T.C.

BATMAN ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

DERİN ÖĞRENME İLE ANLAMSAL BÖLÜTLEME VE PİKSEL GÖRÜNTÜLERİNDEN GERÇEK GÖRÜNTÜ ÜRETİMİ

Emre ARICA

YÜKSEK LİSANS TEZİ

Elektrik-Elektronik Mühendisliği Anabilim Dalı

Haziran-2023
BATMAN

TEZ KABUL VE ONAYI

Emre ARICA tarafından hazırlanan "Derin öğrenme ile Anlamsal Bölütleme ve Piksel Görüntülerinden Gerçek Görüntü Üretimi" adlı tez çalışması 20/06/2023 tarihinde aşağıdaki jüri tarafından oy birliği / oy çokluğu ile Batman Üniversitesi Fen Bilimleri Enstitüsü Elektrik – Elektronik Mühendisliği Anabilim Dalı'nda YÜKSEK LİSANS olarak kabul edilmiştir.

Jüri Üyeleri	**İmza**
Başkan Doç. Dr. Melih KUNCAN	……..
Danışman Doç. Dr. Yılmaz KAYA	……..
Üye Dr. Öğr. Üyesi Davut SEVİM	……..

Yukarıdaki sonucu onaylarım.

Prof. Dr. Osman PAKMA
Lisansüstü Eğitim Enstitüsü Müdür V.

TEZ BİLDİRİMİ

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

DECLARATION PAGE

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

İmza

Emre ARICA

Tarih:

ÖZET

YÜKSEK LİSANS TEZİ

DERİN ÖĞRENME İLE ANLAMSAL BÖLÜTLEME VE PİKSEL GÖRÜNTÜLERİNDEN GERÇEK GÖRÜNTÜ ÜRETİMİ

Emre ARICA

Batman Üniversitesi Lisansüstü Eğitim Enstitüsü
Elektrik-Elektronik Mühendisliği Anabilim Dalı

Danışman: Doç.Dr. Yılmaz KAYA

2023, 76 Sayfa

Jüri
Doç. Dr. Melih KUNCAN
Doç.Dr. Yılmaz KAYA
Dr. Öğr. Üyesi Davut SEVİM

İki bölümden oluşan bu tez çalışmasının ilk bölümünde derin öğrenme metotları ile anlamsal bölütleme işlemi gerçekleştirilmiştir. Anlamsal bölütleme işlemi, bir görüntüdeki her pikselin ilgili bir etiket ile ilişkilendirme işlemidir. Anlamsal bölütleme ile görüntüdeki nesnelerin tespiti, yerinin belirlenmesi mümkün kılınmaktadır. Bilgisayar sistemleri tarafından görüntülerin daha iyi yorumlanması, anlaşılması için anlamsal bölütleme önemlidir. Son yıllarda derin öğrenme metotları ile görüntülerden nesne tespiti nesnelerin yorumlanmasında yaygın bir şekilde kullanılmaktadır. Mevcut araştırmada Resnet-18 transfer yöntemini temel alan Deeplab v3+ CNN ağı ile anlamsal bölütleme işlemi gerçekleştirilmiştir. Bunun için Camvid veri seti kullanılmıştır. 701 yüksek çözünürlüklü görüntüden oluşan veri setindeki görüntülere piksel bazlı anlamsal bölütleme manuel olarak uygulanmıştır. Öncelikli olarak bölütleme işlemi Gretag–Macbeth renk şeması esas alınarak gerçekleştirilmiştir. Ardından Deeplab v3+ gerçek görüntüler piksel görüntülerle eşleştirilerek eğitim işlemi gerçekleştirilmiştir. Modeli test etmek için farklı görüntüler kullanılmıştır. Gözlenen Jaccard, Sørensen-Dice ve BF Skoru metriklerine göre yüksek başarılar gözlenmiştir. Tezin ikinci aşamasında derin öğrenme metotları ile piksel görüntülerden sentetik görüntüler oluşturulmuştur. Bu kapsamda derin öğrenme metotlarından GAN yöntemlerinden faydalanılmıştır. GAN modeller farklı alanlarda sentetik veriler üretmek için yaygın bir şekilde tercih edilmektedir. Araştırmada gerçek görüntüler oluşturmak için Pix2PixHD GAN modeli kullanılmıştır. Pix2PixHD, yüksek çözünürlüklü görüntülerin düşük çözünürlüklü eşlemelerinden gerçekçi ve ayrıntılı görüntüler üretmek için kullanılan bir görüntü çeviri yöntemidir. Bu yöntemin temelinde, derin öğrenme ve özellikle de evrişimli sinir ağları vardır. Pix2PixHD GAN yönteminde CNN ağı olarak VGG19 transfer derin öğrenme metodu kullanılmıştır. Denemeler Camvid veri seti üzerinde gerçekleştirilmiştir. Gerçekleştirilen denemelerde başarılı yüksek çözünürlüklü görüntülerin üretildiği sonucuna varılmıştır.

Anahtar Kelimeler: Anlamsal bölütleme, Derin öğrenme, GAN, Pix2PixHD

ABSTRACT

MS THESIS

SEMANTIC SEGMENTATION WITH DEEP LEARNING AND REAL IMAGE GENERATION FROM PIXEL IMAGES

Emre ARICA

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCE OF BATMAN UNIVERSITY
THE DEGREE OF MASTER OF SCIENCE / DOCTOR OF PHILOSOPHY IN ELECTRICAL- ELECTRONICS ENGINEERING

Advisor: Doç.Dr. Yılmaz KAYA

2023, 76 Pages

Jury
Assoc. Prof. Melih KUNCAN
Assoc. Prof. Yılmaz KAYA
Assist Prof. Dr. Davut SEVİM

In the first part of this thesis, which consists of two parts, semantic segmentation is carried out with deep learning methods. Semantic segmentation is the process of associating each pixel in an image with a corresponding label. Semantic segmentation can be used to detect and locate objects in the image. Semantic segmentation has become an important issue for better interpretation and understanding of images by computer systems. In recent years, object detection from images with deep learning methods has been widely used in the interpretation of objects. Semantic segmentation was performed with the Deeplab v3+ CNN network based on the Resnet-18 transfer method. For this, the Camvid dataset was used. Pixel-based semantic segmentation was applied manually to the images in the dataset consisting of 701 high-resolution images. Primarily, the segmentation process was performed according to the Gretag–Macbeth color scheme. Then, Deeplab v3+ real images were matched with pixel images and the training process was carried out. Different images were used to test the model. High successes were observed according to the observed Jaccard, Sørensen-Dice and BF Score metrics.

In the second stage of the thesis, synthetic images were created from pixel images with deep learning methods. For this, GAN methods, one of the deep learning methods, were used. GAN models are widely preferred to generate synthetic data in different fields. In our thesis study, Px2PxHD GAN model was used to create real images. Pix2PixHD is an image translation method used to produce realistic and detailed images from low-resolution maps of high-resolution images. The basis of this method is deep learning and especially convolutional neural networks. In Pix2PixHD GAN method, VGG19 transfer deep learning method was used as CNN network. Experiments were carried out on the Camvid dataset. It has been observed that successful high-resolution images were produced in the experiments carried out.

Keywords: Deep learning, GAN, Pix2PixHD, Semantic segmentation

ÖNSÖZ

Yüksek lisans eğitimimi tamamlarken, çeşitli zorluklarla karşılaştım ve bu süreçte büyük bir öğrenme deneyimi yaşadım. Bu tez çalışması benim için hem akademik anlamda hem de ilgi duyduğum alanda daha derinlemesine bir araştırma yapma fırsatıydı. Bu tez çalışmasını tamamlarken, birçok kişiye minnettarım ve onlara teşekkür etmek istiyorum. İlk olarak, çok kıymetli danışmanım Doç.Dr. Yılmaz KAYA ’ya en içten duygularımla teşekkürlerimi sunarım. Yüksek lisans ders aşamasında dersini aldığım Batman Üniversitesi’ndeki tüm hocalarıma ilgi ve alakaları için teşekkür ederim. Ayrıca, eğitim hayatım boyunca yanımda olan tüm zorluklar ile mücadelemde desteğini esirgemeyen Doç.Dr. Reşat ARICA ve aileme de teşekkür etmek istiyorum.

Emre ARICA
BATMAN-2023

İÇİNDEKİLER

ÇİZELGELER LİSTESİ

ŞEKİLLER LİSTESİ

SİMGELER VE KISALTMALAR

Kısaltmalar

CNN	: Convolutional Neural Networks
GAN	: Generative Adversarial Networks
CGAN	: Conditional Generative Adversarial Network
UDA	: Unsupervised Domain Adaptation
CT	: Computed Tomography
ReLU	: Rectified Linear Unit
FC	: Fully Connected
OCR	: Optical Character Recognition
IoU	: Intersection Over Union

1. GİRİŞ

Teknolojinin hızlı gelişimi ile farklı buluş, araç, aygıt, sistem, yöntem ve teknikler insan hayatında yer edinmekte, gündelik yaşamın bütünleyicisi olarak kabul görmektedir. Birkaç yıl öncesinde sadece bilim kurgu filmlerinde görülen teknolojik araç, aygıt ve sistemler günümüzde insan hayatında yadsınamaz bir öneme sahip hale gelmiştir. Taşınamaz teknolojilerle başlayan gelişim süreci, taşınabilir teknolojiler, internet, sosyal medya, giyilebilir teknolojiler, gerçeklik teknolojileri, yapay zekâ, blokchain, nano teknoloji, nesnelerin interneti ve metaverse ile günümüz insanının hayatını dönüştüren bir yapıya erişmiştir. Hızlı gelişimi ve insan hayatına etkisi itibariyle teknolojik araç ve sistemler farklı disiplinlerde birçok araştırmaya konu olmaktadır. Teknoloji eksenli gelişim araştırmalarda farklı yönleri ile değerlendirilirken, günümüzün popüler teknoloji araç ve sistemleri arasında yer alan yapay zekâ ve derin öğrenme yöntemleri halihazırdaki teknolojik değişimin anlaşılması ve geleceğe ilişkin perspektif geliştirilmesi noktasında önemli bir araştırma alanı olarak kabul görmektedir.

Yapay zekâ, gündelik yaşantımızda karşılaştığımız problemlerin birçoğuna etkili çözümler üretmektedir. Oluşturduğu verileri ses, görüntü, metin gibi farklı türlerde anlamlandırıp işleyen yapay zekâ uygulamada başarılı sonuçlar vermektedir. Yapay zekâ gelişim sürecinin başlangıcında oluşan problemleri matematiksel işlemler ve matematiksel formüllerle çözülebilme işlevine sahipken, gelişen teknoloji sayesinde yapay zekâ problemlerin çözümünü birbiri üzerine kurulu çok sayıda yapay sinir ağı katmanlarından oluşan ve bu ağların eğitilmesi ile elde edilen deneyim sayesinde gerçekleştirmektedir. Söz konusu yöntemler derin öğrenme yöntemi olarak ta ifade edilmektedir.

Makine öğrenmesinin bir türü olarak açıklanan derin öğrenme, dünyadaki kavramlar hiyerarşisi ve deneyimler ile elde edilmektedir (Sun vd., 2019). Derin öğrenme modelinin eğitim işlemi kısaca, çevresiyle etkileşime giren ve bu etkileşimlerin sonuçlarını gözlemleyerek öğrenilmektedir. Farklı şekillerde temellendirilebilen derin öğrenme yöntemi bir takım avantaj ve dezavantajları barındırmaktadır. Derin öğrenme tekniklerinin çeşitli üst düzey bilgisayarlı görme görevlerindeki ciddi başarısı özellikle görüntü sınıflandırma veya nesne algılama için evrişimli sinir ağları gibi denetimli yaklaşımlar, derin öğrenmenin olumlu çıktılarını geliştirmektedir (Oprea vd., 2017). Derin öğrenmenin başlıca dezavantajı ise yöntemin öğrenilmesinde gereken eğitimin barındırdığı büyük miktarda görüntü ve karmaşık işlemlerdir. Günümüzde oldukça

yaygın olan grafik işleme birimleri, büyük miktardaki verileri bilgisayar işlemcisine oranla çok daha hızlı bir şekilde işleyebilmektedir. Bunun sonucunda hem anlık olarak elde edilen hem de depolama biriminde bulunan verilerin hızlı bir şekilde işlenerek sonuca daha kısa bir zamanda ulaşılması sağlanmaktadır. Derin öğrenmedeki son gelişmeler, özellikle derin evrişimli sinir ağları, önceki anlamsal segmentasyon sistemlerine göre önemli gelişmeleri beraberinde getirmiştir (Asano vd., 2009).

Derin öğrenme yöntemlerinin eğitim işlemi için büyük miktarda veri kümeleri gerekmektedir. Ancak bu konuda birtakım problemlerin olduğu da gerçektir. Bir nesne sınıflandırma probleminde ihtiyaç duyulan veri miktarının toplanması çoğu zaman mümkün değildir. Bu noktada zaman, para ve donanım en belirgin kısıtlayıcılar olmaktadır. Örneğin; birçok biyomedikal görüntünün cep telefonunun kamerası ile toplanması mümkün değildir. Görüntüler toplansa bile etiketlenmesi sürecinde sıradan bir göz yeterli olmayacağından, ilgili görüntü konusunda tecrübe sahibi uzmanlara ihtiyaç duyulmaktadır (Albawi vd., 2017). Bir diğer önemli nokta evrişimli sinir ağlarına verilen görüntülerden elde edilen sınıf etiketleri ile eğitimin gerçekleştirilmesidir. Problemler piksel temelli yaklaşımlarla konumlama bilgisi gerektirmesi, biyomedikal veya savunma gibi önemli yaklaşımlar gerektiren alanlarda piksellerin her birinin sınıf bilgisine ihtiyaç duyulması süreci güç kılmaktadır. Bölütleme güç olmakla beraber uygulamada farklı modellerle karşımıza çıkabilmektedir. Bölütlemenin farklılaşan modellerinden biri Anlamsal Bölütlemedir.

Gelişmekte olan uygulamalar, doğru ve verimli bölütleme mekanizmalarına ihtiyacı arttırmaktadır. Bu ihtiyaç, sahne anlama ve bilgisayarla görme alanlarında uygulamaya yönelik derin öğrenme yaklaşımları ile örtüşmektedir. Bu bağlamda anlamsal bölütleme gibi, bilgisayarla görme ve makine öğrenimi konuları araştırmacılar tarafından daha fazla ilgi görmektedir (Özdemir, 2022).

Anlamsal bölütleme, görüntü işlemede ve bilgisayarlı görü alanında önemli bir araştırma konusudur. Görüntülerdeki nesnelerin ve alanların anlamlı sınıflandırmasını gerçekleştiren anlamsal bölütleme, aynı zamanda görüntüdeki her piksel için bir sınıflandırma etiketi oluşturmayı amaçlar. Örneğin; görüntü içerisindeki araba, insan, bina gibi nesneler daha önceden belirlenmiş piksel gruplarına göre sınıflandırılabilir. Bu sınıflandırma, son derece hassas ve doğru olmalı, çünkü sınıflandırma sonuçları birçok uygulama alanlarında kullanılabilmektedir. Otonom sürüş, haritalama ve görüntü analizi gibi alanlar sınıflandırma sonuçlarının uygulandığı alanlardan bazılarıdır (Hongshan vd., 2018).

Görüntü, birbirinden farklı birçok renkten piksellerin bütünleştiği bir yapıdır ve bilgisayarların görüntüyü işleyebilmesi için, anlamsal bölütleme işleminden geçirilmesi gerekmektedir. Anlamsal bölütleme işlemi, temel olarak iki adımda gerçekleşir. İlki birey bölütlemedir. Birey bölütleme, bir nesneye ait her bir bireyin farklı bir sınıf etiketine atanması işlemidir. Örneğin; bir görüntüde yer alan üç köpeğe üç farklı etiket atanır; ikinci olarak anlamsal bölütlemede, belirli bir nesneye ait ve bireylere ait pikseller aynı sınıf etiketine atanır. Bir görüntüde yer alan üç köpeğe aynı sınıf etiketi atanması bu durumu örneklemektedir.

Anlamsal bölütleme gibi piksel düzeyinde etiketleme sorunları için CNN ağların yeteneklerinin kullanımı onlara geleneksel yöntemlere göre bir avantaj sağlar. Örneğin piksel etiketleme belirli bir veri kümesinde, uçtan uca öğrenme yeteneği sağlar. Modelin belirli bir senaryo üzerinde çalışmasını mümkün kılmak üzere alan uzmanlığı, çaba ve genellikle çok fazla ince ayar gerektiren el yapımı özellikleri kullanılmaktadır. Modern derin öğrenme mimarileri tarafından anlamsal bölütlemenin nasıl ele alındığını anlamlandırmak için, bu alanın izole bir alan olmadığını, kaba çıkarımdan ince çıkarsamanın ilerlemede ciddi bir adım olduğunu bilmek önemlidir. Anlamsal bölütlemenin ince çıkarımında amaç; her piksel için etiketler oluşturan tahminler yapmak ve bu sayede her pikselin içinde bulunduğu nesne veya bölgenin sınıfı ile etiketlenmesidir.

Anlamsal bölütleme, görüntünün aynı nesne sınıflarına ait olan parçalarını bütünleştirme işlemini de gerçekleştirmektedir. Örneğin, görüntüdeki dağ, ağaç, göl gibi nesneleri ve alan yoğunluğunu belirlemek için anlamsal bölütleme yöntemi kullanılmaktadır. Günümüzde anlamsal bölütleme, hareketsiz 2B görüntülere, 3B veya hacimsel verilere de uygulanabilmektedir. Bu yöntem ile görüntü içerisinde daha önceden etiketlenmiş nesneler aranabilmektedir. Bölütleme işleminin başarılı olabilmesi için, bölütlemenin şartlarının doğru bir şekilde sağlaması önemlidir. Evrişimli sinir ağı yaklaşımı olan U-Net ilk olarak biyomedikal görüntüler üzerinde daha iyi bir bölütleme yapma önerisi ile 2015 yılında, Olaf Ronneberger, Phillip Fischer ve Thomas Brox tarafından U-Net: Convolutional Networks for Biomedical Image Segmentation makalesinde ele alınmıştır. Son yıllarda, anlamsal bölütleme alanında çok sayıda önemli gelişme gerçekleşmiştir. Buna örnek olarak birçok araştırmacı, görüntülerdeki nesnelerin ve alanların belirlenmesini iyileştirmek için çoklu modaliteler (örneğin; sıcaklık, radar, LiDAR gibi) kullanmaktadır (Han vd., 2019).

Anlamsal bölütleme, derin öğrenme yöntemleri kullanarak gerçekleştirilmektedir. Bu yöntemler, evrişimli sinir ağları (CNN) veya tam bağlı sinir ağları (FCN) gibi yapay sinir ağlarını kullanmaktadır. Bu yapay sinir ağları, görüntülerdeki nesnelerin veya alanların ne olduğunu tanımlamak için eğitilmektedir. Görüntülere anlamsal bölütleme işlemi uygulamak için kullanılabilecek olan metotlar aşağıda verilmiştir:

Fully Convolutional Networks (FCN): Görüntüleri evrişimli sinir ağları kullanarak etiketlemeyi sağlar. Bu sayede, görüntülerdeki nesneleri veya alanları tanımlamak için eğitilmiş olur ve her piksele bir anlam vererek sınıflandırma işlemini yapar.

Region-based CNNs: Görüntülerdeki nesnelerin tanınmasında ve nesnelerin sınıflarının çizilmesi için region-based CNNs kullanır.

SegNet: Daha yüksek çözünürlüklü etiketlemeler elde etmenizi sağlayarak daha az parametre ve kaynak gerektirir. Bu nedenle, SegNet 'in sınırlı kaynak gerektiren sistemlerde kullanılması daha uygundur.

FCN, Region-based CNNs ve SegNet görüntülere anlamsal bölütleme işlemi uygulamak için kullanılabilen genel metotlar olmasına karşın, söz konusu metotlar kullanılan ağ yapıları, eğitim veri setleri ve performans bakımından farklılık göstermektedir.

Özetle; eğitim aşamasında yapay sinir ağına bir görüntü veri seti ve bunların bağlı olduğu etiketleri verilmektedir. Bu etiketler, görüntülerdeki nesneleri tanımlamamızı sağlamaktadır. Ağ, eğitim verilerinden öğrenerek, görüntülerdeki nesneleri ve nesnelerin sınır çizgilerini tanımlamak için kullanılabilir hale gelmektedir. Ağın test aşamasında ise tanımlanmamış bir görüntü verilerek görüntüdeki nesnelerin sınıflandırılması ve algılanması sağlanmaktadır.

