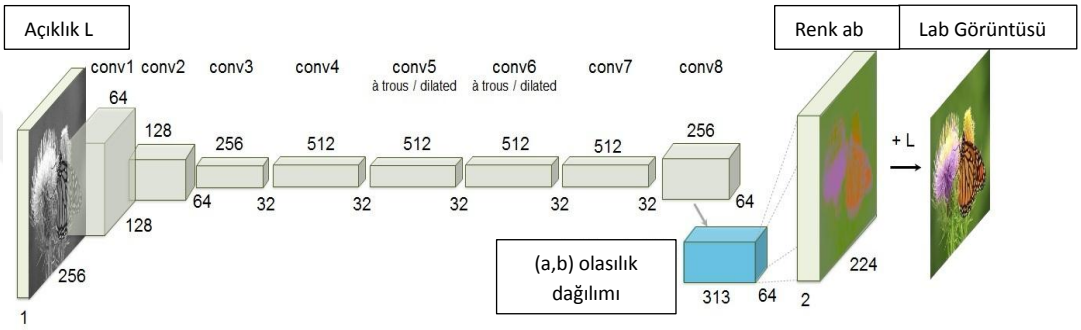
Yatziv ve Sapiro (2006), tüm renkli versiyonun geometrisini ve yapısını temsil ettiği için gradyan bilgileriyle verilen tek renkli parlaklık girdisinin geometrisini ve yapısını göz önünde bulundurarak tek renkli bir görüntüye veya filme renk eklemek için bir yaklaşım sunmuşlardır(Yatziv ve Sapiro, 2006).

Al Azzeh ve ark.(2016), görüntü segmentasyonunu, yama tabanlı örneklemeyi ve olasılıklı akıl yürütmeyi birleştiren yeni bir renklendirme örneği tekniği sunmuşlardır (Al Azzeh ve ark., 2016).

Literatürde ESA kullanılarak renklendirme çalışmaları yapılmıştır. Daha önce derin renklendirme haricinde literatürde bahsedilen tüm yöntemlerde, büyük ölçekli renklendirmede iyi bir sonuç elde etmek için kullanıcı müdahalesi ile çalışmalar yapılmıştır. Derin öğrenme haricindeki literatürdeki yöntemlerin tek bir görüntüyü renklendirmek için çalışma süreleri genellikle dakikalar halinde olduğu düşünülürse, bu durum onları çok fazla görüntüyü renklendirmek için uygunsuz hale getirmektedir. Ancak ESA'nın gelişmesiyle bazı araştırmacılar ESA'yı kullanarak SUN veya ImageNet gibi büyük veri kümeleri üzerinde çalışmışlar ve tam otomatik bir süreç elde etmişlerdir. Böylece daha yüksek bir başarı düzeyine ulaşmışlardır. Şu anda bu yöntemler doğal görüntü renklendirme için en son teknolojilerdir. ESA ağ tarafından öğrenilen ve gerçekleştirilen karmaşık bir yöntem kullanmaktadır. ESA'nın eğitim setinden otomatik

olarak seçilen referanslar ile otomatik renk veya stil aktarma biçimi olarak görülebileceği dikkat edilmesi gereken bir noktadır (Limmer ve Lensch, 2016).

Zhang ve ark. (2016), ImageNet veri kümesinin bir alt kümesinde 22 evrişimli katmanı olan düz bir ESA modeli önermişlerdir. Görevin çok modlu doğasını ele almak için her çıktı pikselinde bir renk histogramını tahmin etmek amacıyla eğitim setinden elde edilen önceki renk dağılımına dayanan sınıfın yeniden dengelenmesi ile özel olarak ayarlanmış bir multinomial çapraz entropi kaybı kullanılmıştır (Zhang ve ark., 2016). Şekil 2.1. Zhang tarafından önerilen örnek bir derin öğrenme tabanlı renklendirme modelini göstermektedir.



Şekil 2.1. Zhang ve ark. tarafından önerilen görüntü renklendirme (Zhang ve ark., 2016).

Benzer şekilde Larsson ve ark.(2016), bir renk histogramını da öngörmektedirler, ancak görüntü sınıflandırma görevi üzerinde önceden çalışılmış ve renklendirme için ince ayarı yapılmış olan piksellerin kromatik değerlerini tahmin etmek amacıyla tam bağlantılı bir hiper sütun katmanına bağlı 16 katmanlı evrişimli bir model kullanmayı tercih etmişlerdir. Yoğun bir şekilde alıştırma yapmak ve bütün görüntünün renklendirilmesini bir geçişte tahmin etmek yerine ESA, merkezi pikselin renk değerini tahmin ederek ağın alıcı alanına eşit boyutta gri ölçekli yamaların uzaysal olarak seyrek örnekleri üzerinde çalışmıştır. Larsson ve ark, renklendirmeyi geliştirmek için bilinen bir kesin referans renk histogramını (global açıklayıcı olarak) aktarma olasılığını da keşfetmişlerdir (Larsson ve ark., 2016).

Iizuka ve ark. (2016) tarafından, biri hedef görüntünün global özniteliklerini öngören diğeri ise lokal öznitelikler konusunda uzmanlaşan iki hesaplama yolunu birleştiren bir ağ önerilmiştir. Bunu başarmak için global öznitelikler renklendirme yerine görüntü sınıflandırması için çalışılmış ve daha sonra L2 Öklidyen kayıp fonksiyonu kullanılarak doğrudan renklendirme için çalışılan lokal özniteliklere

bağlanılmıştır. Bu teknik, çok tutarlı renklendirmeler üreterek modellerinin daha yüksek bir anlamsal görüntü anlayışına sahip olmasını sağlamıştır (Iizuka ve ark., 2016).

L açıklık kanalı göz önüne alındığında derin renklendirme sistemleri, Lab renk uzayında görüntünün karşılık gelen a ve b renk kanallarını tahmin etmektedir. Bu problemi çözmek için büyük ölçekli verilerden faydalanılmaktadır. Renk tahmini, eğitim verilerinin pratik olarak serbest olduğu güzel bir özelliğe sahiptir: herhangi bir renkli fotoğraf, görüntünün L kanalını girdi ve ab kanallarını denetleyici işaret olarak alarak alıştırma örneği olarak kullanılabilmektedir. Cheng ve ark, eğitim verilerinin kolay erişilebilirliğine işaret etmiş ve önceki çalışmalar büyük veri kümeleri üzerinde rengi tahmin etmek için ESA'ları çalışmışlardır. Ancak önceki girişimlerden elde edilen sonuçların solgun görünme eğilimi vardır. Başka bir yöntemle, konservatif tahminleri teşvik eden kayıp fonksiyonlarının kullanılmasıdır. Bu kayıplar, amacın bir tahmin ile kesin referans arasındaki Öklidyen hataları en aza indirmek olduğu standart regresyon problemlerinden kalmaktadır (Cheng ve ark., 2015).

Charpiat ve ark. (2008) tarafından işaret edildiği gibi renk tahmini doğal olarak çok modludur, birçok nesne çeşitli makul renklendirmeyi üzerine alabilir. Örneğin; bir elma tipik olarak kırmızı, yeşil veya sarıdır ancak mavi veya turuncu olması pek mümkün değildir. Problemin çok modlu doğasını uygun şekilde modellemek amacıyla her piksel için olası renklerin dağılımını tahmin etmek gerekmektedir. Nadir renkleri vurgulamak için eğitim süresindeki kaybı yeniden ağırlıklandırmak gerekmektedir. Bu model üzerinde eğitim yapılan geniş ölçekli verilerin tüm çeşitliliğinden yararlanması için modeli teşvik etmektedir. Son olarak, dağılımın ortalamasını alarak son bir renklendirme elde etmekte ve sonuç olarak bu yaklaşım önceki yaklaşımlardan daha canlı ve algısal olarak gerçekçi olan renklendirme elde etmektedir (Charpiat ve ark., 2008).

Gatys ve ark. (2015), bir görüntüden diğerine stil aktarmayı öğrenebilecek derin sinir ağlarının bir uygulamasını sunmuşlardır. İçeriği korunacak bir görüntü (içerik görüntüsü) ve stili aktarılacak başka bir görüntü (stil görüntüsü) göz önüne alındığında, her iki görüntüyü de önceden eğitilmiş bir ESA'ya aktarmışlardır. Ardından içerik gösterimini ve stil gösterimini çıkarmışlardır. Daha sonra bunu gürültülü bir görüntüye yapmışlar ve içerik görüntüsü ve stil görüntüsü ile benzer temsiller elde edene kadar üzerinde değişiklikler yapmışlardır. Aslında bu, amacın içerik ve stilin eşzamanlı olarak yeniden yapılandırılması kaybını en aza indirgemek olduğu bir optimizasyon problemidir. Bu sorunu çözmek için gradyan iniş kullanmışlardır (Gatys ve ark., 2016).

Hwang ve Zhou (2016) ESA'ya dayalı bir resim renklendirme yöntemi önermişlerdir. Çalışmalarında çeşitli ağ mimarilerini, hedeflerini, renk uzaylarını ve problem formülasyonlarını araştırmışlardır. Sundukları yöntemde siyah beyaz resimleri otomatik bir şekilde renklendirmeyi sağlamışlardır (Hwang ve Zhou, 2016).

Nguyen ve ark. (2016) gri tonlamalı görüntüleri renklendirmek için derin öğrenme tekniklerini kullanan yeni bir yaklaşım sunmuşlardır. Başlangıçta görüntü sınıflandırması için tasarlanmış önceden eğitilmiş ESA kullanarak, farklı görüntülerin içeriğini ve stilini ayırabilmiş ve bunları tek bir görüntüde yeniden birleştirebilmişlerdir. Ardından, içeriğini gri tonlamalı semantik benzerliğe sahip renkli bir görüntü stiliyle birleştirerek gri tonlamalı bir görüntüye renk ekleyebilecek bir yöntem önermişlerdir. Bir uygulama olarak, ukiyo-e'nin (bir Japon resmi türü) görüntülerini renklendirmek ve bu yöntemin bilgisayar destekli tasarımın büyüyen alanındaki potansiyelini gösteren ilginç sonuçlar elde etmek için önerilen yöntemi kullanmışlardır (Nguyen ve ark., 2016).

Varga ve Szirányi (2017) girişi gri tonlamalı olan bir görüntünün gerçekçi renklendirmesini sağlayabilen tam otomatik bir yaklaşım önermişlerdir. Görüntü işleme alanındaki derin öğrenme tekniklerinin son zamanlardaki başarısından hareketle, U ve V renk kanallarını öngören ESA'ya dayanan ileriye yönelik iki aşamalı bir mimari önermişlerdir. Önceki çalışmaların çoğundan farklı olarak, bu makalede karmaşık sahnelerde bile yüksek kaliteli ve gerçekçi renklendirme üretebilen tam otomatik bir renklendirme sunmuşlardır. SUN veritabanı görüntüleri ve diğer görüntüler üzerinde kapsamlı deneyler, nitel ve nicel değerlendirmeler yapılmıştır. Kuaterniyon Yapısal Benzerliği'nin bir dereceye kadar nicel değerlendirme için iyi bir temel oluşturduğunu bulmuşlar, bu yüzden renklendirme kalitesi için Kuaterniyon Yapısal Benzerliği'ni bir indeks numarası olarak seçmişlerdir (Varga ve Szirányi, 2017).

Kalyan ve Ramalingam (2019) sıfırdan eğitilmiş derin bir ESA'yı önceden eğitilmiş InceptionResNet-v2 modelinden çıkarılan üst düzey özelliklerle birleştirmişler. Modellerinin amacı giriş görüntüsünün L ışıklı bileşenine göre a* ve b * bileşenlerini tahmin etmektir. Yeniden yapılandırılan görüntünün a * ve b * bileşenleri, tahmini renkli görüntüyü vermek için parlak bileşen ile birleştirilmektedir ve bu metodoloji kullanarak gökyüzü, okyanus, hayvanlar, ağaçlar ve insan derisi gibi yüksek seviyeli resim parçalarını etkili bir şekilde renklendirmektedir (Kalyan ve Ramalingam, 2019).

3. MATERYAL VE YÖNTEM

Görüntü renklendirme, gri tonlamalı görüntüdeki her piksel için bir renk belirleme işlemi olarak adlandırılmaktadır. Genel olarak, gri tonlamalı bir görüntüyü renklendirme işlemi, insan zihni için basit bir görevdir. Erken yaşlardan itibaren boyama kitaplarında eksik renkleri doldurmayı öğreniriz. Çimlerin yeşil olduğunu, gökyüzünün beyaz, bulutların mavi olduğunu veya bir elmanın kırmızı yada yeşil olabileceğini öğreniriz.

Bir görüntüyü renklendirmek için görüntünün doğası gereği içeriğinin derinliği, görüntünün neler içerebileceği ve ayrıca ortak nesnelerin renk bilgisi hakkında detaylı bilgi gerekmektedir. Nesnelerin doğası göz önünde bulundurularak, renklendirme modelinin çeşitli nesneler üzerinde eğitilmesi gerekmektedir, örnek verilecek olursa; renk mavi ise nesne su olmaktadır ya da renk yeşil ise nesne bahçenin rengi olabilmektedir. Benzer şekilde, renklendirme modelinin de renklendirme işlemini gerçekleştirmek üzere eğitilmesi gerekmektedir. Bu konuda yapılan araştırmalarda, elde edilen görüntüler üzerinde uğraşılarak renklendirmeyi titizlikle gerçekleştirmeye çalışmaktadır.

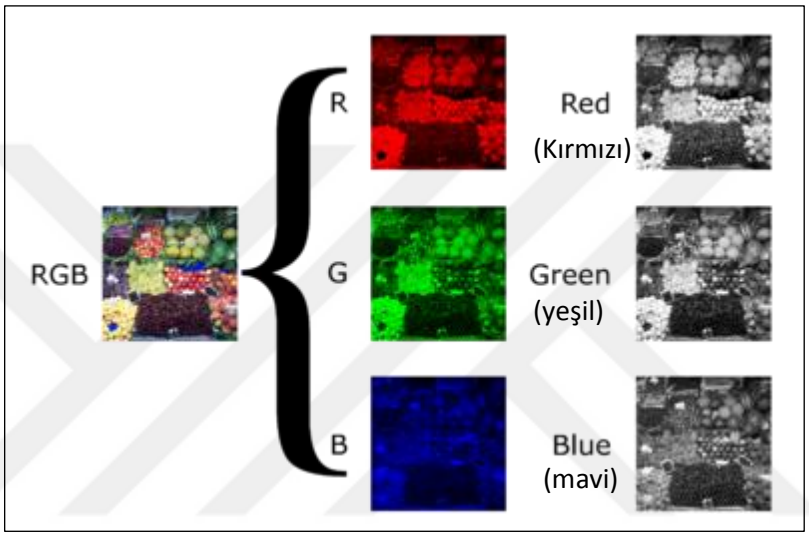
Bu bölümde, görüntü renklendirme yöntemleri ve ESA'nın genel yapısı anlatılmıştır. Tez çalışmasında kullanılan veri seti ve renklendirme kalitesini değerlendirme ölçütleri de bu bölümde sunulmuştur.