Nesnelerin algılanması ve sınıflandırılmasına olanak tanıyan anlamsal bölütleme işlemi, farklı uygulama alanlarında kullanılabilmektedir. Otomotiv endüstrisi bu alanlardan biridir. Otonom araçların çevrelerini anlamaları için kullanılabilen sistem, araçların etrafındaki nesneleri doğru şekilde tanımlamalarını sağlamakta ve güvenli sürüşü mümkün kılmaktadır. Tıbbi görüntüleme alanında ise MR ve CT görüntüleri üzerinde çalışarak, hastalıkları tespit etmek veya dokuları analiz etmek için kullanılabilen sistem, doğal dil işleme, bir cümleyi anlama ve cümledeki her kelimeyi anlamlı bir şekilde sınıflandırma amacıyla kullanılabilmektedir. Ayrıca sistemden çevrimiçi içerik filtreleme, web sitelerindeki içeriği sınıflandırmak ve istenmeyen içeriği engelleme amacı

ile istifade edilebilmektedir. Farklı alanlarda kullanımı anlamsal bölütleme sisteminin birtakım avantajları ve dezavantajları beraberinde getirmesine yol açmaktadır.

Anlamsal bölütlemenin, görüntü işleme ve makine öğrenmesi alanlarında birçok avantaj sağladığı ifade edilmektedir. Bazı avantajları şu şekildedir:

- Nesne tespiti ve sınıflandırma için kullanılabilmektedir. Bu, bir görüntüdeki nesnelerin tespit edilmesi ve sınıflandırılması için daha doğru sonuçlar elde etmemizi sağlamaktadır.
- Anlamsal bölütleme, otomasyon için kullanılabilmektedir. Örneğin, bir otonom aracın çevresindeki nesneleri tanıması ve sınıflandırması için anlamsal bölütlemeden faydalanılmaktadır. Bu, aracın doğru şekilde hareket etmesini sağlayarak trafik güvenliğini artırmaktadır.
- Tıbbi görüntüleme ve bilgisayarlı görü alanında da kullanılabilmektedir. Örneğin, kanser tespiti gibi önemli tıbbi teşhislerde bu sistemden faydalanılmakta, daha doğru bir şekilde teşhis konulmasını olanaklı kılmaktadır.
- Anlamsal bölütleme, derin öğrenme yöntemleri kullanılarak gerçekleştirildiği için, yüksek doğruluk ve veri güvenliği sağlamaktadır. Bu sayede daha doğru sonuçlar elde edilmesi ve daha iyi kararlar alınması mümkün hal almaktadır.
- Tarım sektörü anlamsal bölütleme tekniklerinin kullanıldığı alanlardandır. Bu teknikler tarım sektöründe bitkilerin ve toprakların belirlenmesine yardımcı olurken, çiftçilerin bitki hastalıklarını, zararlıları ve diğer sorunları daha kolay tespit etmelerini sağlamaktadır.
- Anlamsal bölütlemenin yoğun kullanıldığı alanlardan biri de sosyal medyadır. Sosyal medya analizi gerçekleştirmede faydalanılan anlamsal bölütleme teknikleri, sosyal medya kullanıcılarının davranışlarını ve tercihlerini anlamaya yardımcı olurken, markaların daha etkili pazarlama stratejileri geliştirmelerine olanak tanımaktadır.

Anlamsal bölütleme tekniklerinin sahip olduğu avantajlarla beraber, birçok bazı dezavantajları da barındırdığı gerçektir. Bu dezavantajlar aşağıdaki gibi ifade edilmektedir:

- Anlamsal bölütleme işlemi oldukça yüksek hesaplama gücü gerektirmektedir. Bu nedenle anlamsal bölütleme tekniği ile büyük veri kümeleri üzerinde çalışırken, işlem süresi artabilmekte ve sistemlerin çökme riski ortaya çıkabilmektedir.
- Anlamsal bölütleme işlemi sırasında, bazı nesneler ve arka plan yanlış sınıflandırılabilmekte veya hatalı bir şekilde bölünebilmekte. Dolayısı ile anlamsal bölütlemenin hata oranlarının arttırabilirken, sonuçların doğruluğunun azaltılmasına sebebiyet verebilmektedir.
- Anlamsal bölütleme algoritmaları, eğitim verilerinin kalitesine bağlıdır. Eğitim verilerinin yeterli sayıda olmaması (eksik olması) veya kalitesinin düşük olması, algoritmaların doğruluğunu etkileyebilmektedir.
- Anlamsal bölütleme işlemi, farklı kontrast seviyelerine sahip görüntülerle çalışırken daha zor hale gelebilmektedir. Kontrast problemine sahip görüntülerde nesnelerin ve arka planın sınıflandırılması daha zor hal alabilmektedir.

Anlamsal bölütleme tekniklerinin yol açabileceği dezavantajlar, anlamsal bölütleme işlemi gerçekleştirilmeden önce, çalışmada kullanılacak algoritmaların seçimi ve eğitim verilerinin kalitesinin iyileştirilmesi gibi önlemler alınmasını elzem kılmaktadır.

Bu çıkış noktasından hareketle hazırlanan çalışmanın amacı, derin öğrenme yöntemlerini kullanarak görüntüler üzerinde anlamsal bölütleme gerçekleştirmektir. Bununla beraber piksel bazlı görüntülerin gerçek görüntülere dönüştürülme işleminin gerçekleştirilmesi araştırmanın bir diğer amacıdır. Bu işlem, orijinal veri alanının görüntü özelliklerini öğrenerek orijinal verilere benzer yeni görüntüler oluşturma işlemi olarak özetlenebilmektedir (Qin vd., 2022).

Piksel bazlı görüntüler, genellikle düşük çözünürlüklü ve sınırlı renk paletine sahip olabilmektedir. Ancak, gerçek görüntüler daha yüksek çözünürlüklü ve genellikle daha fazla renk seçeneği sunmaktadır. Bu nedenle, piksel bazlı görüntüler gerçek görüntülere dönüştürülerek daha gerçekçi ve ayrıntılı görüntüler elde edilebilmektedir. Piksel bazlı görüntülerin gerçek görüntülere dönüştürülme işlemi, birçok alanda kullanılmaktadır (Güzel ve Bilge, 2020). Sentetik görüntüler üretmek için ise GAN

modellerine ihtiyaç vardır. Örneğin, moda tasarımında yeni kıyafetlerin tasarlanması, mimaride yeni binaların tasarlanması, sanat eserleri üretimi gibi. Ayrıca, oyun geliştiricileri de GAN modellerini kullanarak oyunlarda yeni nesneler üretebilmektedir. Bununla birlikte, GAN modelleri ile üretilen nesnelerin gerçekliği konusunda bazı sınırlamalar ve zorluklarda mevcuttur. Bu bağlamada üretilen nesneler, gerçek nesnelerle tam olarak örtüşmeyebilmekte veya gerçek nesnelerin görünüşüne benzerlik göstermeyebilmektedir. Bu odakta mevcut çalışmada:

- Görüntüdeki detayları artırarak daha doğru ve ayrıntılı bir görüntü elde etmek,
- Görüntü üzerindeki nesneleri çoğaltmak ve farklı görüntüler elde etmek,
- Görüntülerin daha yüksek kaliteli olmasını sağlayarak, kullanıcıların daha iyi bir deneyim yaşamasını sağlamak,
- Görüntülerin tanıma, analiz ve teşhis için daha kullanışlı hale getirilmesi hedeflenmektedir.

Mevcut araştırmaya dayanak oluşturması bağlamında piksel bazlı görüntülerin gerçek görüntülere dönüştürülme işlemi, birçok alanda kullanıldığı görülmektedir. Bunlara örnek verecek olursak:

Medikal görüntüleme: Manyetik rezonans görüntüleme (MRI) veya bilgisayarlı tomografi (CT) taramaları gibi medikal görüntüleme teknikleri, genellikle düşük çözünürlüklü piksel bazlı görüntüler üretmektedir. Bu nedenle, söz konusu görüntüler gerçek görüntülere dönüştürülerek, teşhis ve tedavi için daha fazla bilgi sağlanabilmektedir.

Oyun geliştirme: Piksel bazlı görüntüler, retro tarzı oyunlar veya 8 bit grafikler için kullanılabilmektedir. Ancak, modern oyunlar için gerçekçi görüntüler gerektiğinden, piksel bazlı görüntüler gerçek görüntülere dönüştürülerek daha yüksek kaliteli oyun grafikleri elde edilebilmektedir.

Dönüştürme işleminde kullanılan, koşullu üretken çekişmeli ağlar (Conditional Generative adversarial networks- cGAN) bir tür üretken çekişmeli ağ (GAN) olarak nitelendirilebilmektedir (Li vd., 2021). Üretken çekişmeli ağ, gerçekçi görüntüler, sesler ve diğer verileri üretmek için bir üretici (Generator) ve bir ayrım (Discriminator) ağından oluşmaktadır. Bu ağlar, gerçek ve üretilmiş veriler arasındaki farkı en aza indirmek için birbirleriyle yarışmaktadır. CGAN ‘lar ise, rastgele gürültüye ek olarak, özel bir koşulun (Condition) da generator tarafından kullanılmasına izin vermektedir. Bu koşul, generatorun ürettiği verilerin, koşulun belirttiği özelliklere sahip olmasını sağlamaktadır.

Örneğin, bir cGAN, belli bir renkte bir elbise üretmek için renk bilgisini koşul olarak kullanabilmektedir. Piksel bazlı görüntülerin gerçek görüntülere dönüştürülme işlemi için, koşullu üretken çekişmeli ağlardan pix2pixHD yöntemi kullanılmıştır.

Pix2pixHD yöntemi, üretken çekişmeli ağ yapısına dayanan bir görüntü çevirme modelidir. Bu yöntem, giriş olarak verilen bir görüntüyü, belirli bir çıktı görüntüsüne dönüştürmek için öğrenilmiş bir modeldir (Salehi ve Chalechale, 2015). Pix2pix, birbirine bağlı iki bölümden oluşur: bir generator ve bir discriminator. Generator, giriş görüntüsünden çıktı görüntüsü oluşturmak için bir işlevi öğrenirken, discriminator, oluşturulan çıktı görüntüsünün gerçekçi olup olmadığını belirlemektedir. Bu iki bölüm, birbirleriyle karşılaştırmalı bir şekilde çalışır ve jeneratör, discriminatorun belirlediği gerçekçilik kriterlerini karşılamaya çalışmaktadır.

Pix2pixHD yöntemi, birçok alanda kullanılabilmektedir. Örneğin, birçok doğrusal olmayan görüntü işleme uygulaması için ve özellikle de nesne tanıma, stil transferi, fotoğraf restorasyonu ve renklemesi gibi uygulamalarda oldukça etkilidir (Jiao vd., 2021). CGAN ‘lar ise, görüntü sentezi, stili transferi, yüz tanıma ve diğer birçok alanda kullanılmaktadır. Özellikle görüntü sentezi alanında oldukça başarılı sonuçlar vermektedir.

Mevcut tez çalışmasında anlamsal olarak bölütlenmiş haritalardan derin öğrenme ile yüksek çözünürlüklü sentetik görüntüler oluşturulmuş, bu sayede sentetik görüntülerin gerçek görüntülere dönüştürülmesi işlemini başarılı bir şekilde gerçekleştirilmiştir.

2. KAYNAK ARAŞTIRMASI

Kaynak araştırmasında mevcut araştırmaya ait savların doğrulanmasına dayanak oluşturmak için, konuya ilişkin uluslararası ve ulusal alan yazında yer alan anlamsal bölütleme, derin öğrenme ile anlamsal bölütleme ve piksel görüntülerinden gerçek görüntü üretimi ile ilgili yapılmış çalışmalar incelenmektedir.

Sun ve arkadaşları (2023), zayıf denetimli anlamsal bölümleme için önyargısız kendi kendine öğrenen bir metot üzerine çalışmışlardır. Anlamsal bölütleme etiketi tohumunu oluşturmak için temel olarak sınıf aktivasyon haritası bölgelerini üretilmiştir. Ancak çekirdek fonksiyonun seyrekliğinin, doğruluk oranı az ve yerel ayrımcı bölgelere yol açtığı gözlemlenmiştir. Bu nedenle çalışmada önyargısız kendi kendine öğrenme bölütleme ağı (NBSA) önerilmiştir. Bu odakta ilk olarak, önyargısız katmanlar tasarlanmış ve ağın eğitim sürecinde sınıf aktivasyon haritasının ayırt edici alanını genişletmek için ağa rehberlik etmesi sağlanmıştır. Akabinde görüntü üzerindeki anlamlar arası önyargısız için çizgi evrişimli ağ oluşturularak öğrenme metodu, sınırlı görüntü seviyesi etiketleme örneklerinden tam olarak yararlanmak amacı ile kapsamlı anlamsal bilgileri yakalamak için oluşturulmuştur. Sonuç olarak, oluşturulan yöntemin farklı sınıflarda daha yüksek kalitede çekirdek fonksiyonun üretebilmesi ve böylece anlamsal bölümleme performanslarını etkin bir şekilde geliştirebilmesi mümkün hal almıştır. Yöntemlerinin üstünlüğünü daha iyi göstermek için PASCAL VOC 2012 doğrulama setinde her sınıf için değerlendirme yapılmış ve başarılı sonuçlar elde edilmiştir (Sun vd., 2023).

Piksel bulutlarının ölçek ve yön duyarlı anlamsal bölütlenmesi odağında Wang ve arkadaşları (2023) tarafından gerçekleştirilen çalışmada, piksel bulutlarındaki seyreklik, düzensizlik özelliklerine uyduğu ve 3B anlamsal bölümleme görevi sırasında gürültülü veya sağlam olmayan özelliklere yol açtığını gözlemlenmiştir. Mevcut yaklaşım, irrasyonel özellik öğrenme veya komşuluk seçim şemaları nedeniyle piksel bulutlarının yerel geometrisini ve bağlam bilgisini tam olarak çıkaramamaktadır. Wang vd. (2023) gerçekleştirdikleri çalışmada PASIFTNet ’i önermektedir. PASIFTNet 'in temel bileşeni, PASIFT modülünün Nokta-Atrous oryantasyon kodlama birimidir, alıcı alanlarını genişleterek daha geniş bağlam bilgisini ekleyebilmektedir. Ölçek ve yöne duyarlı özellik noktası bilgisini ayıklayabilmekte, dörtgensel SIFT benzeri noktasal evrişimden yararlanılmaktadır. Araştırmacıların oluşturdukları veri seti ScanNet lazer tarayıcı tarafından oluşturulmuştur. Sonuç olarak, PASIFTNet'in genel olarak %86,8 doğruluğa ulaştığı belirlenmiştir (Wang vd., 2023).

Anlamsal bölütleme ile ilgili Aboobacker ve arkadaşları (2023) tarafından hazırlanan çalışmada, düşük büyütmeli efüzyon sitoloji görüntülerinin anlamsal bölütlenmesi yarı denetimli bir yaklaşım değerlendirilmiştir. Bu odakta etiketlenmemiş düşük büyütmeli efüzyon sitoloji görüntülerinin anlamsal bölümlemesinin yapılması hedeflenmiştir. Araştırmada kötü huylu hücre kümelerini (tümor) düşük bir büyütme oranında bulunmuş ve ardından yüksek bir büyütme oranı için hücre düzeyindeki özellikleri araştırmak için yakınlaştırmak istenmiştir. Veri seti, 1920 X 1440 piksel çözünürlükte 345 efüzyon sitoloji görüntüsünden oluşan 30 hastadan toplanmıştır. Çalışmanın en büyük faydası hücre tarama süresinin azaltmaktır. Bu çalışmada, anlamsal bölümleme için iki yarı denetimli öğrenme modelini, MixMatch ve FixMatch ile görüntüler büyütülmüştür. Modeller, etiketli 10X ve etiketsiz 4X görüntüler kullanılarak eğitilmiştir. Önerilen modeller, MobileUNet ve ResUNet++ adlı iki standart temel model üzerine inşa edilmiştir. Amaç 4x görüntülerdeki iyi huylu ve kötü huylu piksellerin açıklama eklenip tarama listesinden çıkarılmasıdır. Kullanılan modeller sayesinde yaklaşık %9 iyileşme görülmüş ve genişletilmiş MixMatch'te, düşük boyutta büyütmeli görüntülerin %62'lik alt bölgeleri daha yüksek boyutta büyütülmesinde taramadan çıkarılması sağlanmıştır (Aboobacker vd., 2023).

Zhang ve arkadaşları (2023) anlamsal bölümleme ile Piramit Geometrik Tutarlılık Öğrenme için ön gördükleri problem, anlamsal bölütleme uygulanacak görüntüler aynı veri seti ve ağdan olsa bile tahmin elde edilen sonuçları birbirinden farklı olacak olmasıdır. Bu nedenle, örnekler arasında dönüşüm-eşdeğerlik ve temsil tutarlılığının da denetlenmesi gerektiğini düşünmektedir. Bu problemin çözümü için basit bir çapraz veri arttırma önermektedir. Modeli eğitmek için Deeplab V3+ derin öğrenme ağından yararlanılırken, veri seti olarak CamVid kullanılmıştır. Araştırma sonucunda göre, oluşturulan modelin ek hesaplama yükü olmadan mevcut anlamsal bölümleme modellerinin performanslarını önemli ölçüde arttırdığı görülmüştür (Zhang vd., 2023).

Zhu ve Tian (2023), stil transferine dayalı alan uyarlama yönteminde farklı kategori türlerindeki alan boşlukları üzerinde çeşitli etkilerin sorununu çözmek için şekil sağlamlığı geliştirilmiş bir alan uyarlamalı bölütleme algoritması önerilmiştir. Bu çalışmada var olan probleme geliştiricilerin katkıları şunlardır: Şekil sağlamlığını iyileştirmek için stil transferine dayalı bir çapraz alan anlamsal bölütleme yöntemi önermişlerdir. Bu yöntem stil aktarımına dayalı yöntemin performansını korurken, yapılandırılmış bilgilerdeki eksikliklerini azaltmaktadır. Ayarlanabilir stil aktarımı ile kaynak etki alanına stil çeşitliliği enjekte ederek ve şekle duyarlı kategorilerin tutarlılığını

kısıtlamak için farklılaştırılmış çapraz eğitim yöntemini birleştirmektir. Araştırmada kullanılan veri seti, Kaggle's painter by numbers içerisinde 44 stil kategorisini kapsayan 1584 farklı yazardan 79.433 resim görüntüsü içermektedir. İlk olarak, kaynak alan görüntülerinin stil çeşitliliğini artırmak için ayarlanabilir stil aktarım yöntemleri gerçekleştirilmiştir. Ardından GTA5 ve SYNTHIA veri setleri üzerinde başarılı sonuçlar elde edilmiştir. Araştırmanın GTA5 ve SYNTHIA genel veri setleri üzerindeki sonuçları diğer stil aktarım yöntemlerinden daha iyidir. Diğer deneyler, alan uyarlamalı bölütleme görevinde stil geliştirme yönteminin şekil sağlamlığını geliştirdiklerini göstermektedir (Zhu ve Tian, 2023).

Chieh ve arkadaşları (2018), derin sinir ağlarında anlamsal bölütleme görevi için uzaysal piramit havuzlama modülü veya kodlayıcı-kod çözücü yapısı kullanmıştır. İlk aşamada, gelen özellikleri filtreler veya havuzlama işlemleri ile çıkararak çok ölçekli bağlamsal bilgileri kodlayabilirken, ikinci aşamada uzamsal bilgiyi kademeli olarak kurtararak daha keskin nesne sınırlarını yakalayabilmişlerdir. Modelin etkinliği PASCAL VOC 2012 anlamsal görüntü bölütleme veri seti üzerinde gösterilmiş ve test veri setinde herhangi bir son işlem olmaksızın %89'luk bir performans elde edilmiştir (Chieh *vd.*, 2018).