3.1. Görüntü Renklendirme Yöntemleri

3.1.1. Görüntü türleri

Dijital görüntü, doğal nesneleri tanımanın faydası nedeniyle bilim ve mühendislik alanlarında kullanılmaktadır. Bununla birlikte, dijital kelimesiyle, her noktaya bir değer atayan görüntüler ifade edilir. Sonuç olarak görüntüler, farklı yoğunlukta (değerde) çok sayıda nokta içerir. Dijital sistemlerin geliştirilmesiyle birlikte, yeni kameralar eski moda gri tonlamalı görüntüler yerine renkli görüntüleri almak için tasarlanmaktadır. Dijital görüntüler artık çok kanallı net renkli görüntüler sağlamaktadır.

Siyah beyaz görüntüler 0 ve 1 değerlerinden oluşan görüntülerdir, ikili görüntü olarak da adlandırılırlar. Gri tonlamalı görüntüler, değişen gri tonlarını 0-255 aralığında pikseller ile temsil etmektedirler. Bu görüntü türünde siyah 0 piksel değeriyle beyaz ise 255 değeriyle temsil edilir. İlgi çekici görüntüler (renkli görüntüler) renk uzayları ile temsil edilir. RGB yaygın olarak kullanılan renk uzaylarındandır. Kameralara bağlı olarak kırmızı, mavi, yeşil şeklinde 3 kanaldan oluşur. Şekil 3.1'de renkli görüntü olarak nitelendirilen birleşik renkleri göstermektedir.



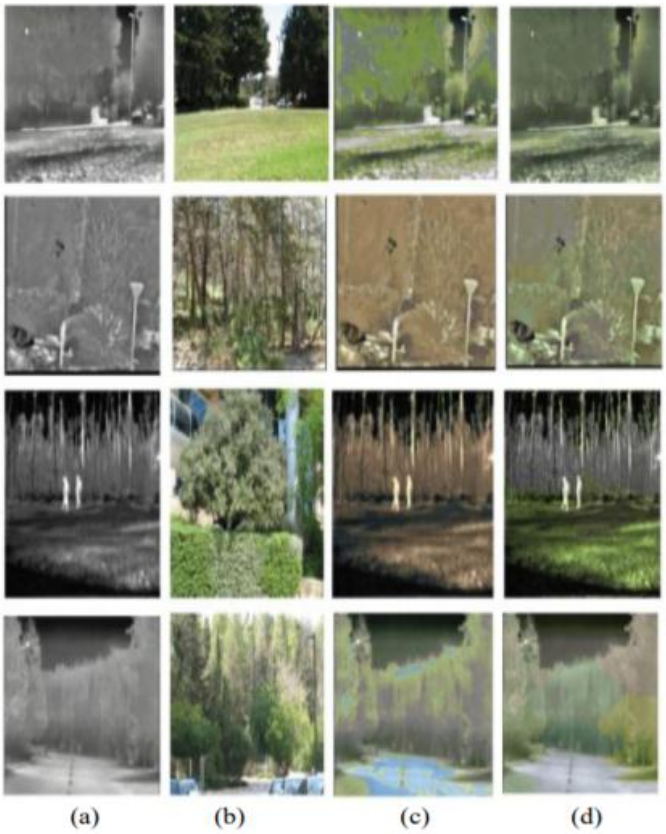
Şekil 3.1. Resmin renkli ve RGB kanalları (Site, 2020)

.

3.1.2. Kızılötesi Görüntü Renklendirme

Kızılötesi görüntüler uydu görüntülerinden jeolojik nesneleri tanımak için yardımcı olan diğer bilgileri verebilmektedir. Uydu görüntü işleme uzaktan algılama gibi özel uygulamalar ile kullanılabilmektedir. Kızılötesi görüntüler izleme, gözetleme cihazları vb. gibi çok çeşitli uygulamalara sahip tek renkli görüntülerdir. Bu görüntüler renkli görüntülere dönüştürülebilirlerse görsel çekiciliklerini ve önemlerini kesinlikle artıracaklardır. Hamam ve ark. (2012) tarafından görüntüdeki pikselin parlaklık değerine ve komşu koordinatlandırmaya (x, y) dayanan bölütleme tekniklerinin kullanımı önerilmiştir. Özniteliklerin çıkarılması sonrasında yazarlar renkli görüntülerin eşleşen segmentini kızılötesi görüntülere tayin etmişlerdir. Bu segmentlerin eşleştirilmesi, HOG yöntemi kullanılarak gerçekleştirilebilir. Segment eşleştirmesi yapılınca, kromatikliğe ve yerel piksel istatistiklerine göre bunları kümeleyerek her

segmentin pikselleri için kromatik kod çizelgesi oluşturulması önerilmektedir. Son olarak, renklendirme işlemini tamamlamak için IR (Kızılötesi) görüntü pikselleri en benzer kod çizelgesi dokusal değerleriyle eşleştirilmektedir. Sonrasında eşleştirilen kromatik değerler α ve β aktarılmakta ve orijinal parlaklık değerleri korunmaktadır (Hamam ve ark., 2012). Şekil 3.2'de tek bant kızılötesi doku tabanlı görüntü renklendirme sonuçları sunulmuştur.

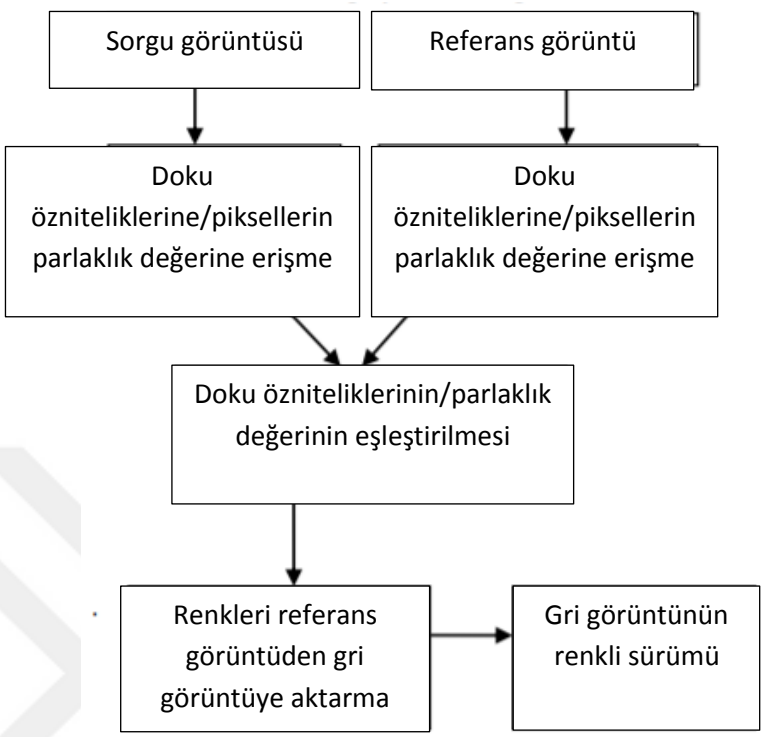


Şekil 3.2. Tek bant kızılötesi doku tabanlı görüntü renklendirme sonuçları (a) hedef IR görüntüleri, (b) renk kaynağı görüntüsü, (c) görüntü renklendirmesi (Toet, 2005), (d) önerilen teknikle renklendirme (Hamam ve ark., 2012)

3.1.3. Referans tabanlı renklendirme

Referans tabanlı renklendirmenin temel çerçevesi Şekil 3.3'te görünmektedir. Önerilen renk aktarma düzenlerinin çoğu için bu işlem neredeyse aynıdır. İlk olarak referans görüntü ve gri ölçekli görüntü öznitelikleri her iki görüntüden çıkarılır ve sonrasında piksellerin hangi özniteliklerinin çıkarıldığına bağlı olarak ya parlaklık ya da doku eşleştirme algoritmaları kullanılarak eşleştirilir. Sonrasında renk, referans

görüntüden gri ölçekli görüntüye aktarılır. Son olarak gri ölçekli görüntünün renkli sürümüne erişilir.



Şekil 3.3. Referans tabanlı renklendirmenin temel çerçevesi

Tüm renklendirme işlemini otomatik hale getirmek için kullanıcı tarafından referans görüntü verilmesi yerine referans görüntüyü insan müdahalesi olmadan sistemin seçmesi gerekmektedir. Büyük bir veri tabanından resim seçmek meşakkatli bir iştir fakat literatürdeki çalışmalarda içerik tabanlı görüntü erişiminin uygulanması önerilmiştir. Bir görüntüye veri tabanından erişildiğinde renklendirme işlemi onun üzerine uygulanabilmektedir. Literatürde görüntüye erişmek için histogram eşleştirme, Temel Bileşen Analizi (TBA) gibi farklı İçerik Tabanlı Görüntü Erişimi (Content-Based Image Retrieval, CBIR) tekniklerini kullanılmıştır. Şekil 3.4 referans tabanlı görüntü renklendirme örneğini gösterilmektedir (Aras ve Gangal, 2009).



Şekil 3.4. Referans tabanlı görüntü renklendirme örneği

Çizelge 3.1'de var olan referans tabanlı görüntü renklendirme yöntemlerini göstermektedir.

Çizelge 3.1. Mevcut referans tabanlı görüntü renklendirme yöntemleri

Kullanılan metodoloji	Değerlendirme Metriği	Kullanılan Renk Alanı
CBIR(Content-Based Image Retrieval) tabanlı görüntü renklendirme (Hadhoud ve ark., 2007)	RMSE, geçen süre	Lab
Görüntülerin doku özelliğinin çıkarılması(Patel ve Parmar, 2013)	MSE, PSNR, MD, Alınan Zaman	YCbCr
Vector quantization algoritması kullanılarak renk kod kitabının üretilmesi(Kumar ve Swarnkar, 2012)	MSE, PSNR	GRGB (Gri, Kırmızı, Yeşil ve Mavi)
Segmentasyon ve kümeleme (Hadhoud ve ark., 2007)	Görüntülerin kalite karşılaştırması	HSV / HSB

3.1.4. Kullanıcı tarafından sağlanan renk ipuçları

Renk karalamalarına dayanan renklendirme yöntemleri, rengi renk yamalarından tüm görüntüye yaymak için genellikle açık parametre öğrenimi olmayan bir optimizasyon çerçevesi kullanmaktadır. Karalamalar genellikle bir renk şeffaflığı maskesi şeklinde ayrı bir görüntü olarak sağlanır ve bu maskede tanımlanmış net bir rengi olmayan görüntünün bölümleri kendilerine yayılmış renk bilgilerine sahip olmalıdır. Bu yöntemlerin çoğunun arkasındaki temel varsayım, benzer yoğunluklardaki yakın piksellerin benzer renklere sahip olması gerektiğidir.

Levin ve arkadaşlarının önerdiği yöntemde renklendirme, komşu pikseller arasındaki yoğunluk farklılıkları olarak elde edilen dışbükey ikinci dereceden bir maliyet fonksiyonunun çözülmesiyle elde edilir (Levin ve ark., 2004). Huang ve arkadaşlarının, nesne sınırları üzerindeki renk akmasıyla ilgili yaygın sorunları azaltmak amacıyla kenar saptamasını kullanmak için daha fazla iyileştirme yöntemi, bu doğal görüntüleri etkileşimli olarak renklendirmek için nispeten popüler bir teknik haline gelmiştir (Huang ve ark., 2005). Luan ve arkadaşları, kabartmaların kullanımını doku benzerliğine genişleten, kabaca benzer renkleri paylaşması gereken pikselleri otomatik olarak etiketleyen ve bunları tutarlı bölgelere gruplandıran bir yöntem sunmuşlardır. Tekniği daha da geliştirerek renklendirmeyi etkili bir şekilde ilerletmek için renk yerellik varsayımını genişleterek renge benzer dokulara sahip uzak pikseller ararlar (Luan ve ark., 2007). Futschik tarafından sunulan renklerin karalamalardan aktarılmasına benzer bir yaklaşım, doku sürekliliğini ölçmek için yerel komşuların istatistiksel desen özelliklerini çıkarmakta daha az karalama gerektirmektedir (Futschik, 2018).

3.1.5. Otomatik renk aktarımı

Karalama tabanlı yöntemlere çok benzer şekilde, görüntüden görüntüye renk aktarımı gerçekleştiren algoritmalar, kullanıcının ek girdiler sağlamasını bekler. Daha basit yöntemler, sadece tek bir görüntüden renklendirmeyi hedef görüntüye aktarır. Ancak istatistiksel özelliklere dayalı renk çıkarma için referans görevi gören bir dizi görüntüyü tanımlamak daha yaygındır. Bazı algoritmalar, hedef görüntüyü daha önce tartışılan renk geliştirme algoritmaları ile işlemeyi, değişen aydınlatma gibi efektleri

kaldırmayı veya sonucun istenen ve bilinen bir renk histogramı gibi global özelliklerini zorlamayı seçer (Deshpande ve ark., 2015).

Çoğu algoritma, SURF, Gabor, yama veya Papatya tanımlayıcıları gibi çeşitli görüntü özelliklerini kümeden çıkarır ve bu özelliklerin renk kanalı verilerine eşlenmesini öğrenir. Bu tanımlayıcılar daha sonra hedef görüntüden çıkarılır ve eşlenmiş renk dağılımları, Welsh ve arkadaşlarının çalışması gibi elde edilen tanımlayıcıların temsil ettiği bölgelere aktarılır ve seçilen özelliklerin 5 × 5 yerel komşunun parlaklık değeri ve istatistiksel özellikleri olduğu bir yöntemi önermektedirler. Hedef görüntüdeki her piksel, kaynak görüntüden çıkarılan bu özelliklerin bir kümesiyle eşleştirilir. Set, titreşimli örnekleme veya manuel olarak tanımlanan dikdörtgen numuneler kullanılarak üretilir. En iyi eşleşen özellikler bulunduktan sonra, renk bilgileri hedef piksele aktarılır. Parlaklık kanalı, çoğu renklendirme yönteminde olduğu gibi değişmeden kalır (Welsh ve ark., 2002).

3.1.6. ESA kullanarak renklendirme

Daha önce tarif edilen yöntemler, derin renklendirme hariç, büyük ölçekli renklendirme için teorik verimleri azaltan, kullanımlarını elverişsiz hale getirebilen ve deneme yanılma yöntemi ile sonuçlanan bir tür kullanıcı yardımı gerektirir. Çalışma sürelerinin genellikle birkaç dakika olduğu göz önüne alındığında, bu onları çok fazla görüntüyü renklendirmek için uygun hale getirmez. Bununla birlikte, bilgisayar görüşündeki ESA devrimi ile bazı araştırmacılar, ESA'ları SUN (Xiao ve ark., 2010) veya ImageNet (Deng ve ark., 2009) gibi büyük veri kümelerinde eğiterek bunu tam otomatik bir süreç olarak daha yüksek bir başarı ile ele almaya başladılar. Şu anda, bu yöntemler doğal görüntü renklendirme için en son teknolojidir. ESA'ların bu göreve uygulanmasının, ağ tarafından orijinal ESA eğitim seti seçimi ile otomatik olarak önceden seçilen referanslarla öğrenilen ve gerçekleştirilen karmaşık bir yöntem kullanılarak otomatik renk veya stil aktarımı biçimi olarak görülebildiğini belirtmek gerekir. Bununla birlikte, çok çeşitli semantik olarak farklı sahneler ve yaygın olarak oluşan nesneler (ImageNet gibi) içeren eğitim setini kullanarak, seçilen renk transferinin en uygun eşleşme olması beklenir.

3.2 Makine Öğrenmesi

Makine öğrenmesi, tanımı gereği, yapay zekadaki örüntü tanıma ve hesaplamalı öğrenme teorisini inceleyerek gelişen bir bilgisayar bilimi alanıdır. Ayrıca veri kümelerinden öğrenebilen ve bunlar üzerinde tahminler yapabilen algoritmaların öğrenilmesi ve oluşturulmasıdır. Bu prosedürler, sağlam statik program talimatlarını takip etmek yerine veriye dayalı tahminler veya seçimler yapmak için örnek girişlerden bir modelin oluşturulmasıyla çalışır. "Bir bilgisayar programı bazı görev T ve bazı performans ölçüsü P ile ilgili deneyimi E ile geliştirmektedir, T üzerindeki performansı ise, P ile ölçülür, deneyimi E ile geliştirir." Dolayısıyla, programımızın yoğun bir düğümdeki trafik görevlerini öngörmesini istiyorsak (görev T), makine öğrenme süresinde önceki alışkanlıklarıyla ilgili veriler basılabilir (deneyim E), eğer başarılı bir şekilde "öğrenmiş" ise o zaman daha iyi (performans ölçütü P) alışkanlıkları tahmin edilebilmektedir.

Makine öğrenmesi birçok alanda kullanılmaktadır. İnsan uzmanlığının olmadığı yerde örneğin Mars'ta gezinme, insanların uzmanlıklarını açıklayamadığı görevlerde örneğin konuşma tanıma, çözümün zamanla değiştiği durumlarda örneğin sıcaklık kontrolü, çözümün belirli durumlara uyarlanması gerekiyorsa örneğin biyometrik gibi alanlarda makine öğrenmesi kullanılmaktadır. Örnek olarak bir akustik konuşma sinyalinin ASCII metnine dönüştürüldüğü sözlü konuşmanın tanınmasını düşünülebilmektedir. Bir kelimenin telaffuzunda, insanların yaşı, cinsiyeti veya telaffuzdaki farklılıklar kişiden kişiye değişebilir. Bu nedenle makine öğrenimindeki yaklaşım, farklı insanlardan geniş örnek ifadeler toplamak ve bunları kelimelere çizmeyi öğrenmektir. Başka bir örnek olarak, paketleri bir bilgisayar ızgarası üzerinden yönlendirmek düşünülebilir. Kaynaktan hedefe hizmet kalitesini en üst düzeye çıkaran iz, sistem trafiği değiştikçe düzenli olarak değişmektedir. Bir öğrenme yönlendirme prosedürü, ağ trafiğini izleyerek en iyi yola uyum sağlayabilmektedir (Mohammed ve ark., 2016).

Makine öğrenimi iki tür görevi içermektedir: ilk olarak danışmanlı makine öğrenimi: burada program, önceden tanımlanmış bir "eğitim örnekleri" kümesi üzerinde eğitilmektedir ve bu da yeni veriler verildiğinde doğru bir sonuca ulaşma yeteneğini kolaylaştırmaktadır. Diğeri ise danışmansız makine öğrenimidir: Programa bir grup veri verilir ve burada kalıplar ve ilişkiler bulunmalıdır. Tahmin aşağıdaki denklemde makine öğrenme algoritmaları kullanılarak yapılabilir.

$$h(x) = \partial 0 + \partial 1.X \quad (3.1)$$

Bu denklemdeki h eğitim çıktısıdır ve istenen gerçek çıkışın y olmasına izin vermektedir, $\partial 0$ ve $\partial 1$ eğitim parametreleridir, X ise girdi verileridir. Makine öğrenimi süresi, h ve y arasındaki en iyi korelasyona ulaşana kadar $\partial 0$ ve $\partial 1$ eğitim parametrelerini ayarlamaktır (Mitchell, 2006).

3.3. Derin Öğrenme

Derin öğrenme (derin yapılandırılmış öğrenme veya hiyerarşik öğrenme olarak da bilinir), göreve özgü algoritmaların aksine, öğrenme verilerini temsil etmeye dayanan daha geniş bir makine öğrenme yöntemleri ailesinin bir parçasıdır.

Derin öğrenme, çoklu işlem katmanlarının hesaplama modellerinin, beynin çok yönlü bilgileri nasıl algıladığını ve anladığını taklit eder, böylece büyük ölçekli verilerin karmaşık yapılarını örtük olarak yakalayan çoklu soyutlama seviyeleriyle verileri öğrenmesini sağlamaktadır. Derin öğrenme, sinir ağlarını, hiyerarşik olasılık modellerini, çeşitli danışmansız ve danışmanlı özellik öğrenme algoritmalarını kapsayan zengin bir yöntem ailesidir. Son zamanlarda derin öğrenme yöntemlerine olan ilgi artmıştır buda çeşitli görevlerdeki en son teknoloji tekniklerinden daha iyi performans gösterdikleri ve farklı kaynaklardan gelen karmaşık verilerin bolluğundan (örneğin, görsel, işitsel, tıbbi, sosyal ve sensör) kaynaklanmaktadır.

İnsan beynini simüle eden bir sistem oluşturma tutkusu, sinir ağlarının ilk gelişimini körüklemiştir. Nöron adı verilen birbirine bağlı temel hücreleri kullanarak beynin nasıl son derece karmaşık desenler üretebileceğini anlamaya dayanmaktadır. MCP modeli olarak adlandırılan nöronun McCulloch ve Pitts modeli, yapay sinir ağlarının geliştirilmesine önemli katkıda bulunmuştur. Literatürde LeNet (Wang ve ark., 2019), ve Uzun Kısa Süreli Bellek (Baynes ve ark., 2018), dahil olmak üzere günümüzün "derin öğrenme çağına" yol açan bir dizi önemli katkı sunmaktadır. Derin öğrenmede en önemli atılımlardan birini 2006 yılında Hinton ve ark. (Bouhal ve ark., 2018), birden fazla Kısıtlı Boltzmann Makineleri katmanıyla Derin İnanç Ağı'nı, danışmansız bir şekilde bir katman eğitimi ile tanıtmıştır. Her düzeyde yerel olarak gerçekleştirilen danışmansız öğrenmeyi kullanarak ara temsil düzeylerinin eğitimine

rehberlik etmek, son on yılın derin mimarilerde ve derin öğrenme algoritmalarındaki artışa neden olan bir dizi gelişmenin arkasındaki temel ilkedir.

Derin öğrenmenin büyük artışa katkıda bulunan en önemli faktörler arasında, CPU tabanlıdan GPU tabanlıya geçişi sağlayan paralel GPU hesaplamanın güçlendirilmesinin yanı sıra büyük, yüksek kaliteli, halka açık etiketli veri kümelerinin ortaya çıkması da bulunmaktadır. Böylece derin modellerin eğitiminde önemli bir gelişme sağlanmaktadır. Aktivasyon fonksiyonlarında hiperbolik tanjant ve lojistik fonksiyon gibi oluşan problemlerin azaltılması için yeni düzenleme tekniklerinin önerilmesi daha hızlı prototip oluşturmaya izin vermektedir.

Derin öğrenme sayesinde nesne algılama, hareket izleme, eylem tanıma, insan poz tahmini ve semantik bölümleme gibi çeşitli bilgisayar görme problemlerinde büyük adımlar atılmıştır. Tekrarlayan Sinir Ağları kategorisinde Uzun Kısa Süreli Bellek (LSTM), derin öğrenme şeması olarak büyük önem taşımasına rağmen, ağırlıklı olarak dil modelleme, metin gibi sorunlara uygulandığından bu derlemede sunulmamaktadır. Genel olarak bilgisayar görme ve multimedya analiz araştırmacılarının yanı sıra, nesne algılama ve tanıma, yüz tanıma gibi bilgisayar görme görevleri için derin öğrenme konusunda en son teknoloji ile ilgilenen genel makine öğrenimi araştırmacıları için de yararlı olmayı amaçlamaktadır (Niu ve ark., 2017).

3.4. Evrişimsel Sinir Ağı (ESA)

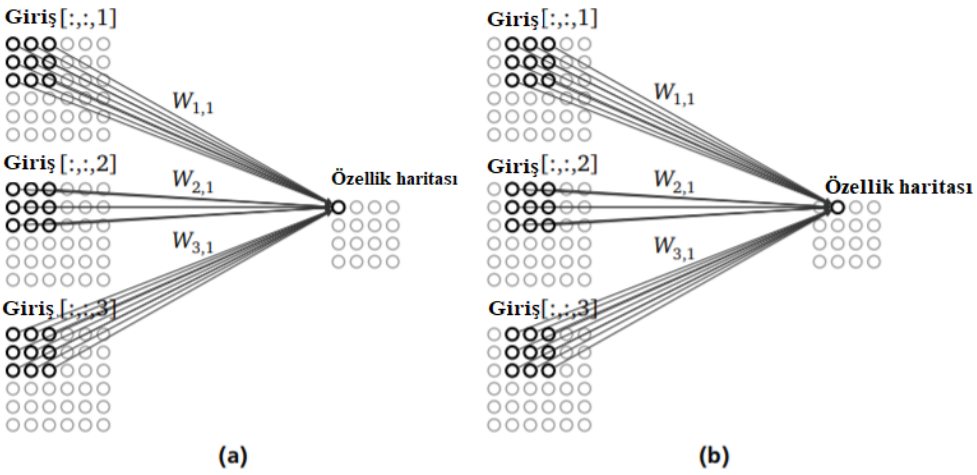
ESA'lar görsel sistemin yapısından ve özellikle de Qin ve ark. (2018)'nın yaptığı çalışmada önerilen modellerden esinlenmiştir (Qin ve ark., 2018), Nöronlar arasındaki bu yerel bağlantılara ve görüntünün hiyerarşik olarak organize edilmiş dönüşümlerine dayanan ilk hesaplama modelleri Neocognitron'da bulunmaktadır (Raseman ve ark., 2017). YannLeCun ve ark. hata gradyanını kullanarak ve çeşitli örüntü tanıma görevlerinde çok iyi sonuçlar elde eden ESA tasarlamışlardır (LeCun ve ark., 1998). Bir ESA, üç ana sinir katmanı türünü içerir, (i) evrişimli katmanlar, (ii) havuzlama katmanları ve (iii) tam bağlantılı katmanlar. Her katman türü farklı bir rol oynamaktadır. Bir ESA'nin her katmanı, giriş hacmini nöron aktivasyonunun bir çıkış hacmine dönüştürür. Sonuçta nihai tam bağlanmış katmanlara yol açar ve bu da giriş verilerinin bir boyutlu özellik vektörüne eşlenmesine yol açmaktadır. ESA'lar yüz tanıma, nesne algılama, robotikte görme gücü ve kendi kendini süren otomobiller gibi bilgisayar görme uygulamalarında son derece başarılı olmuştur (Wu, 2017).

3.4.1. Evrişim katmanı

Bir evrişim katmanı normal bir sinir ağının yanı sıra farklı birimlerden de oluşmaktadır. Diğer birimler farklı düzenleme ve bağlantı mekanizmalarından oluşmaktadır. Sinir ağları arasındaki temel farklar şunlardır:

- Birimlerin yerine bir boyutlu ve üç boyutlu düzenlemeler
- ağırlık paylaşımı
- yerel bağlantı

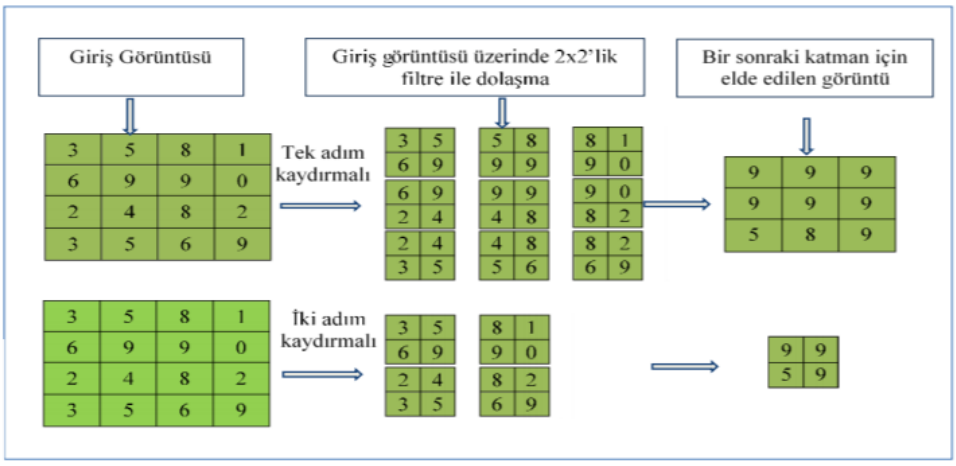
Üç boyutlu düzenleme görüntüden oluşmaktadır. Renkli bir görüntü genellikle üç kanala (kırmızı, yeşil, mavi) sahiptir ve her kanal iki boyutlu bir matris ile tanımlanmaktadır. Bu nedenle, ESA'nın girişi üç boyutlu bir matristen oluşmaktadır. Bir evrişim tabakasının çıktısı, yine iki boyutlu özellikte olan bir matristir. Her filtre bir özellik haritası oluşturmaktadır. Ağırlık paylaşımı, çeşitli çıktı birimleri için aynı ağırlıkları kullanmaktır. Bu nedenle ESA, çeviriye karşı değişmez özellik kazanmaktadır. Bunu Sobel filtresiyle karşılaştırmak için ağırlıklar filtre olur ve bu filtre girdideki çıktıyı üretmek için kullanılır. Şekil 3.5'te gösterilmiş olan iki şekil, evrişim katmanının giriş ve iki çıkış birimi arasındaki bağlantıyı göstermektedir. Ayrıca evrişim katmanı çıktı için sadece bir özellik haritasına sahiptir. Aktivasyon fonksiyonu ve bias daha iyi bir genel bakış için ihmal edilebilir. Yerel bağlantı iki şekilde gösterilmiştir. 3x3 filtre boyutu ile (a) ve (b) özellik haritasının hesaplanması için yeni ağırlıklar kullanılabilir (Aghdam ve Heravi, 2017).



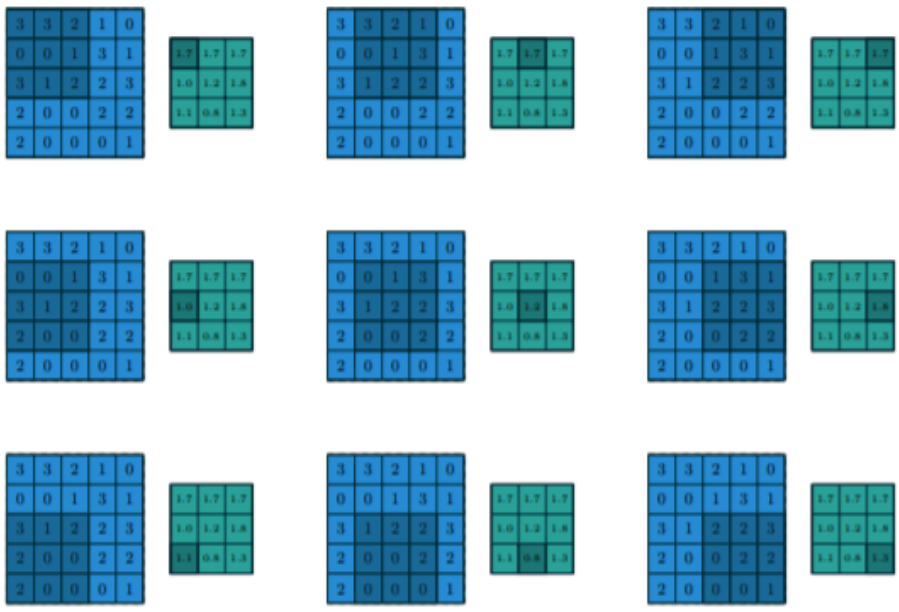
Şekil 3.5. İki şekil, evrişim katmanının giriş ve iki çıkış birimi arasındaki bağlantıyı göstermektedir (Aghdam ve Heravi, 2017).