Brostow ve arkadaşları (2017), Cambridge-sürüş etiketli video veri tabanını (CamVid), meta verilerle birlikte nesne sınıfı anlamsal etiketlere sahip ilk video koleksiyonu olarak sunmaktadır. CamVid veri tabanının potansiyel faydalarını değerlendirmek için araştırmacılar mevcut birkaç algoritmanın performansını ölçmüştür. CamVid veri tabanı, Orijinal yüksek çözünürlüklü (HD) video çekimlerinden ve gönüllülerin 32 nesne sınıfı listesine göre elle etiketlenmiş 10 dakikalık karelerden oluşmaktadır. Çerçevelerdeki nesne etiketlemesinin piksel hassasiyeti, algoritmaların doğru bir şekilde eğitilmesine ve nicel olarak değerlendirilmesine olanak tanımaktadır. Tek bir uygulama için toplanan birçok veri tabanından farklı olarak, CamVid veri tabanı, birden çok alanda kullanılmak üzere tasarlanmıştır. Üç performans deneyinde, veri tabanının nicel algoritma testi için kullanışlılığını incelenmektedir. Algoritmalar sırayla, nesne tanıma, yaya algılama ve videoda bölümlere ayrılmış etiket yayılımını ele almıştır. Sonuç olarak, CamVid veri tabanı, nesne analizi araştırmacıları ile ilgili dört katkı sunar. İlk katkısı, piksel başına 1 veya 15 Hz'de en az 700 görüntü için sınıf etiketleri, videoda çok sınıflı nesne tanıma için mevcut ilk temel gerçeği sağlamasıdır. İkinci katkısı, 30 fps'de çekilen yüksek kaliteli ve büyük çözünürlüklü video görüntüleri, sürüş senaryoları ile ilgilenenler için değerli uzun süreli çekimler sunmasıdır. Üçüncü katkısı, çekimin

gerçekleştiği kontrollü koşullar, sekanslardaki her kare için kamera kalibrasyonu ve 3B poz izleme verilerinin sunulmasına izin vermesidir. İdeal olarak, algoritmalar bu bilgiye ihtiyaç duymaz, ancak otomatik kalibrasyonun alt görevi, daha yüksek seviyeli nesne analizinin önünde durmamalıdır. Son katkısı ise veri tabanı diğer resimler ve videolar için kesin sınıf etiketleri çizmek isteyen kullanıcılara yardımcı olmak için özel yapım yazılımla birlikte sağlanmasıdır (Brostow *vd.,* 2017).

Özdemir (2022), diş çürüklerini anlamsal bölütleme yapay zekâ modeli adlı çalışmasında, 500 adet bitewing radyografi röntgen görüntülerine diş çürüklerinin tespiti için anlamsal bölütleme yöntemini uygulamıştır. Araştırmada uygulanan yapay zekâ modelinin doğruluk oranının %86 olduğu tespit edilmiştir. Buna ek olarak iyileştirilen görseller üzerindeki çürükleri bulmak için eğitilen sistemin doğruluk oranında %90 başarı elde edilmiştir (Özdemir, 2022).

Wang (2018) ve arkadaşları, piksel bazında anlamsal bölütlemenin nasıl geliştirileceğini incelemiştir. Öncelikle çift doğrusal yukarı örneklemede eksik olan daha ayrıntılı bilgileri yakalayıp kodunu çözebilen piksel düzeyinde tahmin oluşturmak için yoğun üst örnekleme evrişimi (DUC) tasarlanmıştır. Ardından kodlama aşamasında bir hibrit genişletilmiş evrişim (HDC) çerçevesi önerilmiştir. Araştırmacılar önceden oluşturmuş oldukları iki proje üzerinde geliştirdikleri yaklaşımı kullanıp başarılı sonuçlar elde etmişlerdir. Bu sonuçlar aşağıdaki gibidir:

- KITTI Yol Bölütlemesi, 289 eğitim görüntüsü ve 290 test görüntüsü dâhil olmak üzere üç farklı yol sahnesi kategorisinin görüntülerini içermektedir. Amaç, görüntülerdeki her pikselin yol olup olmadığına karar vermektir. Test setinde %80,1 mIOU'luk bir sonuç elde edilmiştir.
- PASCAL VOC2012 veri setinde, 1464 eğitim görüntüleri, 1449 doğrulama resimleri ve 1456 görüntüleri test edilmiştir. Sonuçlar; önce artırılmış VOC2012 eğitim seti ve MS-COCO veri setinin bir kombinasyonunu kullanarak 152 katmanlı ResNet-DUC modeline önceden eğitilmiş ve ardından artırılmış VOC2012 trenval setini kullanarak önceden eğitilmiş ağa ince ayar yapılmıştır. Test modelinde %83,1 başarı oranına ulaşılmıştır (Wang *vd.,* 2018).

Zhang ve arkadaşları (2018), çok ölçekli özellikler ve rafine sınırları kullanarak tam evrişimli ağ (FCN) çerçevesi ile pikselsel etiketleme için uzamsal çözünürlüğü iyileştirmede önemli ilerleme kaydetmiştir. Sahnelerin anlamsal bağlamını yakalayan ve sınıfa bağlı özellik haritalarını seçici olarak vurgulayan bağlam kodlama modülünü tanıtarak, küresel bağlamsal bilginin anlamsal bölümlemedeki etkisi araştırılmıştır.

Yaklaşımlarında, PASCAL-Context'te %51,7 mIoU, PASCAL VOC 2012'de %85,9 mIoU sonuçları raporlanmıştır. Araştırmada 14 katmanlı ağımız, 10 kat daha fazla katmana sahip son teknoloji yaklaşımlarla karşılaştırılabilir olan %3.45'lik bir hata oranına ulaşılmıştır (Zhang vd., 2018).

Az çekimli anlamsal bölütleme için hibrit özellikli iyileştirme ağı çalışmasında, az çekimli anlamsal bölütleme yöntemleri bilgisayar ile görme alanında yaygın olarak çalıştırılmasına rağmen hala geliştirilmeye ihtiyacı vardır. Min ve arkadaşları (2022) bu odakta özellik gösterimini doku bilgisi ile zenginleştirmeyi ve kayıplara uyarlanabilir ağırlıklar atamayı önermektedir. Araştırmacılar özellikle doku geliştirme modülü tarafından elde edilen doku bilgisini ResNet' teki katman özellikleri ile birleştirip hibrit bir özellik eklemişlerdir. Doku bilgisinin dâhil edilmesi ile destek kümesi rehberliğini daha etkili hale getirmek için benzerlik hesaplaması geliştirilmiştir. Çalışmada HsNet'in öne çıkan performansı nedeniyle kodlayıcı- kod çözücü ağı olarak HSNET yapısı seçilmiştir. Önerdikleri yöntemin verimliliğini benchmark veri seti PASCAL-5, COCO-20, FSS-100 ile değerlendirmişlerdir. Önerdikleri yöntemin nihai ortalama IoU altında karşılaştırılan yöntemden sürekli olarak daha iyi bölütleme doğruluğu elde ettiği gözlemlenmiştir. Yapılan deneyler sonucunda PASCAL-5 veri setinde önerilen yöntem en iyi yaklaşıma göre %0,4 ve %0,8 iyileştirme elde edildiği gözlenmiştir. COCO-20 veri setinde önerilen yöntem temel HSNet'e göre %2,2 ve %4 iyileştirme; FSS-100 veri seti için önerilen yöntem ile %1,5 ve %0,8 ortalama IoU iyileştirme sağlandığı tespit edilmiştir (Min vd., 2022).

Yuan ve arkadaşlarının (2022) hazırladığı zayıf denetimli anlamsal bölütleme çalışmasında, örüntü yaşı, bölge, piksel ve nesne sınırı seviyelerinde zıt örnek çiftlerinin benzerlik oranlarından yararlanarak gelişmiş özellik temsilleri elde etmek ve WSSS performansını arttırmak için bir MCL kodlayıcı ve BAECON kod çözücüyü içeren stratejili zıt öğrenme metodu MuSCLe önerilmiştir. Bu method ile veri seti olarak PASCAL VOC 20212 kullanılmıştır. Nesne özelliği çıkarımının hem genellemesi hem de ayırt ediciliğini kanıtlamak amacıyla çoklu kontrast öğrenme kodlayıcısı (MCL) oluşturulmuştur. Bu sayede, WSSS yönteminde kullanılan sınıflandırma kaybı terimine ek olarak, kodlayıcı temsilini geliştirmek için eşleştirilmiş rastgele kırpılmış nesne yamalarının örtüşen bölgelerine dayanan görüntü seviyesi kontrastı, piksel kontrastı ve ikili bölgesel kontrast bulunmaktadır. Geliştirilmiş aktivasyon haritasından türetilen sözde maskelerden çıkarılan sınırlar boyunca özellikleri öğrenerek kod çözücüyü geliştirmek için yeni bir sınıf tabanlı kontrast öğrenme yöntemi olan BAECON

(Boundary Enhancement viA Contrastive Orientation Navigation) tasarlanmıştır. MCL, farklı seviyelerde kontrastlı öğrenme yoluyla CAM verilerini iyileştirmek için tasarlanırken, BEACON, kontrastlı bir şema aracığıyla nesne sınırları etrafındaki özellik temsillerini geliştirmek istenmektedir. Muscle metodunun veri seti üzerinde yapılan çalışmalar sonucunda zayıf denetimli anlamsal bölütleme ile geliştirilen diğer modellerden daha başarılı olduğu gözlemlenmiştir (Yuan vd., 2022)

Sijie ve arkadaşlarının (2022) anlamsal görüntü bölütleme için Çok Modlu Denetimsiz Etki Alanı Uyarlaması odağında gerçekleştirdiği çalışmasında, uyarlama performansını artırmak için ek bilgilerden yararlanmayı amaçlayan anlamsal bölütleme için yeni birçok modlu UDA çerçevesi sunulmaktadır. Bunu yapmak için, derinlik görüntüsü yardımcı bir girdi olarak ele alınmış, modeli iki zorlukla karşılaştığı çok modlu bir öğrenme paradigması altında eğitilmiştir. Anlamsal bölümleme için yeni birçok modlu tabanlı Unsupervised Domain Adaptation (UDA) (Denetimsiz Etki Alanı Uyarlaması) yöntemi, son zamanlarda derinliğin, RGB temsilini geliştirmek için geometrik ipuçları sağlamakla ilgili bir özellik olduğu kanıtlanmıştır. Ancak, mevcut UDA yöntemleri yalnızca RGB görüntülerini işlemekte veya ek olarak yardımcı bir derinlik tahmini görevi ile derinlik farkındalığını geliştirmektedir. Yerel şekil ve göreceli konum gibi anlamsal bölütleme için çok önemli olan geometrik ipuçlarının, yalnızca renk (RGB) bilgisi ile yardımcı bir derinlik tahmini görevinden kurtulmanın güç olduğu tespit edilmiştir. Sonuç olarak, önerilen yöntem uyarlama performansını önemli ölçüde arttırmış ve SYNTHIA-to-Cityscapes, GTA5-to-Cityscapes ve SELMA-to-Cityscapes UDA ayarlarında yalnızca RGB tabanlı yöntemlere karşı olumlu performans göstermiştir (Sijie vd., 2022).

Isola (2017), görüntüden görüntüye çeviri sorunlarına genel amaçlı bir çözüm olarak koşullu çekişmeli ağları araştırmıştır. Koşullu çekişmeli ağlar, yalnızca girdi görüntüsünden çıktı görüntüsüne eşlemeyi öğrenmekle kalmaz, aynı zamanda bu eşlemeyi eğitmek için bir kayıp fonksiyonunu da ortaya koymaktadır. Ayrıca araştırmada bu yaklaşımın etiket haritalarından fotoğrafların sentezlenmesinde, kenar haritalarından nesnelerin yeniden yapılandırılmasında ve görüntülerin renklendirilmesinde etkili olduğu belirtilmiştir. Ağ, otomatik kenar algılama konusunda eğitilmiş ve insan çizimleri üzerinde test edilmiş bir model tarafından oluşturulmaktadır. U-Net tabanlı bir mimari kullanmaktadır (Isola, 2017).

Qin ve arkadaşları (2022) tarafından yapılan, manyetik rezonans beyin görüntülerinin çapraz modalite sentezi için koşullu GAN'larda stil aktarımı çalışmasında, manyetik rezonans görüntülemenin teşhis sürecinde çeşitli belirsizlikler nedeniyle yüksek

kaliteli çok modlu MR görüntüleri elde edilmesi zor olduğu için koşullu üretken çekişmeli ağlar (cGAN) mimarisine stil aktarımı tanıtılmıştır. MR görüntülerinin çapraz modalite sentezini ele almak için hiyerarşik özellik haritalama ve füzyona (ST-cGAN) sahip bir cGAN modeli önerilmiştir. Çoğu cGAN tabanlı yöntemin yaptığı gibi piksel bazında benzerliğe odaklanmanın üstesinden gelmek için, önerilen ST-cGAN stil bilgisinden yararlanmış ve bu sentetik görüntünün içerik yapısına uygulanmıştır. Modaliteler arası görüntü sentezi, pikselden piksele bir haritalama problemi olduğundan, koşullu GAN tabanlı yöntemlerin çoğu, ağı tasarlarken sentetik görüntü ile referans modalite görüntüsü arasındaki piksellerin bire bir eşleşmesine odaklanır. Bunlar arasında pix2pix yöntemi kullanılarak iki görüntü arasındaki piksel bazında yoğunluk benzerliğini en üst düzeye çıkarılması amaçlanmaktadır. İki modalitenin görüntülerini koşullu girdi olarak alan ST-cGAN, farklı düzeylerdeki stil özelliklerini göstermekte ve bunları stille zenginleştirilmiş sentetik görüntüyü oluşturmak için içerik özellikleri ile bütünleştirmektedir. Ayrıca önerilen model, generatore gürültü girdisi eklenerek rastgele gürültüye dayanıklı hale getirilmiştir. Önerilen model IXI veri seti kullanılarak eğitilmekte, nihayetinde farklı değerlendirme ölçütlerinin uygulanması, her yöntemin güçlü ve zayıf yanlarını başarılı bir şekilde yansıtmakta ve deneysel sonuçlar ST-cGAN'ı başarısını doğrulamaktadır (Qin vd., 2022).

Yang ve arkadaşları (2022), SAR görüntülerinin neden olduğu benek gürültüsü nedeniyle, insanların zemin nesnelerini karmaşık arka plandan ayırt etmesinin zor olduğunu fark etmiştir. Söz konusu problemde yaygın olarak kullanılan çözüm, SAR görüntülerinin yer nesnelerini zengin renk bilgileri ile net bir şekilde sunabilen optik görüntülere çevirmektir. Bunun için üretken çekişmeli ağlardan (GAN) yararlanılmıştır. Geleneksel GAN tabanlı çeviri yöntemleri, konturun bulanıklaşmasına, dokunun kaybolmasına ve renk tutarsızlığına neden olabileceğinden geliştirilmiş koşullu GAN (ICGAN) yöntemi önermişlerdir. Temel CGAN modeliyle karşılaştırıldığında, geliştirilen ICGAN yöntemi üstünlükleri:

1. **Kontur keskinliği:** Düşük seviyeli ve yüksek seviyeli özellikleri birleştirmek için paralel dallardan yararlanılmış ve böylece görüntü kontur bilgisi, gürültünün etkisi olmadan iyileştirilmiştir. Düşük seviyeli ve yüksek seviyeli bilgileri birbirine bağlamak için bir paralel özellik füzyon üreteci kullanılmıştır.

2. **Doku inceliği:** Görüntünün yerel ve küresel doku özelliklerini zenginleştirmek için çok ölçekli alıcı alanları kullanarak görüntü ayırt edilmiştir.
3. **Renk doğruluğu:** Oluşturulan görüntü ile gerçek optik görüntü arasındaki renk boşluğunu azaltmak için Gauss bulanıklığı konvolüsyonuna dayanan renk sapması kaybı kullanılmıştır.

Yapılan çalışmalar sonucunda, eğitilen VGG19 ağı ile birden fazla yöntem ile karşılaştırılmıştır. Bunlar, pix2pix, NiceGAN, CycleGAN'dır. Başarı oranı %97.73 'tür (Yang vd., 2022).

Sun ve arkadaşları (2023), yüz ifadesi tanıma için geliştirilmiş CGAN ile ayırt edici derin bir füzyon yaklaşımı hazırlamıştır. Araştırmada yüz ifadelerinin soyut temsilini öğrenmek için geliştirilmiş bir koşullu üretken çekişmeli ağa (imcGAN) dayalı, ayrımcı bir şekilde derin füzyon (DDF) yaklaşımı önerilmiştir. Başlangıçta, im-cGAN'ı daha etiketli ifade örnekleri oluşturmak üzere eğitmek için eylem birimleri içeren yüz görüntüleri faydalanılmış ve birleşik öznitelik gösterimini elde etmek için küresel tabanlı modül (yüz görüntülerini girdi olarak alarak genel yüz bilgilerini dikkate alan) tarafından öğrenilen genel öznitelikleri ve bölge tabanlı modül (gözler, burun, kaşlar, ağız ve çene) tarafından öğrenilen yerel öznitelikleri kullanılmıştır. Birleşik özelliklerin ayrımını geliştirmek için sınıf içi mesafeleri en aza indirirken sınıflar arası varyasyonları genişleten ayrımcı kayıp fonksiyonunu tasarlanmıştır. Bu çalışma, JAFFE, CK+, Oulu-CASIA ve KDEF veri kümeleri üzerinde uygulanmıştır. JAFFE veri seti, on Japon kadından alınan 213 yüz görüntüsü mevcuttur. Veri seti içerisinde 7 farklı yüz ifadesi vardır. Öfke, tiksinti, korku, mutluluk, nötr, üzüntü ve sürprizdir. K+ veri seti ise 593 görüntü dizisini içermektedir. Oulu-CASIA veri seti, 80 kişinin görüntü dizisinde oluşmaktadır. Görünür ışık ve güçlü aydınlatma ile yakalanan 6 farklı yüz ifadesi (öfke, tiksinti, korku, mutluluk, üzüntü ve şaşkınlık dâhil) seçilmiş ve CK+ veri setine benzer şekilde, uyarlanmıştır. KDEF veri setinde ise öfke, iğrenme, korku, mutluluk, nötr, üzüntü ve sürpriz olmak üzere 5 farklı bakış açısından 4900 yüz ifadesi görüntüsü bulunmaktadır. Yalnızca yüzün görüneceği görüntülerden faydalanıldığı için veri setinde, 490 yüz ifadesi görüntüsü bulunmaktadır. imcGAN yönteminin JAFFE veri seti üzerindeki başarısı, mutluluk, tarafsız ve sürpriz ifadeleri tanımada ortalama %98,37'lik bir orana sahiptir ve en yüksek doğruluğa ulaştığını görülmektedir. CK+ veri setinde, ortalama %98,10 tanıma doğruluğu elde edildiği görülmektedir. Oulu-CASA veri setindeki karışıklık matrisi, yüz

kaslarının benzer hareketlerine sahip oldukları için ifadeleri (iğrenme, öfke ve üzüntü) karıştırmanın kolay olduğunu göstermiştir. KDEF veri setinde, korku ifadesi, nötr ve üzüntü olarak yanlış sınıflandırılmıştır. Çünkü verilen üç yüz ifadesinin görünümü KDEF veri setinde benzerdir (Sun vd., 2023).

Rahimol ve Maheswari (2022) tarafından yapılan antik tapınak duvar resimlerinin cGAN ve PConv ağları kullanılarak restorasyonu çalışmasında, tapınak duvar resimlerinin çoğu üzerindeki kir ve lekeler yüzünden veya zamanla renklerinin solduğu tespit edilmiştir. Bundan ötürü çizilen tabloların anlaşılması güç hale gelmiştir. Bu nedenle araştırmada, antik tapınak duvar resimlerinin yeniden inşası için bu çoklu rastgele düzensizlikleri göz ardı eden verimli bir iç boyama tekniği önerilmiştir. Önerilen yöntemde, hem maskelerin (beyaz piksellerin biraz düzenleme gerektirdiği, ancak siyah piksellerin gerekmediği siyah beyaz bir görüntü örneği) otomatik olarak geliştirilmesi hem de bozulmuş bölümlerin tanımlanması için cGAN'dan yararlanılmıştır. Bu maskeler, geliştirilen yöntem ile bir sonraki resim parçasının doldurulması gerektiğini tahmin etmek için bir hiper parametre olarak kullanılmış ve bozulan duvar resimlerini, evrişimin maskelendiği ve yalnızca geçerli piksellerde şartlandırılmak üzere yeniden normalleştirildiği kayan pencere tabanlı Derin Evrişim Ağı kullanarak onarmaktadır. Veri seti, 2000 orijinal duvar resmini içermektedir. Her görüntünün boyutu 512×512 ‘ye düşürülmüştür. Veri setinden (%10) rastgele bir eğitim seti (%80), bir doğrulama seti (%10) ve bir test seti seçilmiştir. Sonuç olarak; gerçek hasarlı duvar resimleriyle yapılan karşılaştırmada, önerilen yöntemden çıkan sonucun görüntü yapısı ve dokusu üzerinde yama eşleştirme tekniklerinden daha iyi görsel performans elde ettiği görülmüştür (Rahimol vd., 2022).

3. MATERYAL VE YÖNTEM

Araştırmanın materyaline ilişkin bilgi sunulduktan sonra, yöntem bölümüne geçilmiştir. Bu bölüm kapsamında derin öğrenme, evrişimli sinir ağları, transfer ağları, ResNet, VGG19, üretken çekişmeli ağlar, koşullu üretken çekişmeli ağlar ve Pix2PixHD kavramları ele alınmıştır.