3.4.2. Havuzlama katmanı

Havuzlama katmanları, bir sonraki evrişimsel katmanın girdi hacminin uzamsal boyutlarını (genişlik × yükseklik) azaltmaktan sorumludur. Havuzlama katmanı, birimin derinlik boyutunu etkilemez. Bu katman tarafından gerçekleştirilen işleme, boyut küçültme aynı anda bilgi kaybına yol açtığından alt örnekleme veya aşağı örnekleme de denmektedir. Bununla birlikte, böyle bir kayıp ağ için yararlıdır, çünkü boyuttaki azalma ağın yaklaşan katmanları için daha az hesaplama yüküne yol açar ve aynı zamanda aşırı uyum göstermeye karşı da çalışmaktadır. Ortalama havuzlama ve maksimum havuzlama en sık kullanılan stratejilerdir. Maksimum havuzlama ve ortalama havuzlama performanslarının ayrıntılı bir teorik analizinde verilirken, maksimum havuzlamanın daha hızlı yakınsamaya yol açabileceği, üstün değişmez özellikleri seçebileceği ve genelleştirmeyi geliştirebileceği gösterilmiştir. Ayrıca literatürde havuzlama katmanının her biri farklı motivasyonlardan esinlenen ve stokastik havuzlama, mekansal piramit havuzlama ve def-havuzlama gibi farklı ihtiyaçlara hizmet eden bir dizi başka varyasyon vardır (He ve ark., 2015). maksimum havuzlama ve ortalama havuzlama prensibleri şekil 3.6'da ve Şekil 3.7'de gösterilmiştir.



Şekil 3.6. 1 × 1 ve 2x2 basamak kullanarak 5 × 5 girişinde 2 × 2 maksimum havuzlama işleminin çıkış değerlerinin hesaplanması (Özkan ve Ülker, 2017).



Şekil 3.7. 1 × 1 basamak kullanarak 5 × 5 girişinde ortalama 3 x 3 havuzlama işleminin çıkış değerlerinin hesaplanması (Dumoulin ve Visin, 2016).

3.4.3. Tam bağlantılı katman

Birkaç evrişimsel ve havuzlama katmanını takiben, sinir ağındaki yüksek düzey mantığı tamamen bağlı katmanlar aracılığıyla gerçekleştirmektedir. Tamamen bağlı bir katmandaki nöronlar, adından da anlaşılacağı gibi, önceki katmandaki tüm aktivasyonlarla tam bağlantılara sahiptir. Dolayısıyla bunların aktivasyonu bir matris çarpımı ve ardından bir sapma ile hesaplanabilir. Tam bağlantılı katmanlar sonunda iki boyutlu özellik haritalarını bir boyutlu özellik vektörüne dönüştürür. Elde edilen vektör ya sınıflandırma için belirli sayıda kategoriye ilerletilebilir ya da daha sonraki işlemler için bir özellik vektörü olarak düşünülebilir.

ESA'ların mimarisi üç somut fikir kullanır: (a) yerel alıcı alanlar, (b) bağlı ağırlıklar ve (c) mekansal alt örnekleme. Yerel alıcı alana dayanarak, evrişimsel bir katmandaki her birim, bir önceki tabakaya ait bir dizi komşu birimden girdi alır. Bu şekilde nöronlar, kenarlar veya köşeler gibi temel görsel özellikleri çıkarabilirler. Bu özellikler daha sonra daha yüksek mertebe özellikleri tespit etmek için sonraki evrişimsel katmanlar tarafından birleştirilir. Ayrıca, bir görüntünün bir kısmında yararlı olan temel özellik dedektörlerinin, tüm görüntü boyunca yararlı olmaları fikri, bağlı ağırlıklar kavramı ile uygulanır. Bağlı ağırlıklar kavramı, bir dizi birimi aynı ağırlıklara sahip olacak şekilde kısıtlar. Somut olarak, evrişimsel bir tabakanın birimleri

düzlemlerde düzenlenir. Bir düzlem tüm birimleri aynı ağırlık kümesini paylaşır. Bu nedenle, her düzlem belirli bir özellik oluşturmaktan sorumludur. Düzlemlerin çıktılarına özellik haritaları denir. Her evrişimsel katman birkaç düzlemden oluşur, böylece her konumda birden fazla özellik haritası oluşturulabilmektedir. Bir özellik haritasının oluşturulması sırasında görüntünün tamamı, özellikleri özellik haritasındaki ilgili konumlarda saklanan bir birim tarafından taranmaktadır. Bu konstrüksiyon bir kıvrım işlemine eşdeğerdir, ardından bir ek yanlılık terimi ve sigmoid işlevi gelmektedir (Voulodimos ve ark., 2018).

$$y^{(d)} = \sigma\left(W y^{(d-1)} + b\right) \tag{3.2}$$

Burada d evrişimsel katmanın derinliğini temsil etmektedir, W ağırlık matrisidir ve b bias terimidir. Tam bağlantılı sinir ağları için, ağırlık matrisi doludur, yani her girişi farklı ağırlıklara sahip her birime bağlar. ESA'lar için, bağlı ağırlık kavramı nedeniyle W ağırlık matrisi çok seyrektir.

$$\begin{bmatrix} w & 0 & \dots & 0 \\ 0 & w & \dots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & \dots & 0 & w \end{bmatrix} \tag{3.3}$$

Burada w, birimlerin alıcı alanlarıyla aynı boyutlara sahip matrislerdir. Seyrek bir ağırlık matrisi kullanmak ağın ayarlanabilir parametre sayısını azaltır ve böylece genelleme yeteneğini arttırır. W'yi katman girişleri ile çarpmak, girişi eğitilebilir bir filtre olarak görülebilen w ile birleştirmek gibidir. Eğer $d - 1$ evrişim tabakasına giriş ($N \times N$) boyutundaysa ve belirli bir evrişim tabakası d düzlemindeki birimlerin alıcı alanı ($m \times m$) boyutundaysa, oluşturulan özellik haritası bir boyut matrisi olacaktır (N - m + 1) × (N - m + 1). Özellikle, (i, j) konumundaki özellik eşlemi öğesi, b (bias) teriminin skaler olduğu konumda olacaktır.

$$y_{ij}^{(d)} = \sigma\left(x_{ij}^{(d)} + b\right) \tag{3.4}$$

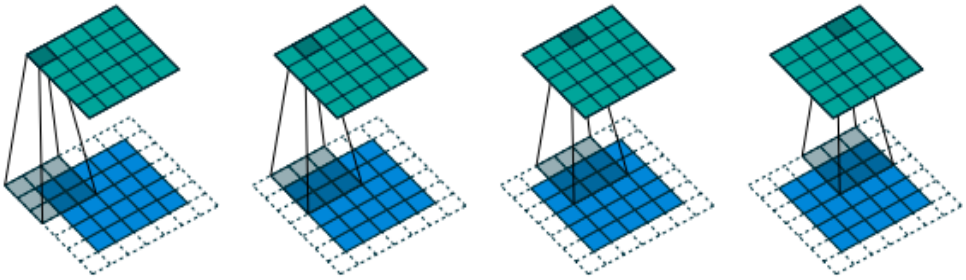
ile

$$x_{ij}^{(d)} = \sum_{\alpha=0}^{m-1} \sum_{b=0}^{m-1} w_{ab}\, y_{(i+\alpha)(j+b)}^{(d-1)} \tag{3.5}$$

Burada$(d-1)$ evrişim tabakasına giriş,b bias, m evrişim tabakasına giriş boyutu, $(i,\ j)$ konumundaki özellik eşlemi öğesi, w ağırlık matrisidir. Girişin tüm $(i,\ j)$ pozisyonları için sırayla denklem 3.4 ve 3.5'i kullanarak, karşılık gelen düzlem için özellik haritası oluşturulur. ESA'ların eğitimi ile ortaya çıkabilecek zorluklardan biri, aşırı uyum problemine yol açabilecek öğrenilmesi gereken çok sayıda parametre ile ilgilidir. Bu amaçla, stokastik havuzlama, bırakma ve veri artırma gibi teknikler önerilmiştir. Ayrıca, ESA'lar genellikle ön eğitime, yani ağı rastgele ayarlanmış parametreler yerine ön eğitimli parametrelerle başlatan bir işleme tabi tutulur. Ön eğitim, öğrenme süresini hızlandırabilir ve ağın genelleme kapasitesini artırabilir. ESA'ların çok çeşitli bilgisayar görme ve örüntü tanıma görevlerinde geleneksel makine öğrenimi yaklaşımlarından önemli ölçüde daha iyi performansa sahip olduğu gösterilmiştir. Olağanüstü performansları, eğitimdeki göreli kolaylık ile birleştiğinde, son birkaç yıldır popülaritesindeki büyük artışı açıklayan ana nedenlerdir (Çalli, 2017).

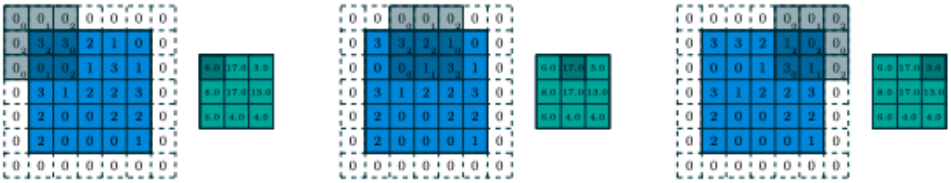
3.4.4. Dolgu ve Adım(Paddings and Stride)

Dolgu, görüntünün etrafına yerleştirilen sıfır değerlerin kenar boşluğudur. Dolgu derinliği, mevcut evrişimsel katmanın çıktısının evrişimden sonra boyutunun küçülmeyeceği şekilde ayarlanabilmektedir. Şekil 3.8'de dolgu işlemine örnek verilmiştir. Birçok ardışık evrişim katmanı, çıktı boyutundaki azalma, her evrişimde bir alan kaybolduğu için sorun haline gelebilmektedir. Birçok katmana sahip bir konvektörde, çıkışın daraltılmasının bu yapısı, girdinin sıfır doldurulmasıyla karşılanabilmektedir. Çok katmanlı ağlarda, çıktıların çok azalmasını önlemek için bu yaklaşım gerekli olabilmektedir (Murphy, 2016).



Şekil 3.8. Dolgu işlemi (Murphy, 2016).

Genellikle S olarak adlandırılan adım, yatay olarak (ve daha sonra bir sonraki satır boyunca kıvrık olarak çevrilirken) kaç piksel değişikliğinin çevrileceğini açıklayan bir parametredir. Bazı ağ mimarilerinde, katman boyutunu azaltmak için maksimum havuzlama yerine adım kullanılmaktadır. Bu yaklaşım GoogLeNet'te uygulanmıştır. Şekil 3.9'da sıfır dolgu ve adım 2 ile konvolüsyon bulunmaktadır ve sadece ilk üç alt grafik gösterilmiştir (Murphy, 2016). Dolgu derinliği 1, adım 2 ve filtre boyutu 3 × 3 kullanılmış ve sonuç matrisi büyüklüğü 3 × 3 olarak elde edilmiştir.

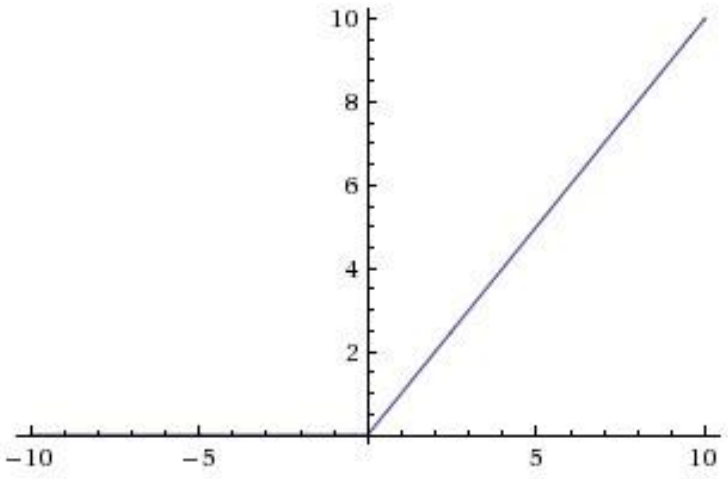


Şekil 3.9. Dolgu derinliği 1, adım 2 ve filtre boyutu 3 × 3 (Murphy, 2016).

3.4.5. Aktivasyon Fonksiyonları

Aktivasyon fonksiyonları derin öğrenmenin önemli bileşenlerindendir, yapay sinir ağlarının verilerdeki karmaşık modelleri öğrenmesine yardımcı olmaktadır. Bu tez çalışmasında ReLU ve Softmax aktivasyon fonksiyonları kullanılmıştır.

Doğrultulmuş Doğrusal Birim katmanı (Rectified Linear Unit - ReLU), son yıllarda çok popüler hale gelmiş aktivasyon fonksiyonlarından biridir. f(x) = maks. (0, x) fonksiyonunu hesaplar. Şekil 3.10'da ReLU aktivasyon fonksiyonu x <0 olduğunda sıfır ve sonra x> 0 olduğunda eğim 1 ile doğrusalı göstermektedir (Ramachandran ve ark., 2017).



Şekil 3.10. ReLU Aktivasyon fonksiyonunun çizimi

Bu aktivasyon fonksiyonunu kullanmanın çeşitli artıları ve eksileri vardı (Abdelbaki, 2018).

(+)Stokastik gradyan inişinin sigmoid / tanh fonksiyonlarına kıyasla yakınlaşmasını büyük ölçüde hızlandırdığı bulunmuştur. Bunun doğrusal, doymamış formundan kaynaklandığı iddia edilmektedir.

(+) Önemli operasyonlar (üslüler vb.) içeren tan / sigmoid nöronlara kıyasla, ReLU aktivasyon matrisini sıfırda eşleştirerek uygulanabilmektedir.

(-)Bazı gradyanlar eğitim sırasında kırılgan olabilir ve yok olabilir. Örneğin, bir ReLU nöronundan akan büyük bir gradyan, ağırlıkların nöronun bir daha hiçbir veri noktasında bir daha aktive olmayacağı şekilde güncellenmesine neden olabilir. Bu durumda, üniteden akan gradyan, o noktadan itibaren sonsuza kadar sıfır olacaktır. Yani, ReLU birimleri çeşitli verilerden çıkarılabildikleri için eğitim sırasında geri dönüşü olmayan bir şekilde yok olabilirler. Örneğin, öğrenme hızı çok yüksek ayarlanırsa, ağınızın% 40'ının yok olabileceğini (yani, tüm eğitim veri kümesinde asla aktif olmayan nöronların) olabileceğini görebilirsiniz. Öğrenme hızının uygun bir şekilde ayarlanmasıyla, bu sorunu azaltır.