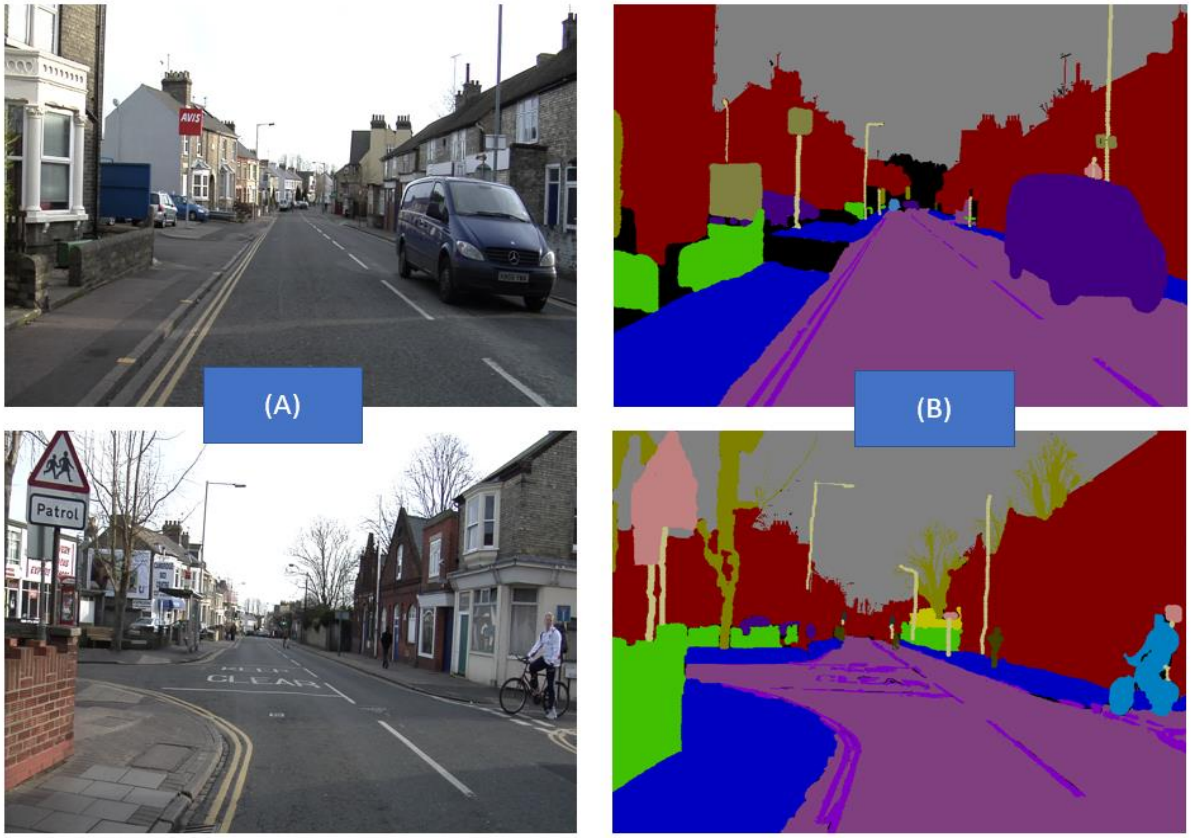
3.1. MATERYAL

Tez çalışmasında CamVid (Cambridge-driving Labeled Video, Cambridge-sürüş Etiketli Video) görüntü veri seti kullanılmıştır (Brostow *vd.*, 2009). Bu veri seti sürüş esnasında kayıt altına alınan videolardan elde edilmiştir. Veri seti yüksek çözünürlüklü kameralar ile elde edilmiştir. Veri setindeki her görüntünün içeriği daha önce belirlenmiş 32 anlamsal sınıf ile ilişkilendirilmiştir. Meta verilerle tamamlanmış, nesne sınıfı anlamsal etiketlere sahip ilk video koleksiyonu olarak belirtilmektedir.

Toplanan görüntülere daha sonra piksel bazlı anlamsal bölütleme manuel olarak uygulanmıştır. Bölütleme işlemi Gretag–Macbeth renk şemasına göre gerçekleştirilmiştir. Veri seti çözünürlüğü 960x720 olan 701 görüntü içermektedir. Veri setindeki etiketler boşluk, bina, duvar, ağaç, bitki örtüsü, çit, kaldırım, park bloğu, sütun/direk, trafik konisi, köprü, işaret, çeşitli metin, trafik ışığı, gökyüzü, tünel, kemer yolu, yol, yol kenarı, şerit işaretleri (sürüş), şerit işaretleri (sürüş dışı), hayvan, yaya, çocuk, araba bagajı, bisikletçi, motosiklet, araba, SUV/pikap/kamyon, kamyon/otobüs, tren ve diğer hareketli nesneler olarak belirlenmiştir. Etiketleme için kullanılan 32 nesne sınıfı ve bunlara karşılık gelen renklerin listesi aşağıda verilmiştir (Bkz: Şekil 3.6.).

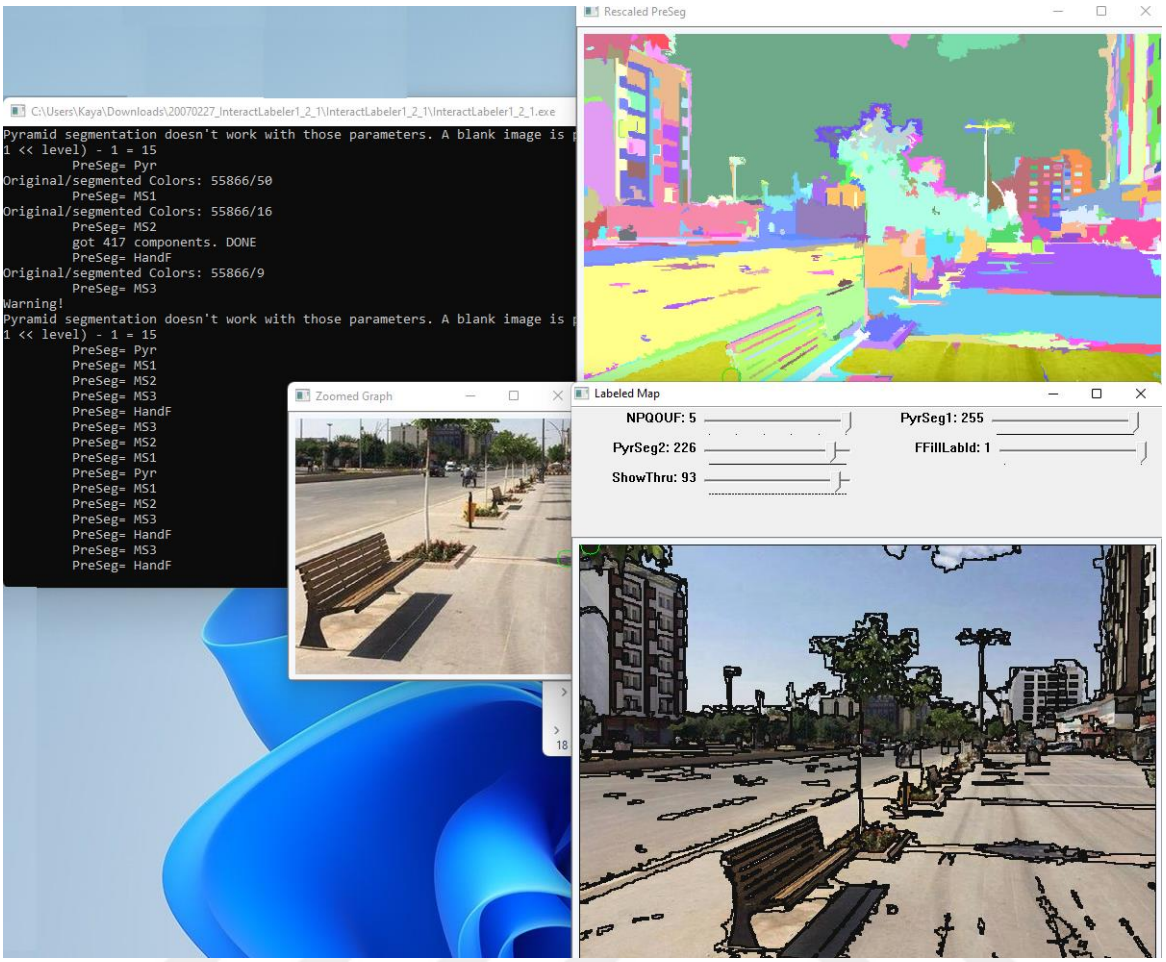
Şekil 3. 1. Etiketleme için kullanılan 32 nesne sınıfı ve bunlara karşılık gelen renklerin listesi (Brostow vd., 2009).

CamVid veri setindeki örnek 2 görüntü ve piksel etiket görüntüye ait bilgiler Şekil 3.2'de verilmiştir.



Şekil 3. 2. Veri setindeki örnek gerçek ve piksel görüntüler. (A) gerçek görüntüler, (B) piksel görüntüler

CamVid veri setindeki gerçek görüntülerin piksel görüntüleri diğer bir ifade ile etiketleri oluşturmak için Şekil 3.3'teki program kullanılmıştır. Her gerçek görüntünün piksel etiketi tek tek oluşturulmuştur.

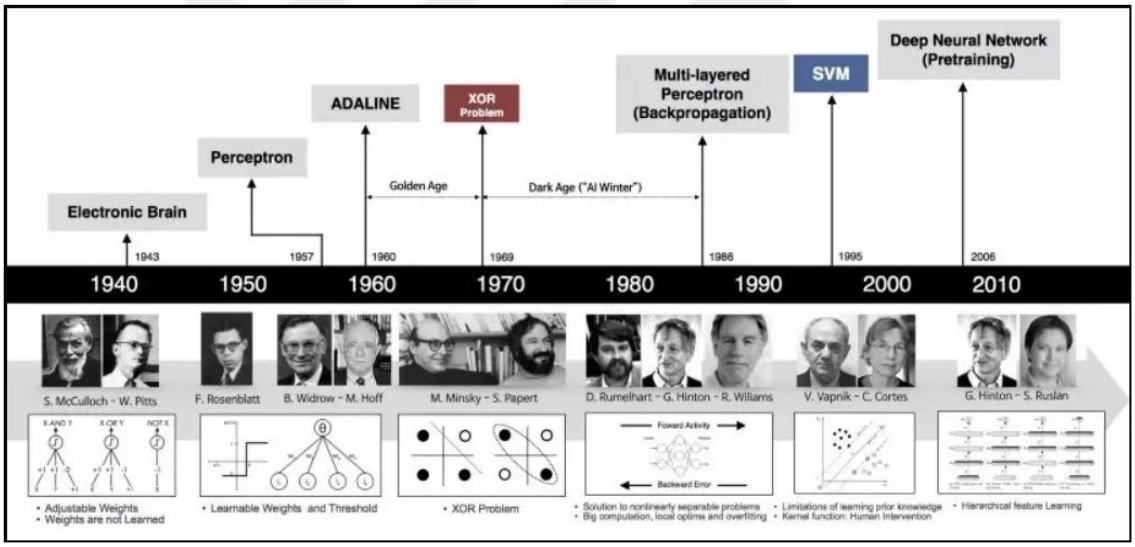


Şekil 3. 3. Görüntülere bölütleme işleminin uygulanması

3.2. YÖNTEM

3.2.1. Derin Öğrenme

Derin öğrenme kavramı, literatürde çeşitli araştırmalarda farklı şekillerde tanımlanmaktadır. Deng ve Yu (2014) derin öğrenmeyi denetimli veya denetimsiz özellik çıkarma, dönüştürme, desen analizi ve sınıflandırma için birçok doğrusal olmayan gizli katmandan yararlanan bir makine öğrenmesi teknikleri sınıfı olarak tanımlamaktadır. Gu vd. (2016) ise derin öğrenmeyi bilgisayarların, deneyimlerden öğrenmelerini ve kavramların hiyerarşisi açısından anlamalarını sağlayan bir makine öğrenimi olarak tasvir etmektedir. Derin öğrenmedeki "Derin" sözcüğü, modelin veri kümesini eğitmek için kullandığı gizli katmanların sayısını belirtmektedir. Derin öğrenme algoritmaları, özellikle görüntü, ses ve metin verileri gibi karmaşık yapıdaki verilerin analizinde etkilidir. Derin öğrenme makine öğrenmesinin bir alt yöntem grubunu ifade etmektedir. Makine öğrenmesinin tarihsel gelişimi Şekil 3.4'de verilmiştir.



Şekil 3. 4. Makine öğrenmesi tarihsel gelişimi

Derin öğrenme, günümüzde çok farklı alanda kullanılmaktadır. Sosyal medya söz konusu alanlardan biridir. Sosyal iletişim platformlarına yüklenen fotoğraflar, derin öğrenme yüz tanıma algoritması sayesinde iletişim ağlarındaki arkadaşlarımızın otomatik olarak etiketlenmesini mümkün kılmaktadır. Siri, Google Asistan gibi dijital asistanlar, doğal dil işleme ve konuşma tanıma yapılarını kullanarak iş ve işlem süreçlerini kolaylaştırmaktadır. Spam e-postalar metin ayırma yöntemleri kullanarak bireylerin zaman ve çaba harcamasına gerek kalmadan kötü amaçlı olduğu anlaşılan mesajları spam klasörüne göndermektedir.

Derin öğrenme, insan beyninin çalışma şekline göre genel hatlarıyla modellenen algoritmalar olan sinir ağlarının katmanları tarafından desteklenen bir yapıdır. Büyük miktarlarda veri ile eğitim, sinir ağındaki nöronları konfigüre eden model, eğitildikten sonra yeni verileri işlemektedir. Derin öğrenme genellikle sınıflandırma işlemleri, yüz tanıma ve nesne tespiti işlemi için kullanılmaktadır (Deng ve Yu 2014).

Dünyada birçok firma (Google başta olmak üzere) derin öğrenme uygulamalarının geliştirilmesi için kullanıcılarına bulut tabanlı hizmetler sunmaktadır. Güçlü bir donanıma sahip olamasak bile bu hizmetten faydalandığımız takdirde modeli geliştirmek daha kolay olacaktır.

Derin öğrenme metotları yeni nesil yapay sinir ağlarıdır. Katmanlardan oluşan makine öğrenmesi yaklaşımlarıdır. Genel olarak aşağıda belirtilen katmanları kullanmaktadır.

3.2.1.1. Derin Öğrenmenin Katmanları

- **Giriş Katmanı**

Nöron katmanı olarak da adlandırılan giriş katmanı, girdi verilerini alarak, gizli katmanlara iletmektedir. Nöronlar arasındaki tüm bağlantıların belirli bir ağırlıktan oluşması, girdi değerlerinin ne kadar önemli olduğunu göstermektedir.

- **Gizli Katman**

Giriş katmanından gelen veriler üzerinde hesaplamalar yapan gizli katmanlar, sinir ağlarının oluşturulmasındaki güçlük, oluşturulacak olan nöronların sayısına ve gizli katman sayısına karar vermektir.

- **Çıkış Katmanı**

Sınıflandırma katmanı olarakta bilinmektedir. Çıkış katmanında gerekli çıktı üretilmekte ve çıkışları standart hale getirerek modelin verimli bir şekilde çalışmasını sağlanmaktadır. Bunun için aktivasyon fonksiyonu kullanılmaktadır.

Giriş katmanı, gizli katman ve çıkış katmanı döngüsü, çalışmanın temel akışıdır. Derin öğrenme modeli örneği Şekil 3.5'te gösterilmiştir.

Şekil 3. 5. Derin öğrenme modeli

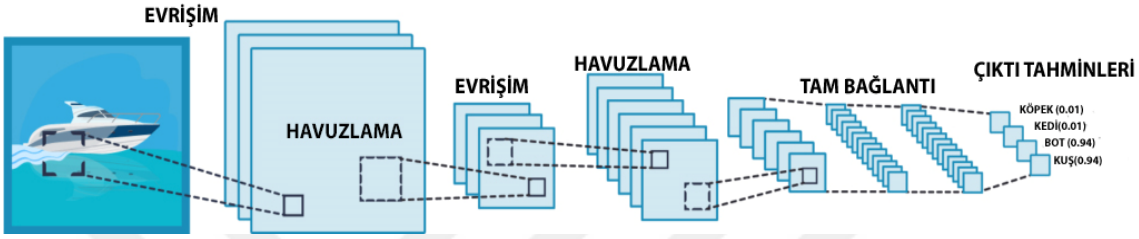
Derin öğrenmedeki "Derin" sözcüğü, modelin veri kümesini eğitmek için kullandığı gizli katmanların sayısını belirtir. Derin öğrenme modelinin başarılı olma oranının artması modele verilen girdi sayısının fazla olması ile doğrudan ilişkilidir. Girdi sayısı arttıkça sonuçların da daha verimli olduğu görülmektedir (Lateef ve Ruichek, 2019). Derin öğrenmede başarı ise örneklem ile yakından ilişkilidir. Buna göre örneklem üzerinden derlenen verilerde farklı uygulamalar kullanılabilmektedir. Derin öğrenmede algoritmalarının temeli, insan beynindeki bir nöronun bilgisayara benzetilmesi temel alınarak ortaya çıkmıştır. Derin öğrenme modeli, Donald Hebb'in sinir hücresi içindeki sinir ağlarının yapısını incelenmesi ile başlamıştır. Devamında bilgisayar ortamında matematiksel olarak modelinin oluşturulması yapay sinir ağlarının çıkış noktası olarak bilinmektedir (Hebb vd., 2005).

3.2.2. Evrişimli Sinir Ağları

Görüntü işleme ile bilgisayarla görü alanları için derin ağların en önemli bileşeni eğitilebilir çok katmanlı evrişimli ağlardır (Szeliski, 2022). Evrişimli sinir ağları ismini, filtreleme özelliği ile verilerden çıkardığı öznitelik işleminden almaktadır. Evrişimli sinir ağları ardı ardına dizilmiş birden çok eğitilebilir katmandan oluşmaktadır. Alex Krizhevsky ve arkadaşları tarafından 2012 yılında geliştirilen AlexNet modeli ImageNet yarışmasını kazanarak Evrişimli sinir ağlarının popüler hale gelmesini sağlamıştır (Krizhevsky vd., 2012).

Evrişimli sinir ağları, gradyan azalması ile eğitilmiş çok katmanlı derin ağların, geniş örnek koleksiyonlarından karmaşık, yüksek boyutlu, doğrusal olmayan eşlemeleri öğrenme yeteneği, onları görüntü tanıma görevleri için bariz adaylar haline getirmektedir (LeCun vd., 1998).

Evrişimli sinir ağları klasik makine öğrenmesinden farklı olarak, modeldeki veri miktarı eğitim ve performansı etkilemektedir. Eğitim çıktıların tahmin doğruluğunu arttırmak için yüksek miktarda veri gerekmektedir. Veri miktarının yüksek olması nedeniyle hesaplamaların yapılması sürecinde güçlü bir donanımın varlığı gereklidir. Evrişimli sinir ağları, öznitelik çıkarımı ve sınıflandırma olarak iki ana bölümden oluşmaktadır. Öznitelik çıkarımı bölümünde, evrişim, havuzlama katmanları bulunmaktadır. Sınıflandırma bölümünde ise tam bağlantı katmanı bulunmaktadır.



Şekil 3. 6. Evrişimli sinir ağları mimarisi

Yukarıda bir evrişimli sinir ağı mimarisi verilmiştir (Bkz: Şekil 3.6.). Buna göre evrişimli sinir ağının girdi verileri işlenebilir bir formata dönüştürülerek giriş katmanına aktarılmaktadır. Sonrasında evrişim katmanında veriler üzerinde filtrelerin dolaşması sonucu özniteliklerine göre işlenmekte, akabinde havuzlama yapılarak devamında sınıflandırılmış çıktılar çıkış katmanına aktarılmaktadır. Elde edilen sonuç ve hedeflenen sonuç arasındaki fark, geri yayılım algoritması ile bütün ağırlıklara sevk edilmektedir. Her bir döngüde bu fark optimize edilerek hedeflenen sonuca erişilmektedir. Standart yapay sinir ağlarında evrişim katmanı bulunmamaktadır. Girdi direkt gizli katmanlara bağlanmaktadır. Evrişimli sinir ağlarında öznitelik çıkarma işlemi evrişim katmanlarında otomatik olarak gerçekleştiği için, gerçek veri doğrudan giriş katmanına verilebilmektedir. Evrişimli sinir ağları katmanlarına ilişkin genel bilgiler aşağıda sunulmuştur.

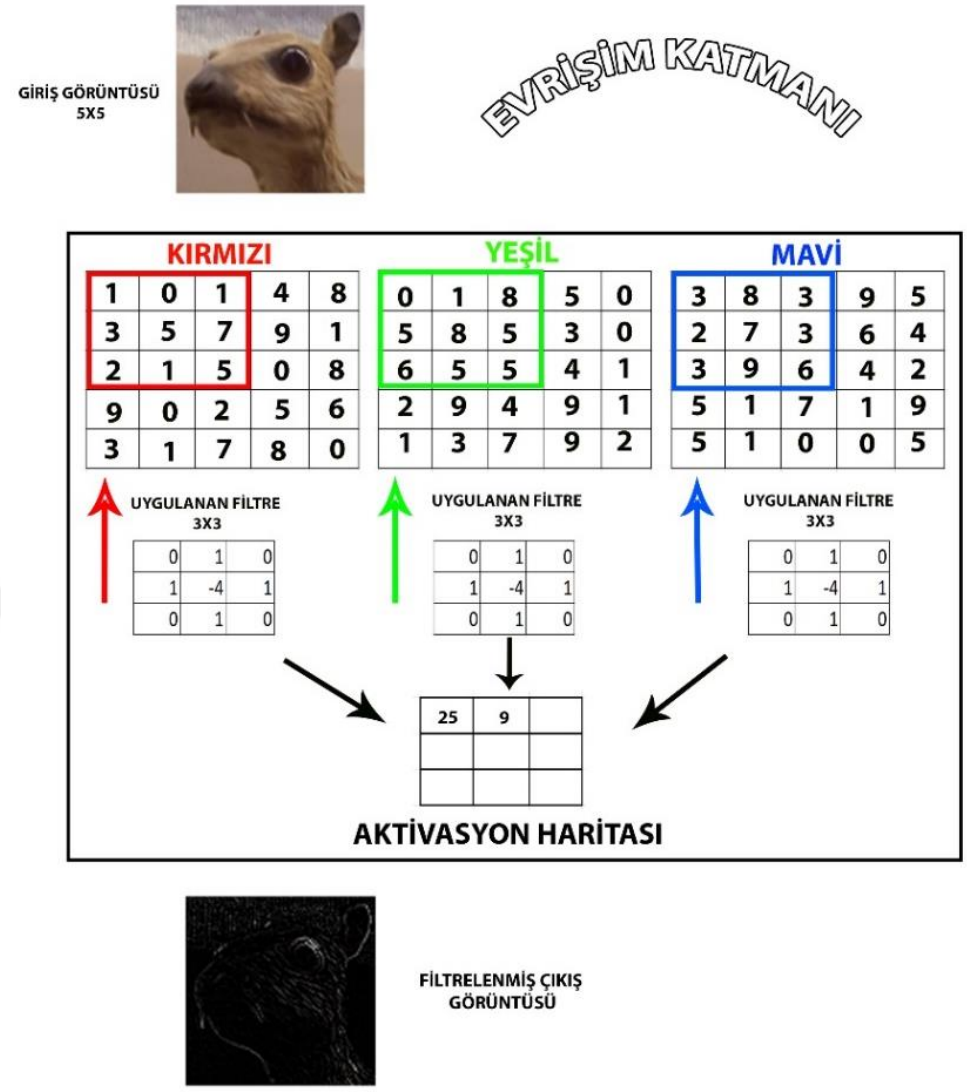
3.2.2.1. Giriş katmanı

Veri (görüntü) içerisindeki gerçek boyuttaki değerler işlenebilir formata getirilerek giriş katmanını beslemektedir. Modelin performansı, verinin boyutu ile yakından ilişkili olduğundan yüksek boyutlu verilerle yüksek doğrulukta çıkarımlar elde edilebilmektedir. Buna karşın, veri boyutunun artışı ile modelin eğitim süresi uzamakta ve depolama alanına ihtiyaç duyulmaktadır.

3.2.2.2.Evrişim Katmanı

Evrişim katmanı en önemli katmandır. Evrişimli sinir ağlarında genellikle çok sayıda evrişim katmanı bulunmaktadır. Evrişim katmanlarındaki özel tasarlanmış filtre değerleri ile matrislerde birden fazla çarpma, çıkarma, toplama işlemleri yapılmakta ve bu işleme evrişim adı verilmektedir. Evrişim işleminin temel amacı, görüntüden öznitelik çıkarımı (kenar, köşe, görüntü içerisinde nesne tanımlama gibi) gerçekleştirmektir. Bu sayede görüntü özelliklerini öğrenilmesi mümkün hale gelmektedir. Evrişim katmanında yapılması gereken Padding, Stride gibi işlemler matrisler kullanılarak gerçekleştirilmektedir. Görüntüye ait piksel değerleri matrislere dönüştürülüp işlemler yapılmaktadır. Görüntüye uygulanan filtrelerin katsayıları, boyutları gibi özellikleri, model mimarisini oluşturan tasarımcıların özgün seçimleridir (Uyanık, 2022). Bu nedenle tasarlanan filtrelerin özellikleri birbirinden farklı olabilmektedir. Ayrıca filtreler üzerinde bulunan katsayılar modelin eğitimi sırasında uygulanan her bir iterasyon sonucu yenilenmektedir. Bu sayede model, özniteliklerin çıkarılması için, verinin hangi bölgelerinin daha önemli olduğunu ortaya koymaktadır. Renk kanallarından (kırmızı, yeşil ve mavi gibi) gelen görüntü üzerinde uygulanan filtreleme işlemi sonucu tek bir matris elde edilmekte, verilerle oluşturulan matrise aktivasyon haritası adı verilmektedir (İnik ve Ülker, 2017). Evrişim işleminde kullanılan filtre çekirdek olarak da isimlendirilebilmektedir. Evrişimli sinir ağında verilen görüntü matrislere dönüştürülerek istenilen filtre uygulanmış ve aşağıdaki sonuçlar elde edilmiştir (Bkz: Şekil 3.7.).

Şekil 3.7’ de verilen giriş matrisinin boyutu 5x5 ve filtre matrisinin boyutu 3x3’tür. Beklenen çıktı matrisinin ise 3x3’tür. Çünkü 3x3 filtrenin 5x5 matris ile oluşturabileceği maksimum pozisyon sayısının 3x3’tür.



Şekil 3. 7. Katmanların filtrelenmesi ile oluşan aktivasyon haritası

Filtreleme işlemi, matrisin sol üst tarafından başlayarak her defasında giriş matrisinin üzerinde sağ tarafa doğru bir piksel (adım) olarak kaymaktadır. Tasarlanmış olan bu filtre, görüntünün üzerinde adım adım kayarken kendi matris değerleri ile örtüşen giriş matris değerlerini çarpmakta ve hepsini toplayarak her örtüşme için tek değer olarak vermektedir.

Evrişimli model, yapısı ve değer hesaplamaları ile birlikte evrişimli katmanı anlamak için bilinmesi gereken iki temel kavram vardır. Bunlar:

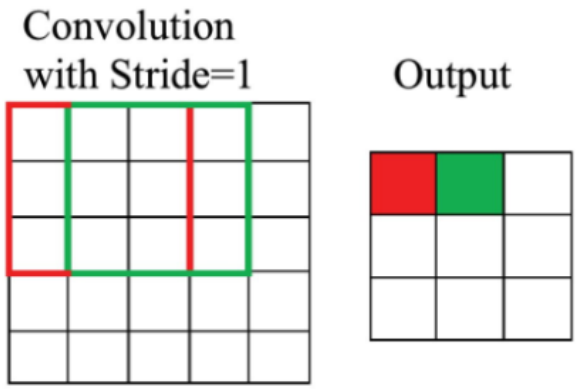
- Padding (Piksel Ekleme)
- Stride (Kaydırma Adımı)'dır.

Padding

Evrişim katmanlarının filtreleri, genellikle girdi verileri boyunca kaydırılırken, kenar noktaları işlemek için yeterli bilgiye sahip olmayabilmektedir. Bu nedenle, girdi verilerine kenarlık (Padding) eklemek gerekebilmektedir. Padding, genellikle verilerin etrafına sıfırlardan oluşan bir çerçeve ekleyerek yapılmaktadır. Bu çerçeve, verinin boyutunu artırmak veya mevcut boyutunu korumak için kullanılabilmektedir. Örneğin, bir 5x5 boyutundaki girdi verisi için, 2 piksel kenarlık eklemek, 9x9 boyutlu bir veriye dönüştürmektedir.

Stride

Veri kümesinde bir filtre boyutu ve kaydırma adımı belirleyerek verinin işlenmesinde kullanılan bir tekniktir. Filtre, boyutu kadar veriyi işlemek üzere belirli bir adım atlayarak pencereyi hareket ettirmektedir. Bu özellikle büyük veri kümesi işlemlerinde hesaplama zamanını azaltabilmekte ve aynı zamanda verinin boyutunu küçültmektedir. Şekil 3.8'de Stride (Kaydırma adımı) işlemine örnek verilmiştir.



Şekil 3. 8. Stride uygulanması

3.2.2.3.Aktivasyon Katmanı

Aktivasyon katman, evrişimli sinir ağının sonraki katmanlarının girdisini oluşturarak, ağın öğrenme sürecinin en önemli bileşenlerindendir. Aktivasyon fonksiyonu, konvolüsyonel katmanın çıkışını sıkıştırarak, verilerin daha iyi anlaşılmasını ve ağın öğrenme sürecinin hızlandırılmasını sağlar. Bu işlem, ağın performansını artırarak, doğru tahminler yapabilmesini mümkün kılar.

Aktivasyon fonksiyonları arasında en yaygın olanlarından biri, ReLU (Rectified Linear Unit) fonksiyonudur. ReLU, pozitif değerlerde doğrusal bir fonksiyondur ve sıfır veya negatif değerlerde sıfırdır. Bu fonksiyon, ağırlıkların güncellenmesindeki doğrusallığı korur ve aynı zamanda ağın derinliğini artırarak gradyanın kaybolma sorununu da önler.

Step Fonksiyonu: İkili sınıflandırma amaçlı kullanılır. Bie eşik değeri alarak 0 veya 1 üretir.

Sigmoid Fonksiyonu: En çok kullanılan fonksiyonlardan biri olup [0,1] arasında bir değer üretir. En çok kullanılan fonksiyonlardan biridir. “S” şekilli çizilen doğrusal olmayan bir fonksiyondur.

TanH: [-1,1] arasında bir değer üreten doğrusal olmayan bir aktivasyon fonksiyonudur. Sigmoid fonksiyonundan daha başarılıdır. Aslında sigmoid fonksiyonunun matematiksel olarak değiştirilmiş versiyonudur. Her ikisi de benzerdir ve birbirinden türetilebilir.

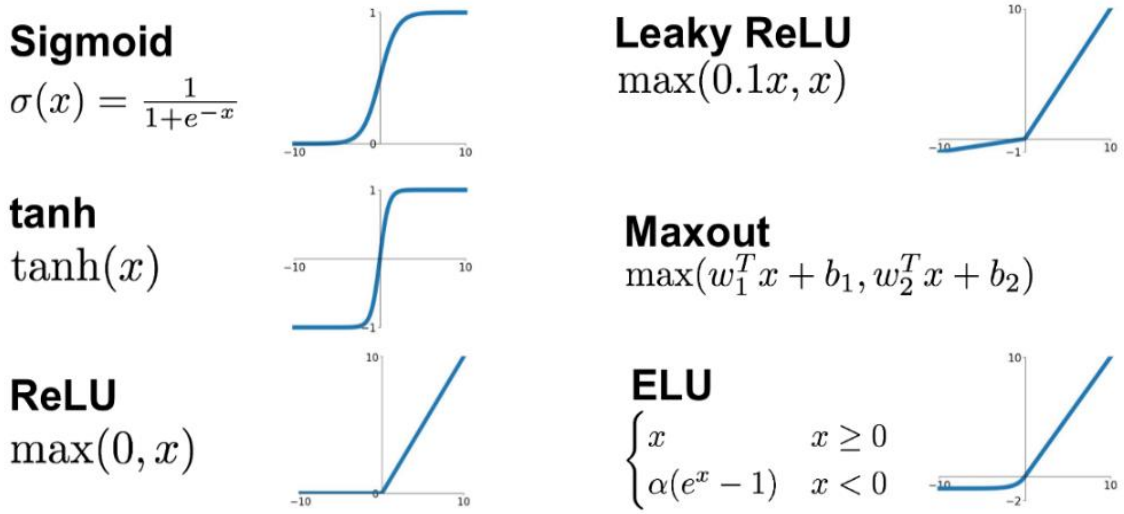
ReLU: ReLU (rectified linear unit) doğrusal olmayan bir fonksiyondur. ReLu doğrultulmuş doğrusal birim anlamına gelir. Derin öğrenme yöntemlerinde en çok kullanılan aktivasyon fonksiyonudur. Genel olarak ağların gizli katmanlarında tercih edilir.

ELU: Exponential Linear Unit-ELU fonksiyonu negatif girdiler hariç ReLU ile benzerdir. Negatif girdilerde ise genellikle 1.0 alınan alfa parametresi ile çarpımı almaktadır.

MaxOut: Bu aktivasyon fonksiyonu ReLU ve Leaky ReLU fonksiyonların bir genellemesidir. Girişlerin maksimum değerini döndüren doğrusal bir fonksiyondur. ReLU ve Leaky ReLU bu fonksiyonun özel durumlarıdır.

Leaky ReLU: ReLU benzeri bir aktivasyon fonksiyonudur. Ancak ReLU'daki gibi negatif değerler için düz bir eğim yerine küçük bir eğim bulunmaktadır. Bu eğim katsayısı modelin eğitiminden önce belirlenir.

Aktivasyon fonksiyonlarına ait grafikler aşağıda verilmiştir (Şekil 3.9).



Şekil 3. 9. Aktivasyon fonksiyonları

3.2.2.4.Havuzlama Katmanı

Evrişim katmanın çıkışını almakta ve veriyi özelliklerine göre küçültmektedir. Bu küçültme işlemi, ağın öğrenme sürecinin hızlanmasına yardımcı olmakta ve aynı zamanda aşırı uyum sorununu önlemektedir. Görüntü çok büyük olduğunda parametre sayısını azaltarak görüntüyü küçültmek işlemi yapmaktadır. Bu sayede hesaplama karmaşıklığı azaltmaktadır. Genellikle bu katmandaki işlemleri gerçekleştirmek için ortalama havuzlama ve maksimum havuzlama olarak isimlendirilen iki yöntem kullanılmaktadır.

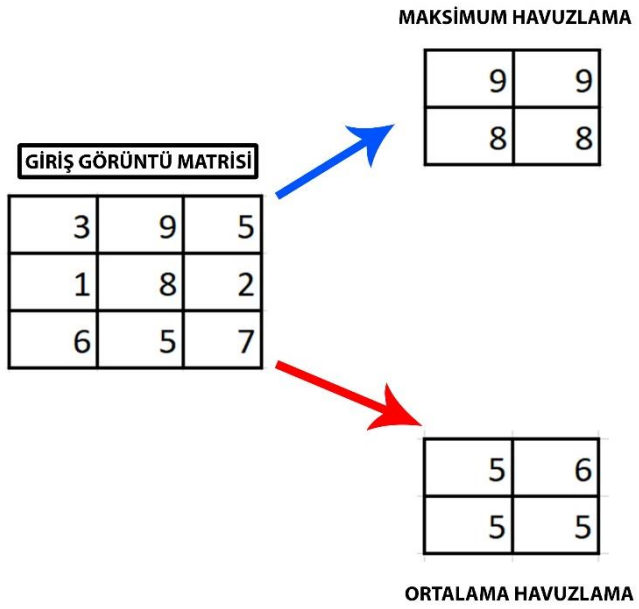
Ortalama Havuzlama

Ortalama havuzlama yöntemi, bir bölgedeki verilerin ortalamasını alarak verileri küçültmektedir. Bazı özel durumlarda, maksimum havuzlama yöntemine göre daha iyi sonuçlar verebilmektedir.

Maksimum Havuzlama

İşlemlerde en çok tercih edilen yöntemdir. Bir bölgedeki verilerin en büyük değerini seçerek verileri küçültmektedir. En büyük değeri seçerek, özellikle resimlerde nesnelerin kenarları veya belirgin özellikleri gibi önemli öznitelikleri korumaya yardımcı olmaktadır.

Ortalama havuzlama ve maksimum havuzlama yöntemleri Şekil 3.10 ‘da gösterilmiştir.



Şekil 3. 10. Maksimum ve ortalama havuzlama yöntemi örneklemesi

3.2.2.5. Tam Bağlantılı Katman

Fully Connected (FC) katmanı, evrişim katmanlardan elde edilen öznitelik haritaların geldiği katmandır. Klasik yapay sinir ağları yapısındadır. FC katmanında meydana gelen işlemler klasik yapay sinir ağlarındaki işlemlerdir. Bu katmandaki nöronlara gelen girişlere ait ağırlıklar ve eşik (bias) değerlerden oluşur. Öznitelikler bu katmana verilmeden önce düzleştirilmektedir. Diğer bir ifade ile matris formundaki öznitelik haritaları vektör haline getirilmektedir. Düzleştirme işleminden sonra bu katmanda öğrenme işlemi gerçekleştirilmektedir. Bir CNN ağın son birkaç katmanını oluşturmakta ve genellikle çıktı katmanından önce gelmektedir. FC katmanı, çıktı katmanı ile birlikte sınıflandırma işlemi amaçlı kullanılmaktadır (Zhou vd., 2017). Diğer bir deyişle bir girdi görüntünün ait olduğu sınıfı belirleyen katmandır.

3.2.2.6. Dropout (Bırakma)

Yapay sinir ağlarında, ağın eğitimi aşamasında overfitting (aşırı öğrenme) olarak adlandırılan istenmeyen bir durum gerçekleşebilmektedir. Bu olay ağın eğitim verilerini çok iyi derecede ezberlemesidir. Bunun yanı sıra ağın eğitim verileri üzerinde sınıflandırma performansının yüksek olduğu ancak aynı durumun test verileri üzerinde gözlenmemesi olayıdır. Bu durum istenmeyen bir problemdir. CNN ağlarda bu sorunun üstesinden gelmek için FC katmanındaki nöronların bir kısmının ağdan çıkarılarak ağın boyutunun küçültmesi ile gerçekleştirilebilir. Bu işlem FC katmanından sonra kullanılan bırakma (Dropout) katmanı ile gerçekleştirilir. Örneğin; 0,3'lük bir bırakma değeri FC katmanındaki nöronların %30'u ağdan çıkarılmasıdır. Bu şekilde ağın ezberleme yapan bir kısım nöronları ağdan çıkarılarak ağın ezberlemesinin önüne geçilmiş olunur (Shen *vd.*, 2019). Bu katmana bir düzenleme katmanı olarak bakılabilir.

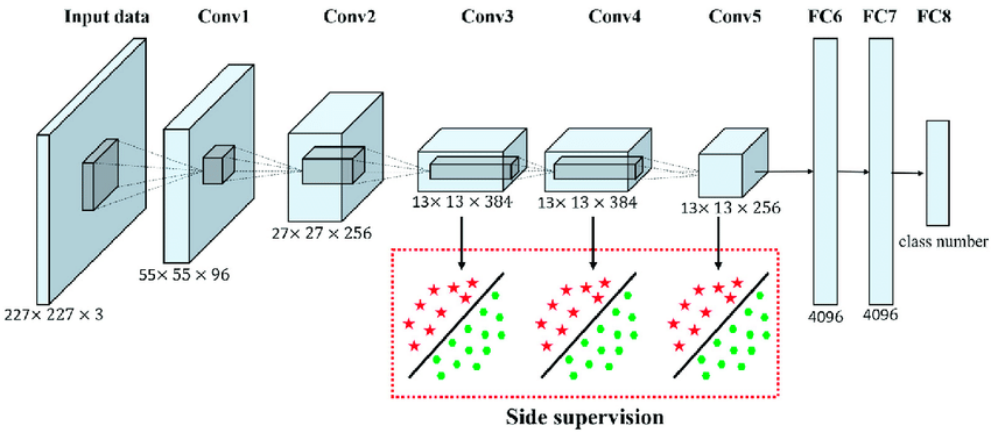
3.2.2.7.Çıkış Katmanı

Çıkış katmanı evrişimli sinir ağlarının işlemi tamamladığı katmandır. Bu katmanda sınıflandırma işleminin yapıldığı görüntüler bulunmaktadır. Sınıflandırılacak olan nesnelerin kaç farklı kategoriye ayrılacağını gösteren nesne adetince düğümler ile bulunmaktadır. En yaygın kullanılan sınıflandırıcı olan softmax, sınıflandırılacak her bir görüntü için kendi fonksiyonundan benzerlik yüzdesi alınır ve benzerlik yüzdesi en çok olan görüntünün sahip olduğu nesne, ağın tahmin ettiği nesne olduğu kabul edilmektedir.