Softmax Genellikle sınıflandırma görevi yapmak için inşa edilmiş çok katmanlı sinir ağlarının çıktı katmanında kullanılır. Aslında her bir sınıfın olasılığını hesaplamaktadır. Olasılıklardan bahsettiğimizden, tüm softmax'ların toplamı bir çıktı

katmanına eşittir. Sınıflar arasında katı bir seçim yapmak yerine, softmax, girdinin büyük olasılıkla hangi sınıfa ait olduğunu tahmin etmektedir. M sınıfının olduğu varsayılarak, softmax denklem (3.6)'da temsil edilebilir (Martins ve Astudillo, 2016).

$$softmx(xk) = \frac{e^{xk}}{\sum_{i=1}^{M} e^{x_i}} \tag{3.6}$$

Burada M sınıf sayısı , e^{xk} ağdan çıkan logits puanları, e^{x_i} ağdan çıkan logits puanlarının toplam sayısı değeridir.

3.4.6. Düzenleme(Regularization)

Araştırmacıların karşılaştığı en yaygın problemlerden biri, aşırı uyum gösterme (overfitting) problemidir. Aşırı uyum göstermeden kaçınmak, modelimizin performansını tek başına artırabilir. L1 ve L2, en yaygın düzenleme türleridir. Bunlar, normalleştirme terimi olarak bilinen başka bir terim ekleyerek genel maliyet işlevini güncellemektedir.

Maliyet fonksiyonu = Kayıp (örneğin, ikili çapraz entropi) + Düzenleme terimi

Bu düzenleyici terimin eklenmesi nedeniyle, ağırlık matrislerinin değerleri azalır çünkü daha küçük ağırlık matrislerine sahip bir sinir ağının daha basit modellere yol açtığını varsayılmaktadır. Bu nedenle, aynı zamanda aşırı uyum göstermeyi de bir ölçüde azaltacaktır, ancak bu düzenlileştirme terimi L1 ve L2'de farklıdır. L2'de:

$$\text{Maliyet fonksiyonu} = \text{Kayıp} + \frac{\lambda}{2m} * \sum \|w\|^2 \tag{3.7}$$

Burada λ (lambda) düzenlenme parametresidir, Kayıp (örneğin, ikili çapraz entropi) w ağırlık değeridir. Daha iyi sonuçlar için değeri optimize edilen hiper parametredir. L2 düzenlenmesi, ağırlıkları sıfıra (ama tam olarak sıfıra değil) çekmeye zorladığı için ağırlık kaybı olarak da bilinir. L1'de:

$$\text{Maliyet fonksiyonu} = \text{Kayıp} + \frac{\lambda}{2m} * \sum \|w\| \tag{3.8}$$

Burada, ağırlıkların mutlak değerini cezalandırılmaktadır. L2'den farklı olarak, ağırlıklar burada sıfıra indirilebilmektedir. Bu nedenle, modeli sıkıştırmaya çalışırken çok faydalı olmaktadır. Aksi takdirde, genellikle L2 tercih edilmektedir (Jain, 2018).

3.4.7. Toplu Normalleştirme (Batch Normalization)

Doğrusal bir model eğitildiğinde, bazı hedefleri optimize etmek için ağırlıkların güncellenmesi gerekmektedir. Doğrusal model için, girdilerin dağılımı eğitim boyunca aynı kalmaktadır. Fakat derin bir sinir ağının ortasındaki bir katmana, örneğin üçüncü katmana odaklanılırsa işler biraz farklı görünebilmektedir. Her eğitim yinelemesinden sonra, birinci ve ikinci dahil olmak üzere tüm katmanlardaki ağırlıkları güncellenmektedir. Bu, eğitim boyunca, ilk iki katmanın ağırlıkları öğrenilirken, üçüncü katmanın girdileri başlangıçta olduğundan çok farklı görünebilir.

Sergey Ioffe ve Christian Szegedy, eğitim sırasında her temsil seviyesindeki özelliklerin her birinin ortalamasını ve varyansını normalleştiren bir teknik olan toplu normallerştirmeyi önermektedir. Bu teknik, her bir mini partideki (mini-batch) örnekler boyunca özelliklerin normalleştirilmesini içermektedir. Deneysel olarak, gradyanı stabilize ediyor (yok olan değerler) ve toplu normalleştirilmiş modeller daha az uyuyor gibi görünmektedir. Toplu normalleştirilmiş modellerde nadiren bırakma (Dropout) kullanılmaktadır.

Toplu Normalleştirme, bırakmanın (Dropout) aksine, etkinleştirme katmanından sonra değil, genellikle etkinleştirme katmanından önce kullanılmaktadır. Temel fikiri normalizasyonu yapmak ve daha sonra doğrusal bir ölçek uygulamak ve mini partiye geçiş yapmaktır (Ioffe ve Szegedy, 2015).

$$\mu_B \leftarrow \frac{1}{m}\sum_{i=1}^{m} x_i \tag{3.9}$$

$$\sigma_B^2 \leftarrow \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu_B)^2 \tag{3.10}$$

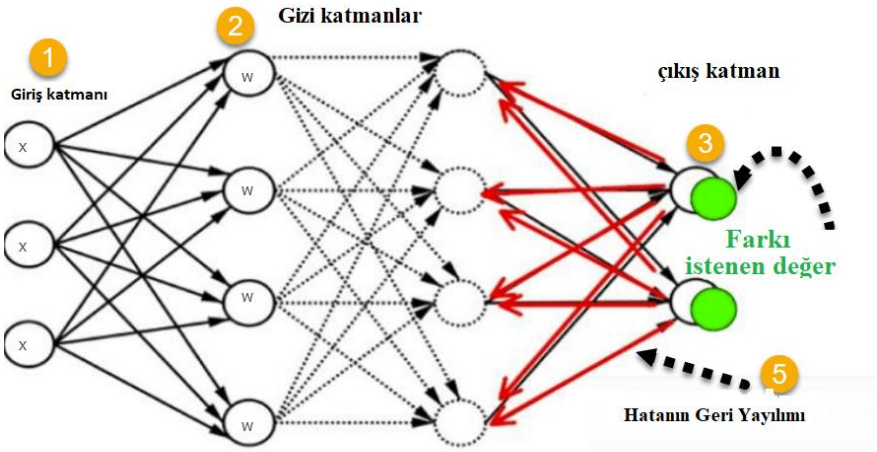
$$\hat{x_i} \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \tag{3.11}$$

$$y_i \leftarrow \gamma \hat{x_i} + \beta \equiv BN_{\gamma,\beta}(x_i) \tag{3.12}$$

Burada μ_B mini toplu ortalama , σ_B^2 mini seri varyans, $\hat{x_i}$ normalleştirmek, y_i ölçeklendirme ve kaydırma , BN Toplu Normalleştirme, γ ve β parametrelerdir , m boyut değeridir.

3.4.8. Geri yayılım

Geri yayılım algoritması, veri madenciliği, makine öğrenimi ve tahminlerin doğruluğunu arttırmada önemli bir matematiksel araç olarak kullanılmaktadır. Geri yayılma esas olarak türevleri hızlı bir şekilde hesaplamak için kullanılan bir algoritmadır. Ayrıca yapay sinir ağları, ağırlıklara göre eğim inişini hesaplamak için bir öğrenme algoritması olarak geri yayılımı kullanmaktadır. İstenen çıkışlar, elde edilen sistem çıkışlarıyla karşılaştırılmakta ve ardından sistemler, ikisi arasındaki farkı mümkün olduğunca azaltmak için bağlantı ağırlıklarını güncellemektedir. Algoritma adını ağırlıkların çıktıdan girdilere doğru geriye doğru güncellemesinden almaktadır. Ağırlıkları bias değiştirmenin yapay sinir ağının genel davranışını nasıl etkilediğini tam olarak anlamanın zorluğu, sinirsel ağ uygulamalarının daha geniş bir şekilde uygulanmasını engelleyen faktörlerden birisidir. Günümüzde, geri yayılım algoritmaları, optik karakter tanıma, doğal dil işleme ve görüntü işleme de dahil olmak üzere birçok yapay zeka alanında pratik kullanımlara sahiptir. Geri yayılma, hata fonksiyon eğimini hesaplamak için her giriş değeri için bilinen ve istenen bir çıktı gerektirdiğinden, genellikle danışmanlı makine öğrenmesi yöntemlerinde kullanılmaktadır. Naive Bayes filtreleri ve karar ağaçları gibi sınıflandırıcılarla birlikte, geri yayılım algoritması, tahmine dayalı analitik içeren makine öğrenme uygulamalarının önemli bir parçası olarak ortaya çıkmıştır. Geri yayılım algoritmasının adımları aşağıda gösterilmektedir (Riedmiller ve Braun, 1993). Şekil 3.11'de Geri yayılımın çalışma şekli görülmektedir.



Şekil 3.11. Geri yayılım (Anonymous, 2019)

1- Ara katmanlara gelen değerler denklem 3.13 kullanırak hesaplanır. (İleri doğru hesaplama)

$$NET_{\mathrm{i}} = \sum_{i=1}^{m} w_{\mathrm{i}}\, x_{\mathrm{i}} + b \tag{3.13}$$

Burada w,ağırlıklar$x$, girdiler b, bias değeridir.

2- Kullanılacak aktivasyon fonksiyonuna göre ara katmanın çıkış değeri denklem 3.14 kullanırak hesaplanır. Aktivasyon fonksiyonu olarak sigmoid aktivasyon fonksiyonu kullanıldığını varsayılarak;

$$\text{Ç}\mathrm{NET}_{\mathrm{i}} = \frac{1}{1 + e^{-(\mathrm{NET}_{\mathrm{i}})}} \tag{3.14}$$

İlk 2 aşama, girdi katmanından alınan değerlerden ara katmandaki çıktı değeri hesaplanırken bir sonraki ara katmanın yeni girdi değerinden çıktı değeri hesaplanacaktır. Bu işlemler sonuncu katmana kadar böyle devam edecektir.

3- Geri doğru hesaplama ile beklenen çıktı ile hesaplanan çıktı arasındaki fark değeri denklem 3.15'e göre hesaplanır.

$$\boldsymbol{ej} = \boldsymbol{Bj} - \text{Ç}\boldsymbol{j} \tag{3.15}$$

Burada e hata değeri, B beklenen çıktı, Ç ağın ürettiği çıktıdır. Toplam hata hesaplanırken hata karelerinin ortalaması denklem 3.16'ya göre alınacaktır.

$$E = \frac{1}{2}\sum_{j} ej^2 \tag{3.16}$$

4- Elde edilen hata değeri geriye doğru dağıtılarak ağırlıkların güncellenmesi sağlanmaktadır. Hata fonksiyonunun türevi ağırlıklara göre hesaplanmaktadır. Bunun için sigmoid fonksiyonu gibi bir aktivasyon fonksiyonu kullanılmaktadır. Ağın ürettiği çıktının aktivasyon fonksiyonu cinsinden yazılması denklem 3.17 kullanılarak hesaplanır. Toplam hatanın herhangi bir ağırlığa göre türevi denklem 3.18'de verildiği gibi zincir kuralına göre alınmaktadır. Ağırlığın yeni değeri denklem 3.19'a göre hesaplanmaktadır. Burada η; öğrenme katsayısıdır.

$$Ç_j = f(a_j) = \sum_j x_i \;\; w_{ji} \tag{3.17}$$

$$\frac{\partial E}{\partial W_{j\mathrm{i}}} = \frac{\partial E}{\partial Ç_j} \cdot \frac{\partial Ç_j}{\partial W_{ji}} \tag{3.18}$$

$$w_{ji} = w_{ji} - \eta \frac{\partial E}{\partial W_{j\mathrm{i}}} \tag{3.19}$$

3.4.9. Gradyan inişi

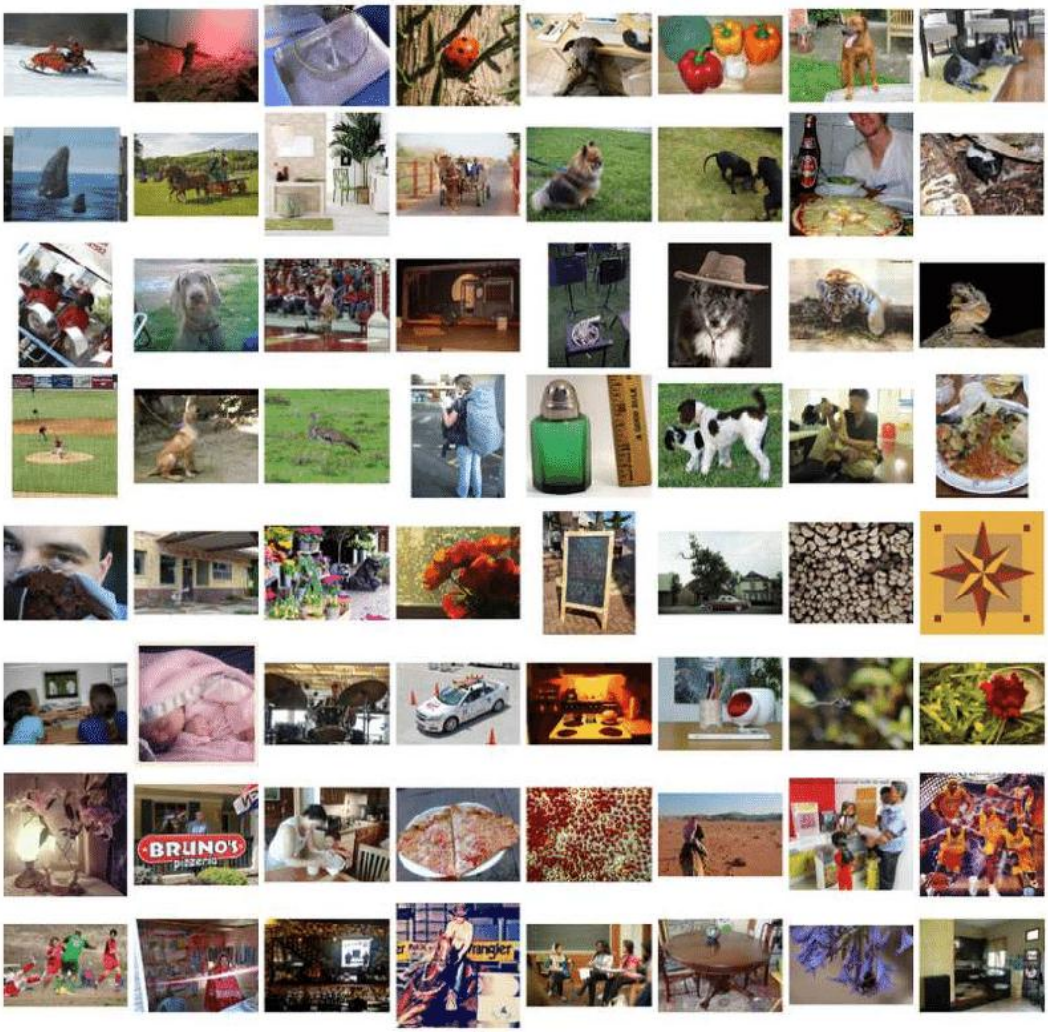
Gradyan inişi, kaybı en aza indirgemek için öğrenilebilir parametreleri tekrar tekrar güncelleyen bir optimizasyon algoritmasıdır, çıktı tahmini ile temel doğruluk etiketi arasındaki mesafeyi ölçmektedir. Kayıp fonksiyonunun gradyanı, fonksiyonun en dik artış hızına sahip olduğu yönü sağlar, ve tüm parametreler bir öğrenme oranına göre belirlenen bir adım boyutu ile degradenin negatif yönünde güncellenmektedir. Gradyan, matematiksel olarak, her öğrenilebilir parametreye göre kaybın kısmi bir türevidir,ve bir parametrenin tek bir güncellemesi denklem 3.20'de gösterilir.

$$w := w - a * \frac{\partial L}{\partial w} \tag{3.20}$$

Burada w her bir öğrenilebilir parametreyi temsil eder, α bir öğrenme oranı anlamına gelir ve L kayıp fonksiyonunu temsil eder. Öğrenme oranı, pratikte eğitim başlamadan önce ayarlanacak en önemli hiperparametrelerden biridir. Uygulamada, bellek sınırlamaları gibi nedenlerle, kayıp fonksiyonunun parametrelere ilişkin gradyanları, mini-batch adı verilen eğitim veri kümesinin bir alt kümesi kullanılarak hesaplanır ve parametre güncellemelerine uygulanır. Bu yönteme mini-batch gradyan inişi denir, genellikle stokastik gradyan iniş (SGD) olarak da adlandırılır, ve her bir mini-batch boyutu da bir hiperparametredir. Ayrıca, Gradyan iniş algoritmasında iyileştirme için birçok algoritma önerilmiş bu algoritmalardan momentum,SGD RMSprop ve Adam yaygın olarak kullanılmıştır (Yamashita ve ark., 2018).