3.3.3. Evrişimli sinir ağları mimarileri

AlexNet

AlexNet, 2012 yılında Alex Krizhevsky, Ilya Sutskever ve Geoffrey Hinton tarafından geliştirilen derin öğrenme algoritmasıdır. AlexNet, görüntü sınıflandırma görevi için kullanılan bir sinir ağıdır ve özellikle ImageNet veri kümesi üzerindeki görevlerde büyük başarı göstermiştir (Alom vd., 2018). AlexNet mimarisi Şekil 3.11'de gösterilmiştir.

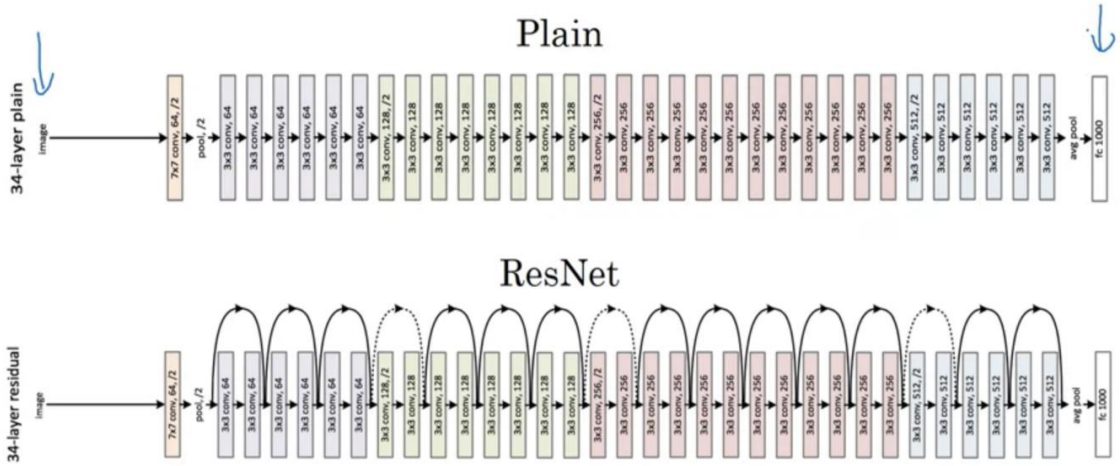


Şekil 3. 11. AlexNet mimarisi

AlexNet mimarisi, 8 katmanlı evrişimli sinir ağı kullanan bir yapıya sahiptir. Bu yapı, tekrarlanan bloklar açısından LeNet'e benzemektedir.

ResNet

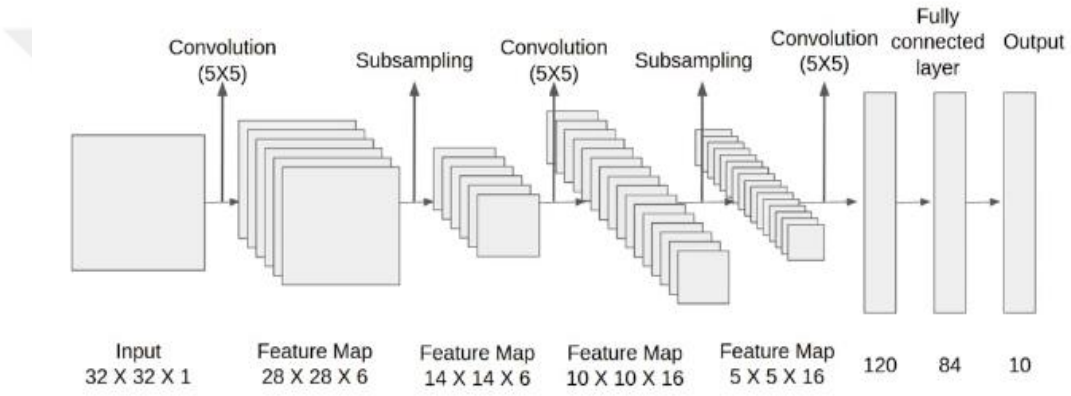
ImageNet yarışmasında Alexnet'in başarısı ile birlikte derin öğrenmede görüntü işleme, nesnelerin belirlenmesi gibi birçok alanda kullanımı hız kazanmıştır. Ağların derinliği arttıkça, yönetimi ve eğitimi de zorlaşmıştır. Bu sebeple eğitimi kolaylaştırmak için ağların optimize edilmesi oldukça önemlidir. ResNet bu bağlamda bir bloktaki değerlerle sonraki katmanları besleyen blokların eklenen yapısı ile önceki ağlara göre daha gelişmiş bir yapıdadır (Kızrak, 2018). ResNet 'in 18 ve 50, 152, 1022 olmak üzere farklı katman yapısına sahip iki farklı modeli vardır. İlk model düz ağ temelli iken, aynı çıktı özelliği haritası için katmanları aynı filtreye sahiptir. İkinci model ise özellik haritası yarıya indirildiğinde filtre sayısı iki katına çıkan bir modeldir. ResNet mimarisi aşağıdaki gibi gösterilebilir (Bkz: Şekil 3.7.).



Şekil 3. 12. ResNet mimarisi

LeNet

LeNet, 1998 yılında Yann LeCun ve arkadaşları tarafından geliştirilmiştir. İlk evrişimli yapay sinir ağı olarakta bilinen LeNet, yapay nöronları kapsama aralığında çevredeki hücrelerin bir kısmına yanıt verebilen ve büyük ölçekli görüntü işlemede iyi performans gösteren ileri beslemeli sinir ağıdır. LeCun banka çekleri üzerindeki sayıları tanımlamak amacıyla geliştirdiği bu ağa LeNet adını vermiştir (LeCun vd., 1998). Model, giriş verilerini evrişimli katmanlar, azaltma katmanları ve tam bağlantılı katmanlar gibi farklı katman türlerinde işleyerek sonuçları elde etmektedir (Baydan, 2021). Şekil 3.13'te LeNet mimarisi verilmiştir.



Şekil 3. 13. LeNet Mimarisi

LeNet, görüntü sınıflandırma, sayı tanıma ve optik karakter tanıma (OCR) gibi çeşitli görevler için kullanılmıştır.

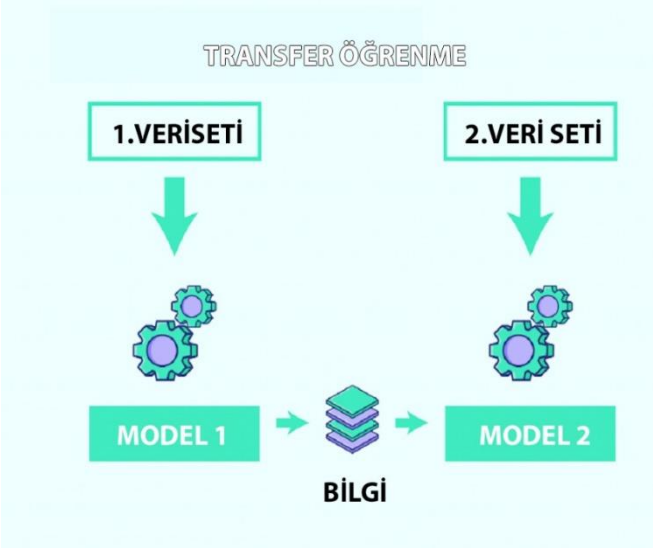
DenseNet

DenseNet, birbirine yoğun şekilde bağlı olan CNN katmanları ile bir araya gelmektedir. Katmanların çıkışları yoğun bir blokta bağlandığı DenseNet, 2017 yılında Gao, Huand ve diğerleri tarafından açıklanmıştır Araştırmacılar modelde yer alan ilk ve son katman arasındaki bağlantıların daha kısa olması gerektiği üzerine çalışma yapmıştır. DenseNet'in verimli bir yeteneği, ağ parametrelerini azaltarak özelliklerin yeniden kullanımıdır. DenseNet'in 40,100,121 ve 169 katmanlı türleri mevcuttur. Katmanların tümü önceki katmandaki özellik haritasını girdi olarak alırken, kendinden sonraki katmanlara da haritalarını girdi olarak göndermektedir (Kılıç, 2021).

3.2.4.Transfer Öğrenme

Transfer öğrenme, öğrenilen bilginin veya becerilerin bir görevden diğerine ya da bir alan veya bağlamdan diğerine aktarılmasını ifade eder. Bu öğrenme biçimi daha önce eğitilmiş bir ağ üzerinden elde edilen model bilgisinin yeni oluşturulan ağ modeline aktarılması prensibi ile çalışmaktadır (Geetharamani vd., 2019). İnsanlığın temellinde bu öğrenme yaklaşımının bulunduğunu görülmektedir. İnsan geçmiş tecrübelerinden yararlanıp karşısına çıkan olaylara farklı çözümler bulmaktadır. Gerçek hayatta ilk defa karşılaştığımız ve ani karar vermek zorunda kaldığımız anlar görülebilmektedir.

Bu model bilginin önceden eğitilmiş ağ modeli üzerinden yeni oluşturulan ağ modeline aktarma işlemi olarakta ifade edilebilir. Modelin temel amacı modellerin yeniden oluşturulması yerine önceden eğitilmiş ağı kullanıp model geliştirme süresini kısaltıp performansı geliştirmektir (Kaya vd., 2019). Transfer öğrenmenin bir amacı da kaynakla hedef arasındaki öğrenme becerilerini geliştirmektir. Aktarım yapılırken birtakım zorluklar gözlenmektedir. Kaynak ile hedef arasında bağlantı zayıf olduğunda negatif transfer oluşmaktadır. Bu nedenle kaynaktaki bilgilerin iyi analiz edilip aktarılması önemlidir. Transfer öğrenme görüntü sınıflandırılması alanında önemli başarılara sahiptir (Oquab *vd.*, 2014; You vd., 2015). Transfer derin öğrenme modeli Şekil 3.14'te görselleştirilmiştir.



Şekil 3. 14. Transfer Derin Öğrenme Modeli

3.2.5.ResNet

ResNet (Residual Neural Network), derin sinir ağlarının performansını arttıran ve derine indikçe ortaya çıkan sorunları çözebilen bir sinir ağı mimarisidir. ResNet, 2015 yılında Kaiming He ve arkadaşları tarafından geliştirilmiştir. Bu mimari, ImageNet veri kümesi üzerindeki büyük ölçekli görüntü sınıflandırmada başarılı sonuçlar elde etmiştir.

ResNet ağının amacı, performansı düşüren katmanları doğrudan atlayarak daha derin ağların öğrenmesini sağlamaktır. Genellikle sinir ağları, her katmanın girişini doğrudan bir sonraki katmana aktarmaktadır. Ancak bu durumda, ağın derinleşmesi ile birlikte ortaya çıkan "*öğrenme engeli*" veya "*değer kaybı*" sorunu ortaya çıkmaktadır. Derin ağlarda, geriye doğru yayılan gradyanlar giderek azalmakta, bu durum ağın daha önceki katmanlarda iyi bir şekilde öğrenememesine neden olmaktadır. Bu sorunu çözmek için "bağlantı (blok) atlamaları" veya "Residual blok" olarak adlandırılan özel bir yapı kullanılmaktadır. Her bir Residual blok, bir katmandaki çıktıyı bir sonraki katmana aktarmak yerine, çıktıyı bir sonraki katmana eklemektedir. Bu, ağın geleneksel bir derin sinir ağından ziyade bir "Residual blok" oluşturmasını sağlamaktadır. Residual bloklar, ağın daha derin katmanlarında oluşan gradyan kaybını önemli ölçüde azaltmaktadır. Eğer bir katman, öğrenmek için yeterince iyi bir optimizasyon gerçekleştirmezse, Residual bağlantılar sayesinde orijinal giriş verisine daha yakın bir dengeleme yapılabilmektedir. Bu sayede, ağın daha derin katmanlarında da iyi bir şekilde öğrenme gerçekleşebilmektedir.

ResNet, çeşitli derinliklere sahip varyasyonlara sahiptir. Örneğin; ResNet-50, ResNet-101, ResNet-152 gibi varyasyonlar daha fazla Residual bloğa sahiptir.

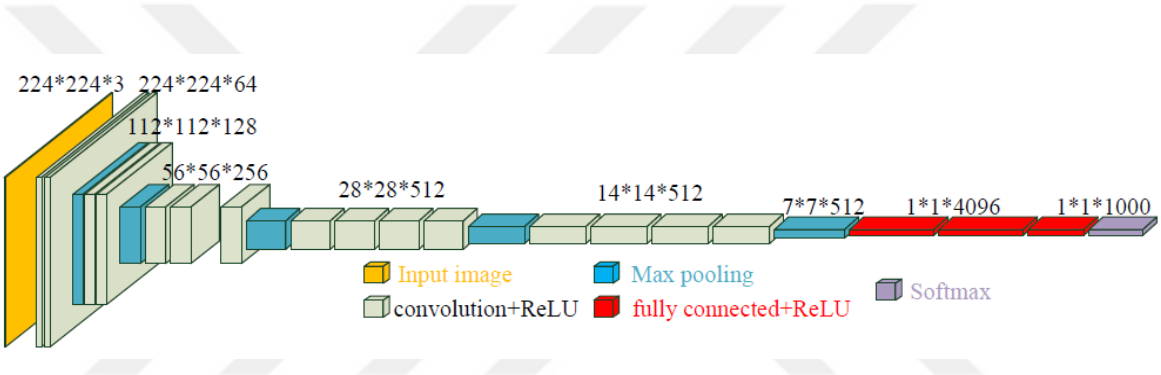
ResNet, anlamsal bölütlemeye uygulanabilen bir sinir ağı mimarisi olarakta kullanılabilmektedir. Özellik çıkarıcı olarak kullanılarak anlamsal bölütlemeyi gerçekleştiren bir kodlayıcı-kod çözücü yapısına sahip bir ağ ile birleştirilebilmektedir. Kodlayıcı, ResNet'in evrişimli katmanlarını içermekte ve görüntüyü daha küçük özellik haritalarına indirgemektedir. Kodlayıcı, daha derin katmanlarda daha genel özellikleri öğrenebilmektedir. Ardından kod çözücü ağ, indirgenmiş özellik haritalarını kullanarak orijinal görüntünün boyutuna geri dönmektedir. Kod çözücü, transpoze konvolüsyon katmanları ve atlamalı bağlantılar kullanarak yüksek çözünürlüklü özellik haritaları üretebilmektedir. Kod çözücü, her piksele sınıf etiketlerini tahmin eden bir çıktı katmanına bağlanmaktadır.

ResNet, çeşitli görüntü analizi ve anlama görevlerinde yüksek performans sağlayabilmektedir. Örneğin; nesne tespiti, görüntü segmentasyonu, yüz tanıma gibi

alanlarda ResNet tabanlı anlamsal bölütleme modelleri yaygın olarak kullanılmaktadır. Bu modeller, görüntü içerisindeki nesneleri tanımlayıp sınıflandırabilmektedir.

3.2.6.Vgg19

VGG-19, önceden eğitilmiş 19 katmanlı bir evrişimli sinir ağıdır ve bir görüntüyü sınıflandırmak için kullanılabilmektedir. Söz konusu 19 katmanın 16'sı evrişimli katmanlar ve geri kalan 3'ü tamamen bağlı katmanlardır (Jaworek *vd.,* 2019). Evrişimli katmanlar, özellikleri çıkarmak için görüntü üzerinde hareket eden filtrelerden oluşmakta ve her katmanda filtre sayısı artmaktadır. Tam bağlı katmanlar, evrişimli katmanlardan çıkarılan özellikleri sınıflandırmak için kullanılmaktadır. Şekil 3.15'te VGG-19 Ağı mimarisi verilmiştir.



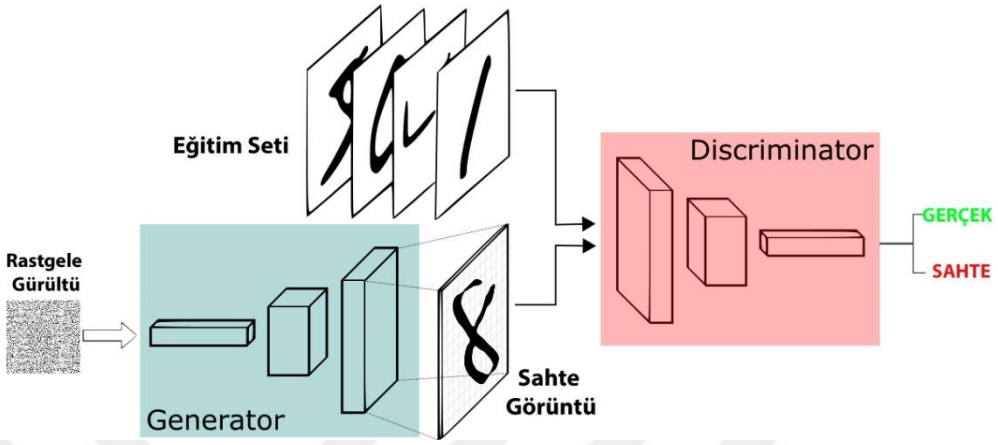
Şekil 3. 15. VGG-19 Ağı Mimarisi

VGG-19, tek bir evrişimden daha iyi olan, çoklu evrişim katmanların ve doğrusal olmayan etkinleştirme katmanlarının değişen bir yapısını kullanmaktadır. Katman yapısı, görüntü özelliklerini daha iyi çıkarabilmekte, alt örnekleme için maksimum havuzlama kullanabilmekte ve aktivasyon işlevi olarak doğrusal birimi (ReLU) değiştirebilmektedir. Bir başka ifade ile alanın havuzlanmış değer olarak görüntü alanındaki en büyük değeri seçmektedir.

3.2.7.Üretken Çekişmeli Ağlar

Üretken çekişmeli ağlar (Generative Adversarial Networks), iki farklı ve birbirine zıt ağ yapıların birleştirilmesi ile üretilmiş ağlardır. Ağlar arasında bir oyun teorisi kurularak Nash dengesine ulaştırılmaktadır (Goodfellow vd., 2014). Generative Adversarial Networks ana odak noktası, çoğunlukla sıfırdan görüntülerdir. Müzik, metin dâhil diğer alanlardan da veri üretmektir. Ancak uygulama kapsamı bundan çok daha büyüktür.

GAN'lar, bir üretici (generator) ağı ve bir ayırt edici (Discriminator) ağı olarak adlandırılan iki sinir ağından oluşmaktadır. GAN'da işlenecek olan bir veri setinin geçtiği adımlar aşağıdaki şekilde özetlenebilir (Bkz: Şekil 3.16.).



Şekil 3. 16. Üretken çekişmeli ağlar örnek eğitim seti işlenişi

Şekil 3.16'da eğitim setinde gerçek ve sahte görüntüyü ayırt etmek için, üretici rastgele resim karelerinden gerçeğe benzer resimler üretmeye başlamaktadır. Eğitim seti içerisindeki gerçek resim ile üretici 'den üretilen gerçeğe yakın resim ayırt ediciye girdi olarak verilmekte ve alınan girdilerin sahte mi yoksa gerçek mi olduğunu sonuç olarak göstermektedir.

Kendinden önceki her bir düğümün değerleri toplanarak geriye yayılma metodu ile gradient'ler hesaplanmaktadır. Bu şekilde rastgele resim karelerinden üretilen resimlerin gerçeğe yakın olanları ile aralarındaki fark hesaplanmaktadır. Bu durumda, üretici ağ her bir düğümden sonra kendini güncelleyip gerçek olan resme yakın resimler üretmemize olanak sağlamaktadır. Modelin ilerleyen düğümlerinde üretici ağ, gerçek olan resimlere daha benzer resimler üretirken, ayırt edici ağ ise gerçek ve sahte resimleri anlamakta zorlanmaktadır.