3.5. Veri Seti

Bu tezde önerilen ESA modelinin eğitim ve validasyon işlemi için ImageNet veritabanındaki görüntüler kullanılmaktadır. ImageNet projesi, görsel nesne tanıma yazılım araştırmalarında kullanılmak üzere tasarlanmış büyük bir görsel veritabanıdır. Proje tarafından hangi nesnelerin resim olduğunu göstermek için 14 milyondan fazla görüntü açıklanmıştır ve görüntülerin en az bir milyonunda sınırlayıcı kutular da bulunmaktadır. ImageNet, "balon" veya "çilek" gibi tipik bir kategoriye sahip, birkaç yüz resimden oluşan 20.000'den fazla kategori içermektedir. Gerçek görüntüler ImageNet'e ait olmasa da, üçüncü taraf resim URL'lerinin ek açıklamalarının veritabanı doğrudan ImageNet'ten ücretsiz olarak edinilebilmektedir (Deng ve ark., 2009). Bu tezde kullanılan veri seti 39,604 görüntüden oluşmaktadır, görüntüler ImageNet veritabanından seçilmiştir, görüntülerin boyutu 256 x256'dır. Bu veriseti tez çalışmasında eğitim ve validasyon için kullanılmıştır. Ayrıca önerilen ESA tabanlı renklendirme modelini test etmek için iki ayrı veri seti kullanılmıştır. Bunlardan birincisi ctest10k veri setidir . Bu veri seti 10,001 görüntüden oluşmaktadır. İkincisi Places205 veri setidir ,bu veri seti 20,500 görüntüden oluşmaktadır ve görüntü boyutu 256 x256'dır.



Şekil 3.12. ImageNet örnek görüntüleri(Abulude ve ark., 2015)

3.6. Renklendirmenin Kalite Değerlendirmesi

Renklendirmenin kalite değerlendirmesi özneldir ve nicelleştirilmesi zordur. Bu durumlarda kalite ölçümleri, en uygun renklendirme sonuçlarına yol açan parametrelerin değerlerini gösterir. Bu nedenle, renklendirme sonuçlarının değerlendirilmesi, iyi renklendirme algoritmasının geliştirilmesinde önemli bir adımdır. Renklendirme ölçümleri, renkli görüntüdeki herhangi bir bozulmaya karşı hassas olan insan algısı ile tutarlı olmalıdır. Sistemin ne kadar başarlı olduğunu değerlendirmek amacıyla performans incelenir. Sistem performansını test etmek için aşağıdaki performans metrikleri kullanılır.

Performans metrikleri doğrudan hata matrisine bağlıdır, hata matrisi çıkarma işlemiyle elde edilebilir. Hata matrisi denklemi 3.21'deki gibi hesaplanır:

$$E[n,m] = x[n,m] - p[n,m] \tag{3.21}$$

Burada $x[n,m]$ sinir ağnın tahmin etmesi gereken hedef görüntü piksel değerini, $p[n,m]$ gerçek görüntü piksel değeri ifade etmektedir.

Bu çalışmada kullanılan performans metrikleri aşağıdaki formüller kullanılarak değerlendirilmiştir.

Ortalama Karesel Hata (Mean-squared error,MSE), bir regresyon çizgisinin bir dizi noktaya ne kadar yakın olduğunu gösterir. Bunu noktalardan regresyon çizgisine kadar olan mesafeleri alıp (bu mesafeler "hata" dır) ve onları kareleyerek yapar. MSE 3.22'de ki denklem uygulanarak hesaplanır. N tüm veri piksel sayısını temsil ettiği bilinmektedir. MSE'nin daha düşük değeri, yüksek kaliteli görüntüyü ifade eder.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(fi - yi)^2 \tag{3.22}$$

Burada N veri piksel sayısını, fi renklendirilmiş görüntünün değeri ve yi gerçek görüntünün değeridir.

Ortalama Mutlak Hata(Mean Absolute Error ,MAE) iki sürekli değişken arasındaki farkın bir ölçüsüdür. 3.23'te ki denklem tarafından verilen ortalama mutlak hata olarak adlandırılır. MAE'nin değerinin büyük olması görüntünün kalitesiz olduğu anlamına gelir.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|(fi - yi)| \tag{3.23}$$

Burada N veri piksel sayısını, fi renklendirilmiş görüntünün değeri ve yi gerçek görüntünün değeridir.

Kök Ortalama Karesel Hata (Root Mean Square Error ,RMSE) (tahmin hataları) denklem 3.24'te gösterildiği gibi ortalama mutlak hata fonksiyonunun karekökü kullanılarak değerlendirilebilir;

$$RMSE = \sqrt{\sum_{I=1}^{N} \frac{(fi - yi)^2}{N}} \tag{3.24}$$

Burada N veri piksel sayısını, fi renklendirilmiş görüntünün değeri ve yi gerçek görüntünün değeridir.

Tepe Sinyali-Gürültü Oranı(Peak Signal to Noise Rratio,PSNR) görüntüdeki piksel sinyal gürültü oranını temsil eder. PSNR'nin düşük değeri, görüntü kalitesinin düşük olmasına neden olur

$$\mathrm{PSNR} = 10 \log_{10}\left(\frac{\mathrm{MAX^2_i}}{\mathrm{MSE}}\right) \tag{3.25}$$

Burada $\mathrm{MAX^2_i}$ gerçek görüntü veri tipindeki maksimum dalgalanmadır ve MSE ortalama karesel hatadır.

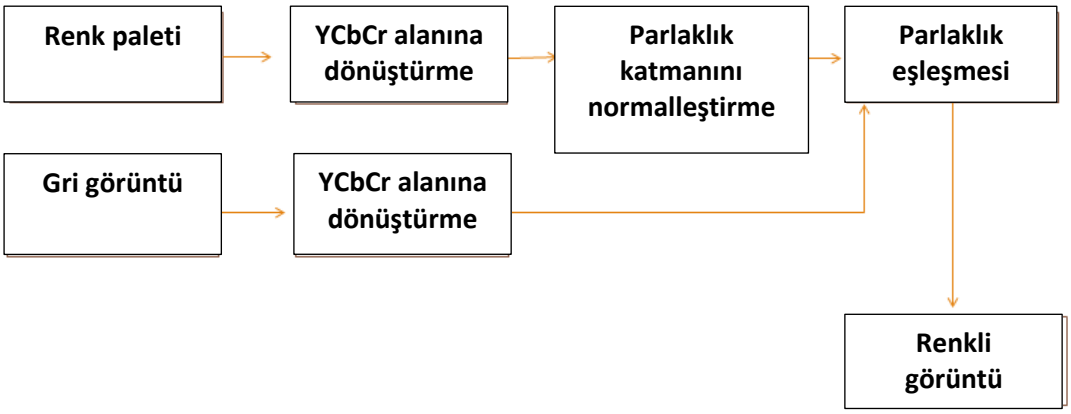
Son performans kriteri olarak renklendirme süresi kullanılmıştır.Simülasyonun (yalnızca test süresi) her görüntü için renklendirmenin sürdüğü saniye cinsinden süre miktarını temsil eder.

4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA

Bu tez çalışmasında referans kullanarak ve ESA tabanlı renklendirme yöntemleriyle otomatik renklendirme gerçekleştirilmiştir. Bu bölümde kullanılan renklendirme yöntemleri anlatılmakta ve önerilen ESA tabanlı tam otomatik görüntü renklendirme mimarisi tartışılmaktadır.

4.1. Referans Kullanarak Otomatik Renklendirme

Referans tabanlı görüntü renklendirme ile, renk paleti olarak bir referans görüntü kullanarak gri tonlamalı görüntüleri otomatik bir şekilde renklendirme yapılmıştır. Görüntüler ilk olarak YCbCr renk uzayına dönüştürülür, sonra parlaklık katmanı normalize edilir ve son olarak parlaklık eşleştirmesi yapılır. Şekil 4.1'de işlemin şeması gösterilmektedir.



Şekil 4.1. Referans kullanarak otomatik renklendirme yönteminin akış şeması

Şekil 4.2'de görüntüyü renklendirmek için girdi olarak gri tonlamalı bir görüntü ve renk paleti olarak bir referans görüntü kullanan yöntemin sonucu gösterilmektedir.



Şekil 4.2. Referans kullanarak otomatik renklendirme örneği

Çalışmamamızda ilk olarak orijinal görüntüler gri tonlamalı görüntülere çevrilmiştir. Ardından referans görüntü verilerek renkli görüntü tahmini yapılmıştır. Referans görüntü otomatik renklendirme yöntemi kullanılarak elde edilen renkli görüntü ile orijinal renkli görüntü arasında kalite değerlendirmek için RMSE, MSE, MAE ve PSNR kullanılmıştır. Çizelge 4.1'de renklendirme sonuçları gösterilmektedir.

Çizelge 4.1. Referans kullanarak otomatik renklendirme sonuçları

Gerçek Görüntü	Gri Görüntü	Referans Görüntü	Renkli Görüntü	RMSE değeri	MSE değeri	MAE değeri	PSNR değeri	Süre
				9.11	83.04	240.99	31.23	14.28
				8.86	78.62	249.61	29.18	13.69
				8.91	79.49	222.50	29.13	13.81
				9.26	85.79	234.99	28.8	14.12
				8.89	79.17	226.38	29.15	14.25

Çizelge 4.1de referans kullanarak otomatik renklendirme yönteminin sonuçları bulunmaktadır. Gerçek görüntü, gri görüntü, referans görüntü ve modelin elde etiği renkli görüntü gösterilmektedir. Kullanılan görüntü boyutu ise 308x230 olarak belirlenmiştir. Renklendirme sonucunu değerlendirmek içinde RMSE, MSE, MAE, PSNR ve referans verildikten sonraki renklendirme süresi kullanılmıştır.

4.2. Önerilen ESA Modeli Kullanarak Renklendirme

Bu tez çalışmasında, ESA tabanlı tam otomatik bir görüntü renklendirme modeli önerilmiştir. Literatürdeki ESA kullanılarak yapılan başlıca renklendirme çalışmalarında (Zhang ve ark., 2016), toplu normalizasyonu, örnekleme katmanı ve evrişimsel katmanları bulunmaktadır. Önerilen ESA yapısı da bu katmanlar kullanılarak oluşturulmuştur. Mimari 23 evrişimsel katmandan oluşmaktadır. Her evrişimsel katman, 0.001 değerine sahip olan L2 Düzenleyici ile düzenlenmiştir. Kullanılan evrişim

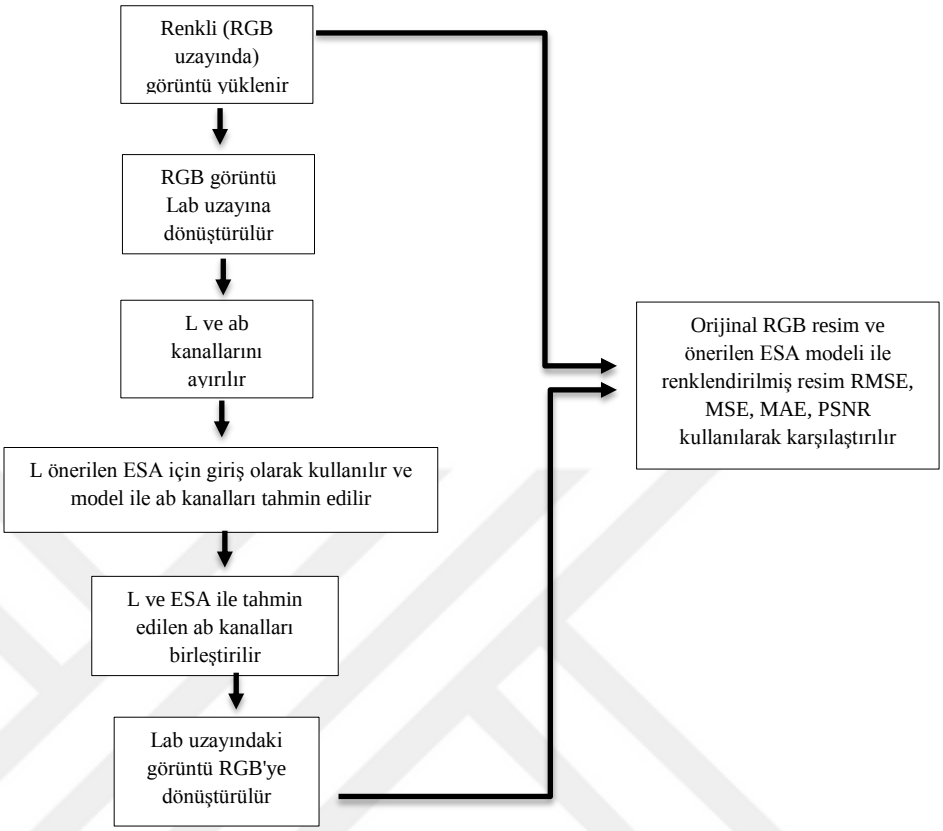
tabakasının çekirdek boyutu 3x3'tür, evrişimsel tabakaların hepsinde ReLU aktivasyonu kullanılmaktadır. Kullanılan dolgu değeri 1'dir ve tüm katmanlarda aynıdır. Modelin son katmanı softmax aktivasyonunu kullanır. Tablo 4.2 mimariyi göstermektedir.

Çizelge 4.2. Renklendirme modeli mimarisi

Katmanlar	Filtreler	Aktivasyon	Düzenlilik
Conv2D	64	ReLU	L2
Conv2D	64	ReLU	L2
Toplu Normalleştirme	Yok	Yok	Yok
Conv2D	128	ReLU	L2
Conv2D	128	ReLU	L2
Toplu Normalleştirme	Yok	Yok	Yok
Conv2D	256	ReLU	L2
Conv2D	256	ReLU	L2
Conv2D	256	ReLU	L2
Toplu Normalleştirme	Yok	Yok	Yok
Conv2D	512	ReLU	L2
Conv2D	512	ReLU	L2
Conv2D	512	ReLU	L2
Toplu Normalleştirme	Yok	Yok	Yok
Conv2D	512	ReLU	L2
Conv2D	512	ReLU	L2
Conv2D	512	ReLU	L2
Toplu Normalleştirme	Yok	Yok	Yok
Conv2D	512	ReLU	L2
Conv2D	512	ReLU	L2
Conv2D	512	ReLU	L2
Toplu Normalleştirme	Yok	Yok	Yok
Conv2D	256	ReLU	L2
Conv2D	256	ReLU	L2
Conv2D	256	ReLU	L2
Toplu Normalleştirme	Yok	Yok	Yok
Örnekleme 2D	Yok	Yok	Yok
Conv2D	128	ReLU	L2
Conv2D	128	ReLU	L2
Conv2D	128	ReLU	L2
Toplu Normalleştirme	Yok	Yok	Yok
Conv2D	313	Softmax	Yok

Model, ImageNet veri kümesinden alınan görüntüler kullanılarak eğitilir, eğitim kümesindeki verilerin %80'i eğitim %20 si validasyon için olacak şekilde bölünmüştür. Eğitim için 31683 görüntü ve validasyon için 7921 görüntü kullanılmıştır. SGD optimizer modeli eğitmek ve modelin hiperparametrelerini (ağırlıkları) optimize etmek için kullanılır. Optimize edici 0.001'e ve itme değeri 0.9'a ayarlanmıştır. Eğitim için kullanılan kayıp fonksiyonu kategorik çapraz entropidir.