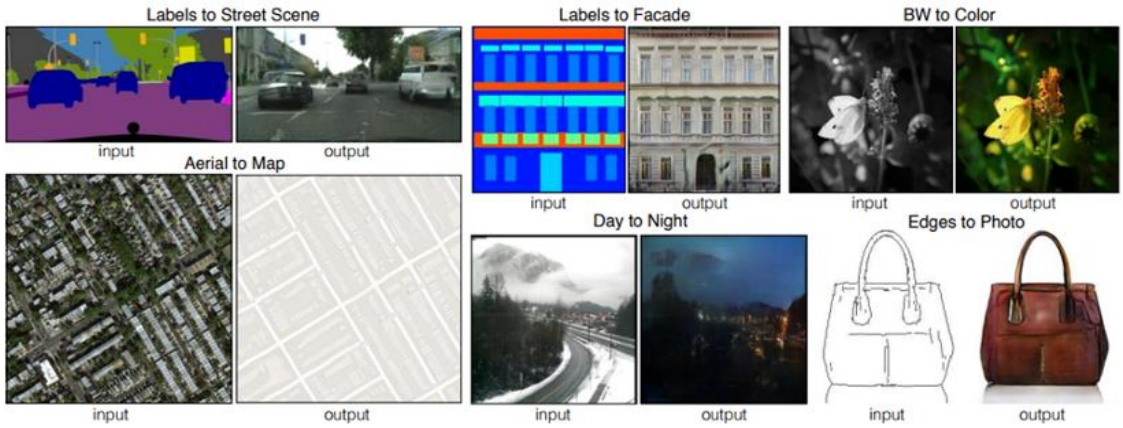
Üretici ağın amacı, ayırt edici ağın hata yapmasını sağlayacak resimler üretmek iken; ayırt edici ağın amacı, gerçek resimleri sahte olanlardan daha iyi ayırt etmektir. Ayırt edici tıpkı bir derin ağ sınıflandırıcısı gibi eğitilmelidir. Girdi gerçek ise, (x)=1 olmalıdır. Oluşturulursa, (x)=0 olmalıdır. Bu süreç boyunca, ayırt edici gerçek görüntülere katkıda bulunan özellikleri tanımlar. Öte yandan, üretici (x) = 1 (gerçek görüntüyle eşleşen) ile görüntüler oluşturulması beklenmektedir. Böylece, bu hedef

değeri geri yayılımla üreteci eğitilebilir, yani üreticinin ayırt ediciye gerçek olduğunu düşündüğü şeye doğru görüntüler oluşturmak için eğitilmelidir.

Üretken çekişmeli ağlar ile yapılabilecek bazı şeyler aşağıda sıralanmaktadır.

- Birden fazla türde çiçek resmini, GAN'a girdi olarak verip eğittikten sonra önceden hiç görmediğimiz türde çiçek resimlerine sahip olunabilmektedir.
- GAN modeline şiirlerden alınan kesitler verilerek yeni bir şiir kesiti oluşturmasını sağlayabilmektedir.
- Günümüz popüler modellerinden biri olan pix2pixHD ile yapılan birkaç çalışma, siyah beyazdan renkli resme çevirme, kurşun kalem ile çizilen resmi gerçeğine yakın boyama, gündüz çekilen fotoğrafları geceye veya gece çekilen fotoğrafları gündüze çevirme ve uydudan alınan görüntüleri haritaya dönüştürme gibi.

Şekil 3.17'de Üretken çekişmeli ağlar modelinin ürettiği örnek görüntüler verilmiştir.

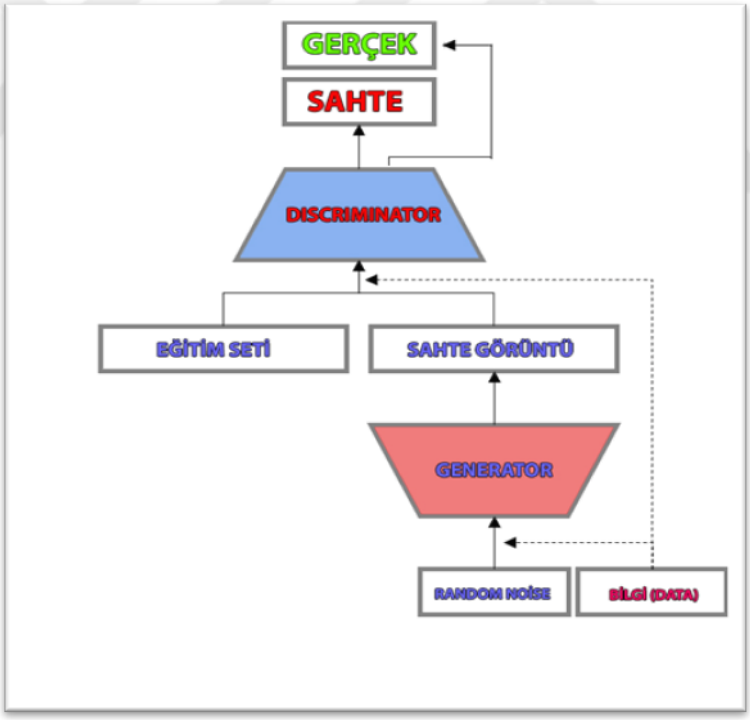


Şekil 3. 17. Generative Adversarial Network modelinden örnek görüntüler

3.2.8. Koşullu Üretken Çekişmeli Ağlar

Koşullu üretken çekişmeli ağlar (Conditional Generative Adversarial Network), süper çözünürlük, görüntü içi boyama ve stil aktarımı gibi görüntü işleme alanında büyük ilerleme kaydetmiştir. GAN modelinden farklı olarak koşul parametresi ekleyerek, üretilen verinin bazı karakteristiklerini kontrol etmeyi amaçlayan bir ağdır. Hem üretici ağa hem de ayırt edici ağa ekstra bir girdi verisi ekleyerek, üretilen verinin belirtilen koşula uyması sağlanmaktadır.

GAN modelinin çıktısında rastgele bir şekilde gerçekçi veriler üretmek amaçlanmaktadır. CGAN 'da ise üretici ve ayırt edici derin ağlarının girişine bilgi verilmesi gerekmektedir. Üretici, verilen bilgi doğrultusunda gerçek veya sahte görüntünün hangi sınıf ve etikete ait olduğunu bulabilecektir. Ayırt edici, resimlerin sınıf bilgilerine ulaşarak gürültü verisiyle birlikte resim üretme işlemine başlayacaktır. CGAN'da işlenecek olan bir veri setinin geçtiği adımlar aşağıdaki şekilde özetlenebilir (Bkz: Şekil 3.18.).



Şekil 3. 18. Koşullu üretken çekişmeli ağlar örnek eğitim seti işlenişi

Koşullu üretken çekişmeli ağlar, görüntü sınıflandırma, görüntü etiketleme ve sentetik veri oluşturma gibi alanlarda kullanılmaktadır. Örneğin, bir CGAN modeli, belirli bir etikete göre görüntüleri üretebilmekte veya belirli bir sınıfa ait görüntülerin sentetik versiyonlarını oluşturabilmektedir.

3.2.9. Pix2PixHD

Pix2PixHD, yüksek çözünürlüklü görüntülerin düşük çözünürlüklü eşlemelerinden gerçekçi ve ayrıntılı görüntüler üretmek için kullanılan bir görüntü çeviri yöntemidir. Bu yöntemin temelinde, derin öğrenme ve özellikle de evrişimli sinir ağları bulunmaktadır.

Pix2PixHD, yüksek çözünürlüklü görüntü sentezlemek için bir eğitim veri kümesine dayanmaktadır. Bu eğitim veri kümesi, düşük çözünürlüklü giriş görüntülerini yüksek çözünürlüklü hedef görüntülerine eşleştiren çiftlerden oluşmaktadır. Yöntemin temel çalışma mantığı, giriş görüntüsünü alır ve onu yüksek çözünürlüklü bir çıktıya dönüştürmek için birçok evrişimli sinir ağı katmanı kullanılmaktadır. Bu katmanlar, giriş görüntüsündeki özellikleri yakalar ve yüksek çözünürlüklü çıktıyı oluşturmak için bu özellikleri mevcuttur. Pix2PixHD, evrişimli sinir ağları, toplu normalizasyon, residual bağlantılar gibi teknikleri kullanılmaktadır. Bu yönteme, fotoğraf restorasyonu, stilde transfer, renklendirmeli görüntü sentezi gibi örnekler verilebilmektedir. Mevcut tez çalışmasında, pix2pixHD yöntemi kullanılarak görüntü içerisindeki nesneleri (yaya, yol, araba vb.) üretme işlemi gerçekleştirilmiştir.

4. DENEYSEL BULGULAR

4.1. Anlamsal Bölütleme İşlemi Sonuçları

Araştırmanın bu bölümünde sokak görüntülerine anlamsal bölütleme işlemi uygulanarak görüntülerdeki nesneler sınıflandırılmıştır. Görüntü temel olarak piksel denilen çeşitli tonlara bağlı olarak sayısal değerlerden oluşur. Görüntüler kullanım alanlarına göre belirli işlemlerden geçmesi gerekebilir. Bölütleme bu işlemlerden biridir. Bölütleme işlemi görüntü üzerindeki piksellerin belirli gruplara/sınıflara ayrılma işlemidir. Temel olarak iki tür bölütleme işlemi bulunmaktadır. Birey (Instance) bölütleme işleminde her nesne ayrı sınıflandırıldığı gibi benzer nesnelerde farklı etiketler ile sınıflandırılmaktadır. Örneğin; bir görüntüdeki 3 ağaca ayrı ayrı sınıf etiketi atanmaktadır. Bir diğer bölütleme yöntemi anlamsal bölütleme (Semantic Segmantation) yönteminde belirli nesnelere ait piksellere aynı etiket atanmaktadır. Görüntüdeki 3 ağaç aynı sınıf bilgisi ile etiketlenmektedir. Mevcut tez çalışmasında görüntüler derin öğrenme ile anlamsal bölütleme işlemi gerçekleştirilmiştir. Anlamsal bölütleme otonom araçlarda, tıbbi teşhis alanlarında, uzay bilimlerinde, ziraat vs gibi çeşitli alanlarda yaygın bir şekilde kullanılmaktadır.

Derin öğrenme ile ilgili olarak çeşitli CNN tabanlı yöntemlerle anlamsal bölütleme çalışmaları yapılmıştır. Yaygın bir şekilde kullanılanlar: Deeplab v3+, FCN (Fully convolutional networks), SegNet ve U-Net yöntemleridir. Mevcut çalışmada daha önceden eğitilmiş ResNet-18 transfer derin öğrenme metodu ile DeepLab v3+ ağı oluşturularak anlamsal bölütleme işlemi gerçekleştirilmiştir. ResNet-18, sınırlı sayıda görüntüye sahip uygulamalar için uygun bir transfer derin öğrenme ağıdır. DeepLab v3+ ağında diğer transfer derin öğrenme metotları da kullanılabilmektedir. Yaklaşımı test etmek için Cambridge Üniversitesi'nden CamVid veri seti kullanılmıştır. Bu veri seti, sürüş sırasında elde edilen sokak görüntülerinden oluşmaktadır. Veri seti, araba, yaya ve yol dâhil olmak üzere 32 anlamsal sınıf için piksel düzeyinde etiketlenmeler yapılmıştır. Anlamsal bölütleme işlemi Matlab yazılımı ile gerçekleştirilmiştir. Deeplab v3+ ağına ait eğitim parametreleri Çizelge 4.1’de verilmiştir.

Çizelge 4. 1. Deeplab v3+ ağına ait eğitim parametreleri

Parametre	Değer
Optimizasyon Algoritması	Sgdm
LearnRateDropPeriod	10
Momentum	0.9
InitialLearnRate	1e-3
L2Regularization	0.005
MaxEpochs	30
MiniBatchSize	8

Veri setinde 32 sınıf bilgisi bulunmaktadır. Ancak mevcut çalışmada bir kısım etiketler birleştirilerek tek etiketle ifade edilmiştir. Örneğin; tren, otobüs, araba ve diğer hareket eden taşıtlar araba etiketi ile belirtilmiştir. Bu 11 etiket {gökyüzü, bina, direk, yol, kaldırım, ağaç, işaret/ sembol, çit, araba, yaya ve bisiklet} olarak belirtilmiştir.

Şekil 4.1'de bir görüntünün anlamsal bölütlenme işlemine örnek verilmiştir.

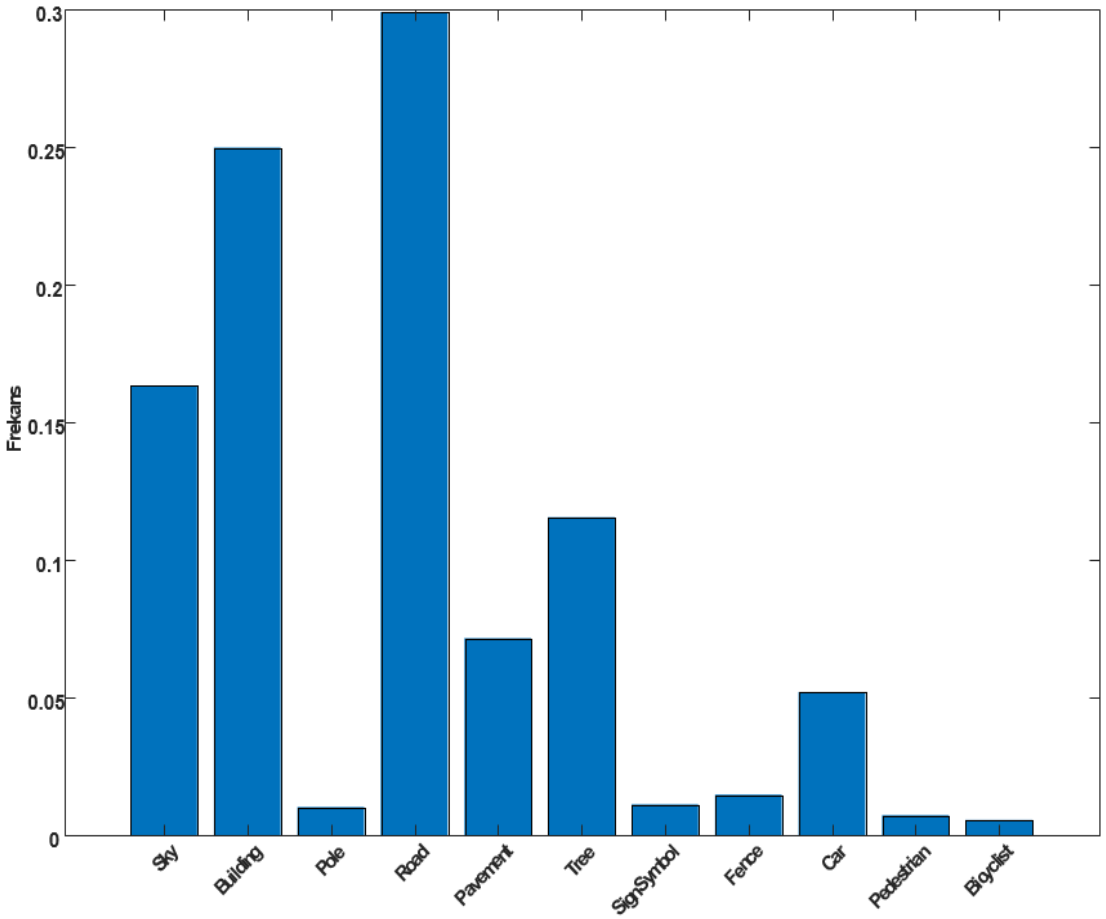
Şekil 4. 1. Örnek bir görüntünün anlamsal bölütlenmesi

Anlamsal bölütleme işleminden sonra elde edilen görüntüdeki her etiketin piksel dağılımı ve tüm veri setindeki piksel dağılımları Çizelge 4.2'de gösterilmiştir.

Çizelge 4. 2. CamVid veri setindeki etiketlerin piksel dağılımları

Etiket	Bulunan Piksel Sayısı	Resim Piksel Sayısı
Gökyüzü (Sky)	7.6801e+07	4.8315e+08
Bina (Building)	1.1737e+08	4.8315e+08
Direk (Pole)	4.7987e+06	4.8315e+08
Yol (Road)	1.4054e+08	4.8453e+08
Kaldırım (Pavement)	3.3614e+07	4.7209e+08
Ağaç (Tree)	5.4259e+07	4.479e+08
İşaret/Sembol (Sign)	5.2242e+06	4.6863e+08
Çit (Fence)	6.9211e+06	2.516e+08
Araba (Car)	2.4437e+07	4.8315e+08
Yaya (Padestrian)	3.4029e+06	4.4444e+08
Bisiklet (Bicycle)	2.5912e+06	2.6196e+08

Tüm veri setindeki 11 etikete ait piksel dağılımları Çizelge 4.2'de verilmiştir. Değerler normalize edildikten sonra gösterilmiştir.



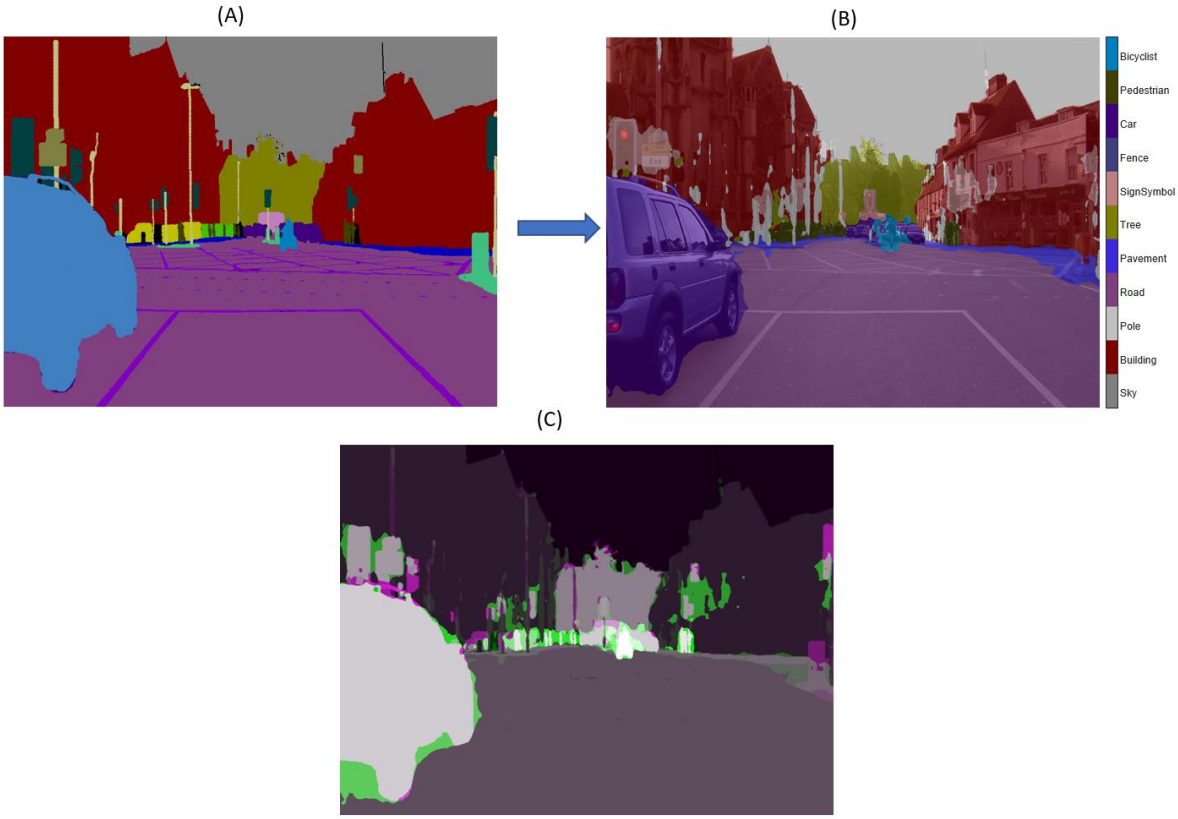
Şekil 4. 2. CamVid veri setinde 11 etiketin piksel dağılımları

DeepLab v3+ transfer derin öğrenme metodunun eğitim sürecinden sonra test veri setinden rastgele resimler seçilerek denemeler gerçekleştirilmiştir. Örnek 2 denemeye ait sonuçlar aşağıda verilmiştir. İlk örneğimizde gerçek görüntü ve anlamsal bölütleme işleminden sonra elde edilen görüntü Şekil 4.3'te verilmiştir.

Şekil 4.3'e bakıldığında gökyüzünün, yol, direk, ağaç gibi nesnelerin doğru tespit edildiği ve etiketlediği görülmektedir. Ancak daha küçük yaya gibi nesnelerin bölütlenmesi daha zor olmaktadır. Bölütleme işleminin başarısını ölçmek için gerçek etiket bölütlenmiş görüntü ile transfer derin öğrenme metodun anlamsal bölütleme görüntülerinin karşılaştırılması gerekmektedir.

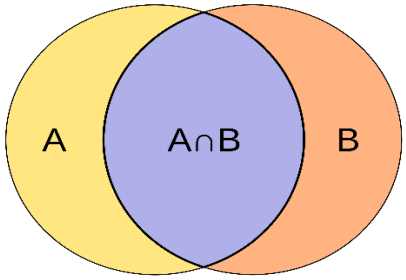
Şekil 4. 3. Örnek bir görüntü için anlamsal bölütleme

Gretag–Macbeth renk şemasına göre bölütlenmiş görüntü ile derin öğrenme metodu ile bölütlenmiş görüntüler aşağıda verilmiştir (Bkz: Şekil 4.4). Etiket/sınıf başına başarıyı veya örtüşmeyi ölçmek için birleşim üzerinden kesişme (Intersection Over Union, IoU) metriği hesaplanmalıdır. Bu metrik Jaccard indeksi olarak da bilinmektedir. IoU metriği gökyüzü, yol araba gibi büyük nesneler/etiketler için yüksek değerler elde edilmişken yaya, bisiklet gibi küçük nesneler için küçük değerler ölçülmüştür.