Model performansını hesaplamak için RMSE, MSE, MAE, PSNR gibi çoklu yöntemler kullanılmıştır. Önerilen ESA tabanlı renklendirme yönteminin akış şeması şekil 4.3'te gösterilmektedir.



Şekil 4.3. Önerilen ESA tabanlı renklendirme yönteminin akış şeması

Önerilen ESA tabanlı renklendirme yönteminde renkli görüntüler (RGB uzayında) yüklenir ve sonra (Lab) uzayına dönüştürülür. Daha sonra görüntü kanalları L ve ab parçalarına ayırılır. ab kanalının ağ çıkışı olarak, ağın girişi L kanalı olarak kullanılmaktadır. Ardından L kanalı ağa giriş olarak verildikten sonra ESA kullanark tahmin edilen ab kanalıyla birleştirilir böylece Lab uzayında bir renkli görüntü elde edilir. Elde ettiğimiz Lab uzayındaki görüntütü RGB uzayına dönüştürülür. Daha sonra renklendirme işlemini test etmek için modelin oluşturduğu renkli görüntü orjinal görüntü ile karşılaştırılır ve RMSE, MSE, MAE, PSNR fonksiyonlarını kullanarak sonuç değerlendirilir.

Önerilen otomatik renklendirme modeli ESA kullanarak ImageNet veri setinden rastgele seçilen görüntüler üzerinde eğitilmiştir. Veri setinin üç defa rassal % 80’i eğitim için ve % 20'si validasyon için olacak şekilde bölünmesiyle elde edilen sonuçlar aşağıdaki Çizelge 4.3'te gösterilmektedir.

Çizelge 4.3 Validasyon verilerinden elde edilen performans ölçüm sonuçları

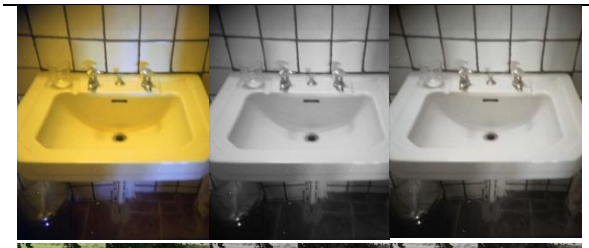
Çalışmalar	RMSE	MSE	MAE	PSNR	Renklendirme süresi
Birinci çalışma	7.20	53.72	124.12	31.13	1.03
Ikinciçalışma	7.16	53.10	124.04	31.19	0.84
Üçüncü çalışma	7.24	54.29	124.09	31.09	0.86
Ortalama	7.20	53.70	124.08	31.13	0.91

Validasyon sonuçları çizelge 4.3'te gösterilmektedir. RMSE'nin maksimum değeri 7.24 iken minimum değeri 7.16'dır ve ortalama değeri ise 7.20'dir. MSE'nin maksimum değeri 53.72'iken minimum değeri 53.10'dur ve ortalama değeri ise 53.70'tir. MAE'nin maksimum değeri 124.12' iken minimum değeri 124.04'tür ve ortalama değeri ise 124.08'dir. PSNR'nin maksimum değeri 31.13' iken minimum değeri ise 31.09'dur ve ortalama değeri ise 31.13'tür. Son olarak işlem süresi (renklendirme süresi) her bir görüntü için renklendirmenin sürdüğü saniye cinsinden süre miktarı 3 ayrı çalışmadan elde edilen sürelerin ortalama değeri ise 0.91 saniye olarak bulunmuştur.

Çizelge 4.4'te önerilen ESA kullanılarak rassal seçilen 15 örnek görüntü ile renklendirme işleminden elde edilen renklendirme sonuçları sunulmuştur.

Çizelge 4.4. Önerilen modeli kullanarak otomatik renklendirme sonuçları

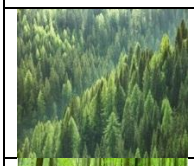
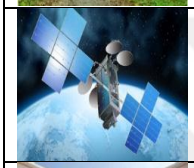
Gerçek görüntü	**Gri tonlamalı görüntü**	**Çıktı görüntü**	**RMSE**	**MSE**	**MAE**	**PSNR**	**Renklendirme süresi**
			7.38	54.48	125.9	30.77	1.22
			7.20	51.81	114.7	30.99	0.80

	3.85	14.81	111.7	36.42	0.78
	5.76	33.19	106.7	32.92	0.76
	7.91	62.52	121.9	30.17	0.76
	5.26	27.65	117.4	33.71	0.76
	6.60	43.59	119.2	31.74	0.76
	6.94	48.17	127.0	31.30	0.76
	7.48	55.98	141.6	30.65	0.75

	6.72	45.19	125.4	31.58	0.75
	8.15	66.48	138.2	29.90	0.78
	7.87	62.00	93.0	30.21	0.75
	6.61	43.66	120.2	31.73	0.76
	6.54	42.73	107.6	31.82	0.75
	8.75	76.50	108.6	29.29	0.77
Ortalama	**6.86**	**48.58**	**118.6**	**31.54**	**0.79**

Çizelge 4.5'te referans kullanarak ve önerilen ESA tabanlı renklendirme modeli kullanarak renklendirme sonuçlarının karşılaştırılması gösterilmektedir. Aynı görüntüler 2 yöntem kullanılarak renklendirilmiş ve RMSE, MSE, MAE, PSNR ve renklendirme süresi bakımından karşılaştırılmıştır.

Çizelge 4.5. Referans kullanarak ve önerilen ESA tabanlı renklendirme modeli kullanarak renklendirme sonuçları

Görüntüler	**1.Yöntem (Referans) kullanarak renklendirme**					**2.Yöntem Önerilen (ESA) tabanlı renklendirme**				
	RMSE	MSE	MAE	PSNR	Süre	RMSE	MSE	MAE	PSNR	Süre
	9.11	83.04	240.99	31.23	14.28	8.70	75.69	154.20	29.34	0.91
	8.86	78.62	249.61	29.18	13.69	8.90	79.29	143.10	27.98	0.85
	8.91	79.49	222.50	29.13	13.81	8.29	68.76	90.99	28.97	0.81
	9.26	85.79	234.99	28.8	14.12	8.14	66.33	134.28	29.91	0.76
	8.89	79.17	226.38	29.15	15.39	7.10	50.47	111.42	31.09	0.77
Ortalama	**9.006**	**81.22**	**234.89**	**29.49**	**14.25**	**8.22**	**68.10**	**126.798**	**29.45**	**0.82**

Önerdiğimiz ESA mimarisini test etmek için iki ayrı veri seti daha kullanılmıştır. Bunlardan birincisi ctest10k ve ikincisi ise Places205 veri setidir. Test sonuçlarını değerlendirmek için RMSE, MSE, MAE, PSNR ve renklendirme süresi değerleri bulunmuştur. Modelin elde ettiği sonuçlar çizelge 4.6'da bulunmaktadır.

Çizelge 4.6. Önerilen ESA modelinin ctest10k ve Places205verisetleri üzerinde elde etiği test sonuçları

veri seti	RMSE	MSE	MAE	PSNR	Renklendirme süresi
ctset10k	**5.189889**	**29.14883**	**114.5571**	**34.23678**	**0.832537**
Places205	**7.098634**	**52.15844**	**120.4613**	**31.28532**	**0.840554**

Önerdiğimiz ESA tabanlı renklendirme modeli kullanılarak ctest10k veri setinde rassal seçilmiş 10 görüntünün renklendirme sonuçları Çizege 4.7'de verilmiştir.

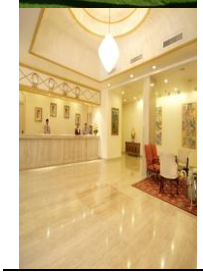
Çizelge 4.7. Ctest10k veri setinde rassal seçilmiş 10 görüntünü

Gerçek görüntü	Gri tonlamalı görüntü	ESA modeli ile renklenen görüntü	RMSE	MSE	MAE	PSNR	Renklendirme süresi
			4.36	19.01	120.7	35.33	0.83
			7.84	61.49	147.3	30.24	0.81
			5.83	34.05	143.6	32.80	0.82
			3.83	14.67	110.5	36.46	0.84
			6.08	36.98	109.8	32.44	0.84
			5.31	28.25	100.8	33.61	0.82

			5.41	29.31	126.7	33.45	0.81
			5.38	28.95	105.3	33.51	0.84
			7.27	52.85	117.8	30.89	0.82
			7.98	63.71	130.0	30.08	0.84
	Ortalama		**5.93**	**36.93**	**121.2**	**32.88**	**0.83**

Önerdiğimiz ESA tabanlı renklendirme modeli kullanılarak Places205 veri setinde rassal seçilmiş 10 görüntünün renklendirme sonuçları Çizelge 4.8'de verilmiştir.

Çizelge 4.8. Places205 veri setinde rassal seçilmiş 10 görüntünün

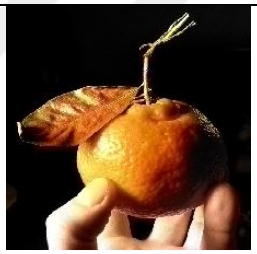
Gerçek görüntü	Gri tonlamalı görüntü	ESA modeli ile renklenen görüntü	RMSE	MSE	MAE	PSNR	Renklendirme süresi
			7.16	51.38	129.7	31.02	0.98
			7.16	51.36	135.3	31.02	0.79
			5.79	33.60	121.0	32.86	0.82
			8.520	72.60	95.97	29.52	0.82
			6.38	40.82	105.2	32.02	0.79
			8.17	66.79	141.3	29.88	0.81

			7.13	50.87	132.3	31.06	0.81
			5.98	35.84	108.9	32.58	0.84
			7.85	61.66	141.0	30.23	0.811
			8.921	79.58	151.5	29.12	0.79
	Ortalama		**7.31**	**54.45**	**126.2**	**30.93**	**0.82**

Zhang ve ark. (Zhang ve ark., 2016) tarafından ctest10k veri setini kullanarak gerçekleştirilmiş renklendirme işlemini önerilen yöntemle karşılattırmak için 10 rassal resim seçilmiştir (Zhang, 2016). Zhang ve ark. (Zhang ve ark., 2016) tarafından elde edilen sonuçlar ve önerilen ESA tabanlı renklendirme sonuçları Çizelge 4.9'da gösterilmiştir.

Çizelge4.9. Ctest10k veri setini kullanarak litertürde bulunan ve ESA modeli kullanarak renklendirme sonuçları

Gerçek görüntü	(Zhang ve ark., 2016)	Önerilen ESA ile renklendirme sonuçları
	 **MSE= 93.15** **PSNR= 28.44**	 **MSE= 54.6** **PSNR= 30.76**
	 **MSE= 83.98** **PSNR= 28.89**	 **MSE= 56.02** **PSNR= 30.65**
	 **MSE= 87.24** **PSNR= 28.72**	 **MSE= 72.5** **PSNR= 29.53**
	 **MSE= 87.28** **PSNR= 28.72**	 **MSE= 77** **PSNR= 29.27**
	 **MSE= 85.3** **PSNR= 28.82**	 **MSE= 74.19** **PSNR= 29.43**

	 **MSE=88.24** **PSNR=28.67**	 **MSE= 107.27** **PSNR= 27.83**
	 **MSE= 81.38** **PSNR= 29.03**	 **MSE= 72.22** **PSNR= 29.54**
	 **MSE= 67.67** **PSNR= 29.83**	 **MSE= 57.56** **PSNR= 30.53**
	 **MSE= 38.07** **PSNR= 32.32**	 **MSE= 38.54** **PSNR= 32.27**
	 **MSE= 66.88** **PSNR= 29.88**	 **MSE= 68.16** **PSNR= 29.8**
Ortalama	**MSE= 77.919** **PSNR= 29.332**	**MSE= 67.806** **PSNR= 29.961**

5. SONUÇLAR VE ÖNERİLER

5.1 Sonuçlar

Bu tez çalışmasında gri tonlamalı (seviye) görüntülerin otomatik renklendirme işlemi gerçekleştirilmiştir. Renklendirme için iki model önerilmiştir, bunlardan ilki renk paleti olarak bir referans görüntüsü kullanarak gri tonlamalı görüntüleri otomatik bir şekilde renklendirme yapmaktadır. Görüntüler YCbCr renk uzayına dönüştürülür sonra parlaklık katmanı normalize edilir ve parlaklık eşleştirmesini yaparak renklendirme işlemi gerçekleştirilir. Kullanılan bu yöntemde görüntü boyutu 308x230 olarak belirlenmiş olup çoğunlukla yöntem az miktarda görüntü renklendirmede iyi sonuç vermektedir. Ayrıca renklendirme işleminin aldığı süre diğer modele göre daha fazladır ve referans görüntünün kullanıcı tarafından belirlenmesi gerekmektedir.

İkinci model olarak ESA tabanlı renklendirme yöntemi kullanılmıştır. Kullanılan veri seti 39.604 görüntüden oluşmakta olup görüntülerin boyutu 256 x256'dır. Bu görüntülerin % 80'i eğitim için % 20'si validasyon için ayrılmıştır. SGD iyileştirici, modeli eğitmek ve modelin hiperparametrelerini (ağırlıkları) optimize etmek için kullanımlıştır. Elde edilen renklendirme sonuçları, RMSE, MSE, MAE ve PSNR kullanılarak değerlendirilmiştir. Orijinal resimle modelin çıkışından elde edilen resimler arasındaki benzerliğin büyük olması modelimiz renklendirme işlemini doğru bir şekilde yapmış olduğunu gösterir. Dolayısıyla benzerlik ne kadar fazlaysa sonuç o kadar başarılı elde edilmiştir.

İki yöntemi karşılaştırmak için rassal seçilen 5 görüntü bu yöntemler kullanılarak renklendirilmiş ve sonuçları Çizelge 4.5'te verilmiştir. İki yöntem arasında bir karşılaştırma yapılırsa ESA modeli kullanılarak yapılan görüntü renklendirme insan müdahalesini büyük oranda ortadan kaldırmaktadır. ESA modeli ortalama 0.82 saniyede bir görüntü renklendirme işlemini yaparken, referans modeli uygun referans bulunduktan ve yönteme verildikten sonra sadece renk aktarma işlemini ortalama 14.25 saniyede gerçekleştirmektedir. Referans bulma ve ekleme süreleri düşünüldüğünde referans ile renklendirme yönteminin süresi daha da artacaktır. Sonuç olarak ESA modeli daha kısa sürede kaliteli sonuçlar vermektedir. En büyük özelliğiyse yüzlerce resimleri kısa bir süre içerisinde iyi sonuç vererek renklendirmesidir. Dezavantajı ise eğitim süresi uzun zaman almaktadır. Referans kullanarak yapılan renklendirme sonuçları iyi performans göstermesine rağmen referansı insan tarafından verilmesi

gerekmektedir. Ayrıca her bir resim için uygun bir referans görüntü gerektiğinden dolayı, çok sayıda görüntü renklendirme işlemi yapılmak istendiğinde bu yöntem önerilmemektedir.

Bu tezde önerilen ESA kullanılarak renklendirme modelini test etmek için iki ayrı veri seti kullanılmıştır. Bunlardan birincisi ctest10k ve ikincisi Places205'tir. Test sonuçlarının performansını değerlendirmek için RMSE, MSE, MAE, PSNR ve renklendirme süresi değerleri kullanılmıştır. ctest10k veri seti ile elde edilen ortalama RMSE, MSE, MAE, PSNR ve renklendirme süresi değerleri sırasıyla 5.189889, 29.144883, 114.5571, 34.23678, 0.832537 olarak, Places205 veri seti ile elde edilen ortalama RMSE, MSE, MAE, PSNR ve renklendirme süresi değerleri sırasıyla 7.098634, 52.15844, 120.4613, 31.28532, 0.840554 olarak elde edilmiştir. Çizelge 4.7'de ctest10k veri setinden ve Çizelge 4.8'de Places205 veri setinden rassal seçilen 10 görüntüyü önerilen ESA modeli kullanarak renklendirme sonucunda elde edilen görüntüler gösterilmektedir. Çizelge 4.9'da ctest10k veriseti kullananan literatür çalışmasıyla 10 rassal görüntü seçilerek karşılaştırma yapılmıştır. Elde edilen sonuçlara bakıldığında önerilen modelin başarılı bir şekilde renklendirme işlemini gerçekleştirdiği gözükmektedir.

5.2 Öneriler

Bu çalışmada, derin öğrenme yaklaşımına dayalı otomatik bir görüntü renklendirme tanıtılmıştır. Gri tonlamalı görüntülere renk eklemek için otomatik renklendirme tekniğinin kullanılması, zaman ve insan emeğinden tasarruf sağlayacaktır. Çünkü kullanıcı tabanlı renklendirme, görüntüler daha fazla nesneye ve karmaşık bir sahneye sahip olduğunda zaman alıcı bir görevdir. Bu tekniğin geliştirilmesi ve çok miktarda veri seti ve daha karmaşık görüntüler kullanılması, renklendirme sonuçlarını iyileştirecektir.

Gelecekteki çalışmalarda ResNet kullanarak daha fazla derinliği olan(daha fazla hesaplama kaynağı gerektirir) bir model oluşturup gri tonlamalı resimleri renklendirmek planlanmaktadır. Bu modeller doğal resimlerde uygulandığında oldukça etkileyici sonuçlar vermektedir. Ayrıca resim üzerinde boyanmasını istediğimiz alanları karalama yaparak, ESA ile sadece karalanmış olan bölgedeki nesneleri tanıyan bir model geliştirmek hedeflenmektedir.

KAYNAKLAR

Abdelbaki, A., 2018, ConvNet Features for Lifelong Place Recognition and Pose Estimation in Visual SLAM.

Abulude, F., Akinyinka, A. ve Adeyemi, A., 2015, GLOBAL POSITIONING SYSTEM AND IT'S WIDE APPLICATIONS, *Continental J. Information Technology*.

Aghdam, H. H. ve Heravi, E. J. J. N. Y., NY: Springer. doi, 2017, Guide to Convolutional Neural Networks, 10, 978-973.

Al Azzeh, J., Alhatamleh, H., Alqadi, Z. A. ve Abuzalata, M. K. J. I. J. o. C. A., 2016, Creating a Color Map to be used to Convert a Gray Image to Color Image, 153 (2), 31-34.

Anonymous, 2019, Back Propagation Neural Network: Explained With Simple Example.

Aras, S. ve Gangal, A., 2009, Görüntü Renklendirme Teknikleri ve Uygulamaları Üzerine Bir Araştırma.

Baynes, E. R., Van de Lageweg, W. I., McLelland, S. J., Parsons, D. R., Aberle, J., Dijkstra, J., Henry, P.-Y., Rice, S. P., Thom, M. ve Moulin, F. J. E.-S. R., 2018, Beyond equilibrium: re-evaluating physical modelling of fluvial systems to represent climate changes, 181, 82-97.

Bouhal, T., Agrouaz, Y., Kousksou, T., Allouhi, A., El Rhafiki, T., Jamil, A., Bakkas, M. J. R. ve Reviews, S. E., 2018, Technical feasibility of a sustainable Concentrated Solar Power in Morocco through an energy analysis, 81, 1087-1095.

Çalli, E., 2017, Faster Convolutional Neural Networks,.

Charpiat, G., Hofmann, M. ve Schölkopf, B., 2008, Automatic image colorization via multimodal predictions, *European conference on computer vision*, 126-139.

Cheng, Z., Yang, Q. ve Sheng, B., 2015, Deep colorization, *Proceedings of the IEEE International Conference on Computer Vision*, 415-423.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K. ve Fei-Fei, L., 2009, Imagenet: A large-scale hierarchical image database, *2009 IEEE conference on computer vision and pattern recognition*, 248-255.

Deshpande, A., Rock, J. ve Forsyth, D., 2015, Learning large-scale automatic image colorization, *Proceedings of the IEEE International Conference on Computer Vision*, 567-575.

Dumoulin, V. ve Visin, F. J. a. p. a., 2016, A guide to convolution arithmetic for deep learning.

Futschik, D., 2018, Colorization of black-and-white images using deep neural networks.

Gatys, L. A., Ecker, A. S. ve Bethge, M., 2016, Image style transfer using convolutional neural networks, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2414-2423.

Gupta, R. K., Chia, A. Y.-S., Rajan, D., Ng, E. S. ve Zhiyong, H., 2012, Image colorization using similar images, *Proceedings of the 20th ACM international conference on Multimedia*, 369-378.

Hadhoud, M., ElKilani, W., Semary, N. ve Ismail, N., 2007, A texture recognition coloring technique for natural gray images, *2007 International Conference on Computer Engineering & Systems*, 237-245.

Hamam, T., Dordek, Y. ve Cohen, D., 2012, Single-band infrared texture-based image colorization, *2012 IEEE 27th Convention of Electrical and Electronics Engineers in Israel*, 1-5.

He, K., Zhang, X., Ren, S., Sun, J. J. I. t. o. p. a. ve intelligence, m., 2015, Spatial pyramid pooling in deep convolutional networks for visual recognition, 37 (9), 1904-1916.

Hsieh, C.-H., Lin, C.-M. ve Chang, F.-J., 2009, Pseudo-coloring with Histogram Interpolation, *2009 Ninth International Conference on Hybrid Intelligent Systems*, 9-12.

Huang, Y.-C., Tung, Y.-S., Chen, J.-C., Wang, S.-W. ve Wu, J.-L., 2005, An adaptive edge detection based colorization algorithm and its applications, *Proceedings of the 13th annual ACM international conference on Multimedia*, 351-354.

Hwang, J. ve Zhou, Y., 2016, Image colorization with deep convolutional neural networks, In: Stanford University, Tech. Rep., Eds, p.

Iizuka, S., Simo-Serra, E. ve Ishikawa, H. J. A. T. o. G., 2016, Let there be color! Joint end-to-end learning of global and local image priors for automatic image colorization with simultaneous classification, 35 (4), 1-11.

Ioffe, S. ve Szegedy, C. J. a. p. a., 2015, Batch normalization: Accelerating deep network training by reducing internal covariate shift.

Ironi, R., Cohen-Or, D. ve Lischinski, D., 2005, Colorization by Example, *Rendering Techniques*, 201-210.

Jain, S. J. A. V., 2018, An Overview of Regularization Techniques in Deep Learning (with Python code), 19.

Ji, Y. ve Chen, Y., 2008, Rendering greyscale image using color feature, *2008 International Conference on Machine Learning and Cybernetics*, 3017-3021.

Kalyan, A. ve Ramalingam, M., 2019, Image Colorization Using Convolutional Neural Networks.

Kok, C., Hui, Y. ve Nguyen, T., 1996, Medical image pseudo coloring by wavelet fusion, *Proceedings of 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 648-649.

Kumar, S. ve Swarnkar, A., 2012, Gray image colorization in ycbcr color space, *2012 1st International Conference on Emerging Technology Trends in Electronics, Communication & Networking*, 1-6.

Kuzovkin, D., Chamaret, C. ve Pouli, T., 2015, Descriptor-based image colorization and regularization, *International Workshop on Computational Color Imaging*, 59-68.

Larsson, G., Maire, M. ve Shakhnarovich, G., 2016, Learning representations for automatic colorization, *European Conference on Computer Vision*, 577-593.

LeCun, Y., Bottou, L., Bengio, Y. ve Haffner, P. J. P. o. t. I., 1998, Gradient-based learning applied to document recognition, 86 (11), 2278-2324.

Levin, A., Lischinski, D. ve Weiss, Y., 2004, Colorization using optimization, In: ACM SIGGRAPH 2004 Papers, Eds, p. 689-694.

Limmer, M. ve Lensch, H. P., 2016, Infrared colorization using deep convolutional neural networks, *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 61-68.

Luan, Q., Wen, F., Cohen-Or, D., Liang, L., Xu, Y.-Q. ve Shum, H.-Y., 2007, Natural image colorization, *Proceedings of the 18th Eurographics conference on Rendering Techniques*, 309-320.

Martins, A. ve Astudillo, R., 2016, From softmax to sparsemax: A sparse model of attention and multi-label classification, *International Conference on Machine Learning*, 1614-1623.

Mitchell, T. M., 2006, The discipline of machine learning, Carnegie Mellon University, School of Computer Science, Machine Learning …, p.

Mohammed, M., Khan, M. B. ve Bashier, E. B. M., 2016, Machine learning: algorithms and applications, Crc Press, p.

Murphy, J. J. M. I., 2016, An overview of convolutional neural network architectures for deep learning.

Nguyen, T., Mori, K. ve Thawonmas, R. J. a. p. a., 2016, Image colorization using a deep convolutional neural network.

Niu, Z., Zhou, M., Wang, L., Gao, X. ve Hua, G., 2017, Hierarchical multimodal lstm for dense visual-semantic embedding, *Proceedings of the IEEE International Conference on Computer Vision*, 1881-1889.

Özkan, İ. ve Ülker, E. J. G. B. A. D., 2017, Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri, 6 (3), 85-104.

Patel, D. ve Parmar, S., 2013, Image retrieval based automatic grayscale image colorization, *2013 Nirma University International Conference on Engineering (NUiCONE)*, 1-6.

Premoˇvze, S. ve Thompson, W. B., 2002, Automated Coloring of Panchromatic Orthoimagery, *3rd ICA Mountain Cartography Workshop, Mt. Hood, Oregon, USA*.

Qin, W., Wang, L., Lin, A., Zhang, M., Xia, X., Hu, B., Niu, Z. J. R. ve Reviews, S. E., 2018, Comparison of deterministic and data-driven models for solar radiation estimation in China, 81, 579-594.

Ramachandran, P., Zoph, B. ve Le, Q. V. J. a. p. a., 2017, Searching for activation functions.

Raseman, W. J., Kasprzyk, J. R., Rosario-Ortiz, F. L., Stewart, J. R., Livneh, B. J. E. S. W. R. ve Technology, 2017, Emerging investigators series: a critical review of decision support systems for water treatment: making the case for incorporating climate change and climate extremes, 3 (1), 18-36.

Riedmiller, M. ve Braun, H., 1993, A direct adaptive method for faster backpropagation learning: The RPROP algorithm, *IEEE international conference on neural networks*, 586-591.

Ruderman, D. L., Cronin, T. W. ve Chiao, C.-C. J. J. A., 1998, Statistics of cone responses to natural images: implications for visual coding, 15 (8), 2036-2045.

Rujuta, 2013, R. M. J. I. J. o. C., Information Technology ve Bioinformatics, 2013, Converting Grayscale Image to Color Image, 1.

Site, 2020, https://en.wikipedia.org/wiki/Grayscale, [Ziyarat Tarihi: 13 Ocak 2019],

Tai, Y.-W., Jia, J. ve Tang, C.-K., 2005, Local color transfer via probabilistic segmentation by expectation-maximization, *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 747-754.

Varga, D. I. ve Szirányi, T., 2017, Convolutional Neural Networks for automatic image colorization.

Vieira, L. F. M., Vilela, R. D., Nascimento, E., Fernandes, F., Carceroni, R. L. ve Araújo, A. d. A., 2003, Automatically choosing source color images for coloring grayscale images, *16th Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI 2003)*, 151-158.

Voulodimos, A., Doulamis, N., Doulamis, A., Protopapadakis, E. J. C. i. ve neuroscience, 2018, Deep learning for computer vision: A brief review, 2018.

Wang, C., Wang, Z., Kong, Y., Zhang, F., Yang, K. ve Zhang, T. J. S. r., 2019, Most of the northern hemisphere permafrost remains under climate change, 9 (1), 1-10.

Welsh, T., Ashikhmin, M. ve Mueller, K., 2002, Transferring color to greyscale images, *Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, 277-280.

Wu, J. J. P. o. a. h. c. n. e. c. w. t. C. p., 2017, Convolutional neural networks.

Xiao, J., Hays, J., Ehinger, K. A., Oliva, A. ve Torralba, A., 2010, Sun database: Large-scale scene recognition from abbey to zoo, *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 3485-3492.

Yamashita, R., Nishio, M., Do, R. K. G. ve Togashi, K. J. I. i. i., 2018, Convolutional neural networks: an overview and application in radiology, 9 (4), 611-629.

Yatziv, L. ve Sapiro, G. J. I. t. o. i. p., 2006, Fast image and video colorization using chrominance blending, 15 (5), 1120-1129.

Zhang, R., Isola, P. ve Efros, A. A., 2016, Colorful image colorization, *European conference on computer vision*, 649-666.

Zhang, R., Isola, P. ve Efros, A. A., , 2016, https://richzhang.github.io/colorization/ [Ziyarat Tarihi: 07 Mart 2020], .

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı	**:**	Omar Abdulwahhab Othman OTHMAN
Uyruğu	**:**	Irak
Doğum Yeri ve Tarihi	**:**	Irak, Kerkük. 05.05.1992
Telefon	**:**	05380609521
Faks	**:**	
e-mail	**:**	omer.salihi92@gmail.com

EĞİTİM

Derece		**Adı, İlçe, İl**	**Bitirme Yılı**
Lise	:	Merkezi Kerkük Lisesi,IRAK, KERKÜK	2011
Üniversite	:	Al-Qalem Ünversitesi ,IRAK, KERKÜK	2015
Yüksek Lisans	:	Konya Teknik Üniversitesi, Selçuklu, KONYA	
Doktora	:		

İŞ DENEYİMLERİ

Yıl	**Kurum**	**Görevi**

UZMANLIK ALANI

YABANCI DİLLER

Arapça, Türkçe, İngilizce

BELİRTMEK İSTEĞİNİZ DİĞER ÖZELLİKLER

YAYINLAR

Othman, O. , Uymaz, S. , Uzbaş, B.. (2019). Automatic Black & White Images colorization using Convolutional neural network. Academic Perspective Procedia, 2 (3), 1189-1195. DOI: 10.33793/acperpro.02.03.131