Şekil 4. 4. Bir test görüntüsüne ait (A) Gretag–Macbeth ile bölütlenmiş etiket, (B) derin öğrenme ile anlamsal bölütlenmiş görüntü, (C) A ve B görüntülerinin kesişimi

Tablo 4.3'te örnek görüntünün anlamsal bölütleme başarı metrikleri verilmiştir. Metrikler Şekil 4.4 (A) ve Şekil 4.4 (B) görüntülerin örtüşen piksel oranlarını belirtmektedir. Jaccard benzerlik indeksi iki küme arasındaki benzerliği ifade etmektedir. Tahminler ve etiketler arasında benzerliği ifade etmek için bölütleme işlemlerinde yaygın bir şekilde tercih edilmektedir. Söz konusu benzerlik indeksi Şekil 4.5'te gösterilen iki kümenin küme kesişimi ve birleşimi bilgilerinden Eşitlik (4.1)'e göre hesaplanmaktadır.



Şekil 4. 5. Küme Kesişim Gösterimi

$$IoU = \frac{|A \cap B|}{|A| \cup |B|} \quad (4.1)$$

Bir diğer ölçüt Sørensen-Dice katsayısıdır. İki kümenin benzerliğini ölçmek için kullanılan bir istatistiktir. İndeks, özellikle Sørensen–Dice indeksi, Sørensen indeksi ve Dice katsayısı olmak üzere birçok başka adla bilinir. Eşitlik (4.2) ile hesaplanır.

$$DSC = \frac{2|A \cap B|}{|A| + |B|} \quad (4.2)$$

Çalışmada son olarak kullandığımız ölçüt BF metriğidir. Bu metrik aşağıdaki gibi hesaplamaktadır.

$$P^c = \frac{1}{|B_{ps}^c|} \sum_{x \in B_{ps}^c} \left[\left[d\left(x, B_{gt}^c\right) < \partial \right] \right] \quad (4.3)$$

$$R^c = \frac{1}{\left|B_{gt}^c\right|} \sum_{x \in B_{gt}^c} \left[\left[d\left(x, B_{ps}^c\right) < \partial \right] \right] \quad (4.4)$$

$$BF^c = \frac{2.P^c.R^c}{P^c + R^c} \quad (4.5)$$

Burada;

B_{gt}^c: İkili görüntüdeki c sınıfın sınırlarını göstermektedir.

B_{ps}^c: Tahmin edilen görüntüdeki piksel görüntüdeki c sınıfın sınırlarını belirtir.

∂ : Uzaklık eşik değeri olarak belirtilmektedir.

P^c Ve R^c : C sınıfı için sırası ile kesinlik ve duyarlılık değerlerini ifade etmektedir.

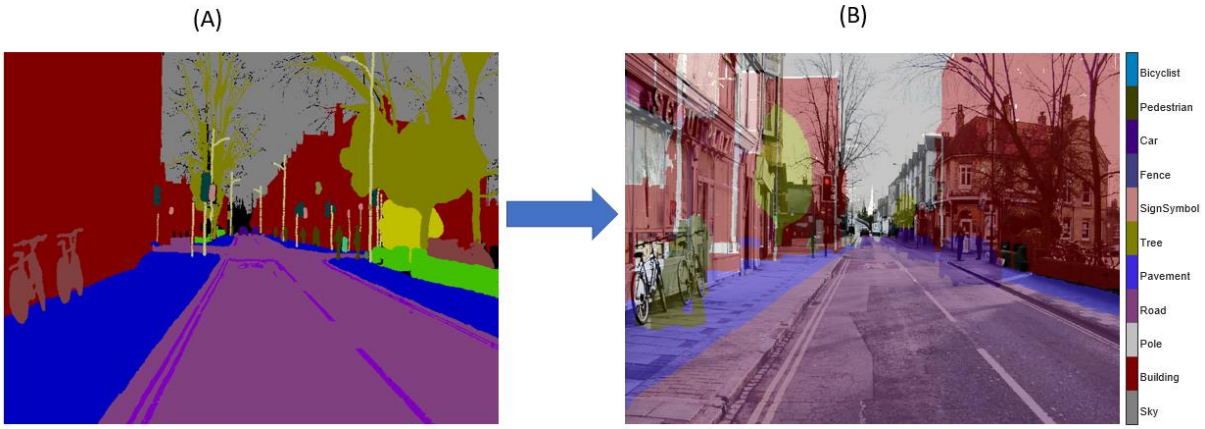
Örnek görüntü için belirlenen sınıflara ait performans değerleri Çizelge 4.3'te verilmiştir. Çizelgeye göre görüntüde geniş bir bölgeye sahip sınıfların performans değerleri yüksek elde edilmiştir. Daha küçük nesneler için bu ölçütler düşük kalmıştır.

Çizelge 4. 3. Örnek 1 görüntü için performans değerleri

Etiket	Jaccard Indeksi	Sørensen-Dice	BF Skoru
Gökyüzü (Sky)	0.90979	0.94272	0.90853
Bina (Building)	0.7916	0.81488	0.63963
Direk (Pole)	0.24632	0.75997	0.58505
Yol (Road)	0.92638	0.93955	0.80615
Kaldırım (Pavement)	0.73874	0.90048	0.74538
Ağaç (Tree)	0.7746	0.88173	0.72892
İşaret/Sembol (Sign)	0.42338	0.76491	0.53707
Çit (Fence)	0.57442	0.83661	0.5567
Araba (Car)	0.79441	0.79441	0.74331
Yaya (Padestrian)	0.47077	0.47077	0.64357
Bisiklet (Bicycle)	0.64779	0.64779	0.59473

İkinci bir deneme için farklı bir görüntü kullanılmıştır. İkinci görüntüde ağaç, yol, gökyüzü, bina, kaldırım gibi nesneler ön plandadır. Derin öğrenme ile elde edilen anlamsal bölütleme görüntüsü ve gerçek görüntü Şekil 4.6'da verilmiştir. Anlamsal bölütleme sonucu elde edilen görüntü ve etiket görüntü de Şekil 4.7'de verilmiştir. Anlamsal bölütleme sonucu elde edilen görüntü ve etiket görüntü arasındaki benzerlik ölçütleri her sınıf için ayrı ayrı olarak Çizelge 4.4'te verilmiştir.

Şekil 4. 6. Örnek bir görüntü için anlamsal bölütleme



Şekil 4. 7. Bir test görüntüsüne ait (A) Gretag–Macbeth ile bölütlenmiş etiket, (B) derin öğrenme ile anlamsal bölütlenmiş görüntü

Çizelge 4. 4. Örnek 1 görüntü için performans değerleri

Etiket	Jaccard Indeksi	Sørensen-Dice	BF Skoru
Gökyüzü (Sky)	0.90979	0.94272	0.90853
Bina (Building)	0.7916	0.81488	0.63963
Direk (Pole)	0.24632	0.75997	0.58505
Yol (Road)	0.92638	0.93955	0.80615
Kaldırım (Pavement)	0.73874	0.90048	0.74538
Ağaç (Tree)	0.7746	0.88173	0.72892
İşaret/Sembol (Sign)	0.42338	0.76491	0.53707
Çit (Fence)	0.57442	0.83661	0.5567
Araba (Car)	0.79441	0.92588	0.74331
Yaya (Padestrian)	0.47077	0.86718	0.64357
Bisiklet (Bicycle)	0.64779	0.88881	0.59473

4.2. Derin Öğrenme ile Bölütleme Haritasından Sentetik Görüntü Oluşturma

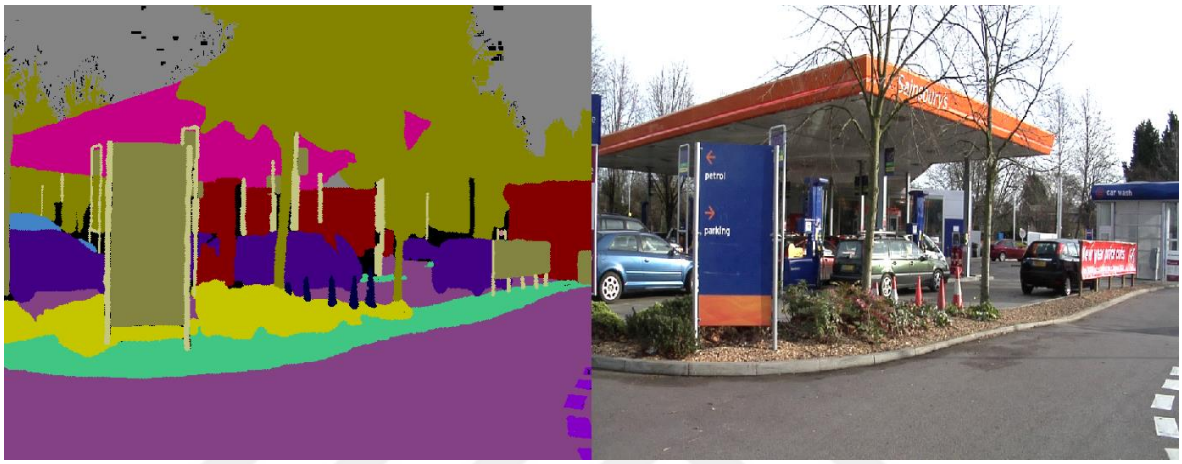
Mevcut bölümde anlamsal olarak bölütlenmiş haritalardan derin öğrenme ile yüksek çözünürlüklü sentetik görüntüler oluşturulmuştur. Bu tür yaklaşımların yaygın uygulama alanları bulunmaktadır. Örneğin; oyun veya film senaryoları için önce piksel görüntüleri oluşturduktan sonra derin öğrenme yaklaşımları ile piksel görüntüler üzerinden gerçekçi görüntüler oluşturmak daha kolay olacaktır. Anlamsal bölütleme yöntemlerini kullanarak, görüntüleri anlamsal bir etiket alanına dönüştürebilir, etiket alanındaki nesneleri düzenleyebilir ve ardından onları tekrar görüntü alanına dönüştürebiliriz. Bu yaklaşımlar görüntülere nesneler eklemek veya mevcut nesnelerin görünümünü değiştirmek gibi bize daha üst düzey görüntü düzenleme için yeni araçlar sunmaktadır. Bunun için koşullu üretken çekişmeli ağlardan pix2pixHD yöntemi kullanılmıştır. Bu yöntem yüksek çözünürlüklü ince ayrıntılar ve dokulara sahip görüntüler sentezlemek için kullanılmaktadır. Araştırmanın bu bölümünde görüntüden-görüntü oluşturmak için çekişmeli ağlar kullanılmıştır. Koşullu üretici ağlar görüntü-görüntü çevirisi uygulamalarda başarılı olmaktadır. Pix2PixHD görüntüden görüntüye çeviri için kullanılan çekişmeli bir ağdır.

Pix2pixHD yöntemi bir çekişmeli bir yöntem olduğundan dolayı eğitilmiş iki ağdan oluşmaktadır. İlk ağ üretici (Generator) anlamsal bölütleme haritasından bir sentetik görüntü oluşturacak kodlayıcı-kod çözücü tarzında bir ağdır. Üretken çekişmeli ağ üretici ağı görüntü oluşturmak için eğitir. Ancak bu görüntünün diğer ağ olan ayrıştırıcı (discriminator) ağı tarafından doğru olarak yanlış sınıflandırılması gerekmektedir. Ayrıştırıcı ağ ise oluşturulan sentetik görüntü ile karşılık gelen görüntüyü karşılaştırıp gerçek ve sahte şeklinde sınıflandıran, tamamıyla evrişimli bir ağdır. Üretici ve ayrıştırıcı ağlar eğitim sürecinde birbirleri ile rekabet ederler. Üretici ağın amacı, anlamsal piksel haritalarını gerçekçi görünen görüntülere çevirmektir. Ayrıştırıcı ağın amacı ise gerçek görüntüleri üretilen görüntülerden ayırmaktır. Bu bölümde sentetik görüntü oluşturmak için Pix2pixHD ağı CamVid veri seti ile eğitilmiştir.

Araştırmadaki sentetik görüntüler aşağıdaki aşamalara göre oluşturulmuştur.

1. **Piksel Etiketleme**

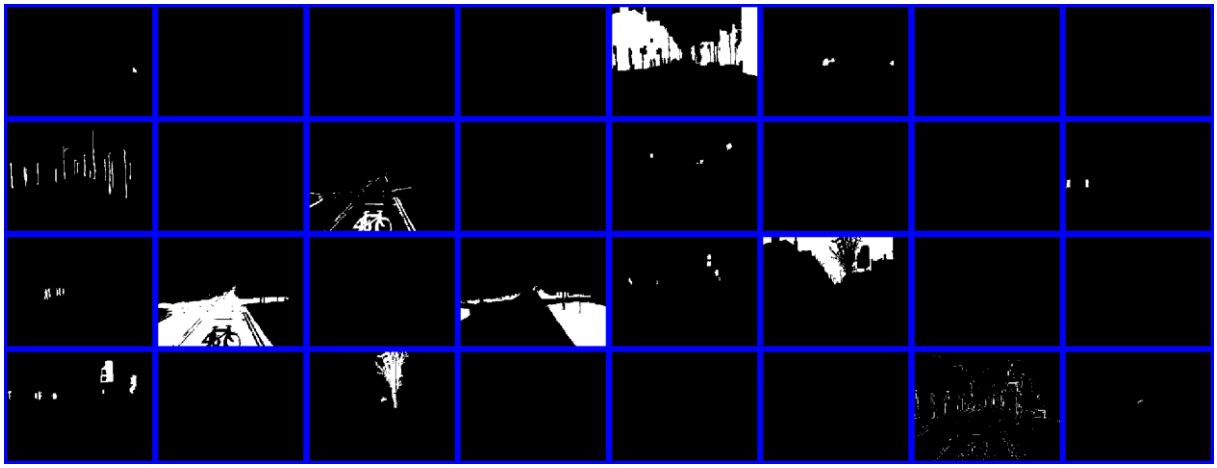
Veri setindeki her sınıf için piksel etiket bilgileri belirtilmiştir. Her etiketin piksel renk bilgisi piksel etiket olarak alınmıştır. Piksel etiketler ve karşılık gelen görüntüler haritalanmıştır. Örnek bir piksel etiket ve karşılık gelen gerçek görüntü aşağıda verilmiştir (Bkz: Şekil 4.8.).



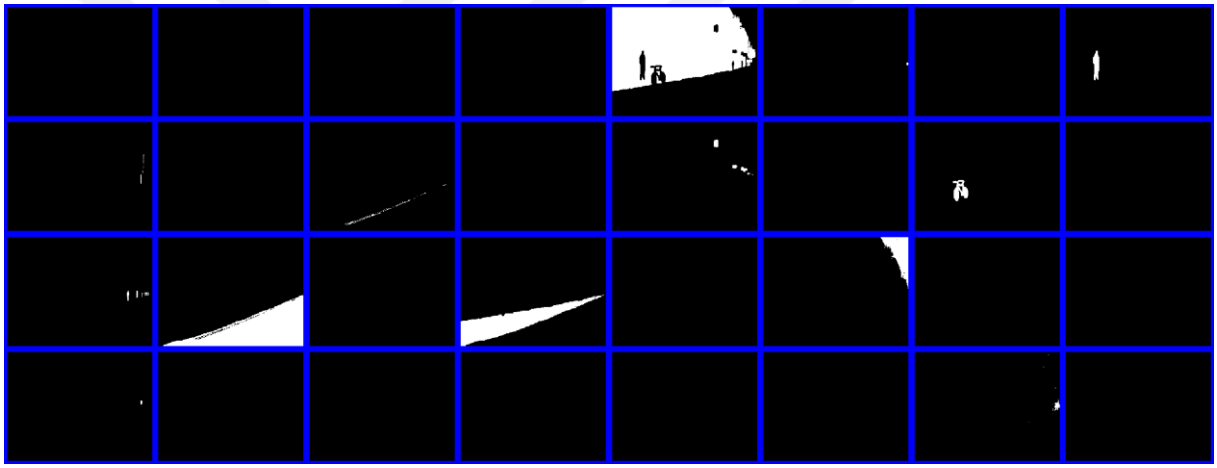
Şekil 4. 8. Örnek bir piksel etiket ve gerçek görüntüsü

Veri seti %70-%30 eğitim test oranlarına bölünmüştür. Veri setinin boyutu yetersizliğinden dolayı çoğaltma (data augmentation) gerçekleştirilmiştir. Araştırma için yetersiz veri olduğunda çeşitli teknikler ile sentetik veri oluşturma veya çoğaltma işlemine data augmentation (veri artırma) denilmektedir. Derin öğrenme modellerin başarısı veri boyutuna bağlıdır. Aksi takdirde ezberleme (overfitting) problemleri gerçekleşmektedir. Bu gibi ezberleme durumlarla karşılaşmamak ve araştırma modelini iyi bir şekilde öğrenip, performansını artırmak için veri artırma yöntemlerine başvurulur. Görüntüler [-1,1] aralığında ölçeklendirilerek, etiket ve görüntülerin boyutu değişmeden bicubic ve K-nn downsampling yöntemleri ile boyut değiştirme, tek kanallı bölütleme haritasını 32 kanallı bölütleme haritasına dönüştürme ve etiket-gerçek görüntü çiftlerini yatay yönde rastgele çevirme işlemleri ile veri çoğaltma yapılmıştır.

Eğitim ve test setleri için piksel etiketlerin tek kanallı bölütlenmiş ya da 32 kanallı bölütleme işleminden sonra her sınıf için one-hot haritası Şekil 4.9 ve Şekil 4.10'da verilmiştir. Her kanal her sınıf için piksel değerlerine karşılık gelmektedir.



Şekil 4. 9. Eğitim veri seti için One-Hot kodlama bölütleme kanallarına ait görünüm



Şekil 4. 10. Test veri seti için One-Hot kodlama bölütleme kanallarına ait görünüm

2.Üretici Ağın Oluşturulması

Bu aşamada one-hot bölütleme haritasından görüntüler oluşturacak pix2pixHD üretici ağı tanımlanmıştır. Girişleri orijinal segmentasyon haritasıyla aynı yükseklik ve genişlikte ve sınıflarla aynı sayıda kanaldan oluşmaktadır. Mevcut çalışmada pix2pixHD üretici ağı 576x768 boyutlarında görüntüler oluşturulacaktır. Yüksek çözünürlükte görüntüler de oluşturulabilir.

3.Ayrıştırıcı Ağın Oluşturulması

Araştırmanın bu bölümünde ayrıştırma işlemini gerçekleştiren GAN kısmı oluşturulmuştur. Bu bölüm giriş olarak verilen görüntüyü sahte (0) veya gerçek (1) olarak sınıflandırmaktadır. Ele aldığımız GAN ayrıştırıcı çok ölçekli görüntüleri bile ayrıştırabildiğinden dolayı çok ölçekli ayrıştırıcı olarak ifade edilir. Bunun için iki ayrıştırıcı ağ kullanılmaktadır. İlk ölçek, giriş görüntüsü boyutu ile aynı boyuttadır ve ikinci ölçek, görüntü boyutunun yarısı kadardır. Ayrıştırıcı ağa giriş olarak One-hot bölütlenmiş haritalar ve sınıflandırılacak görüntü birlikte verilir. Ayrıştırıcı giriş kanal sayısı toplam sınıf sayısı ve giriş görüntü kanal sayısı toplamına eşittir.

4.Model Kayıp Fonksiyonları

Üretici Ağ için Kayıplar

Model kayıp fonksiyonları üretici ağ için VGG derin transfer öğrenme yöntemi ile öznitelikler uyum kayıplarını hesaplamaktadır. Üretici ağın amacı üretilen sentetik görüntünün ayırıcı ağ tarafından gerçek olarak sınıflandırmasıdır. Üretici ağın kayıpları üç (3) tanedir. Birincisi bir (1) değeri ile üretilmiş resim için ayırıcı tahmin değerleri $\hat{Y}_{\text{üretici}}$ arasındaki farkın karesidir. Aşağı eşitlikte gösterilmiştir.

$$Adversarial\ Kayıp\ (L1) = (1 - \hat{Y}_{\text{üretici}})^2 \tag{4.6}$$

İkinci kayıp fonksiyonu öznitelik eşleştirme kayıp değerini ölçer. Bu değer ayırma ağından tahminler olarak elde edilen üretilmiş ve gerçek öznitelik haritaları arasındaki uzaklığı belirtir. T, $Y_{gerçek}$, $Y_{\text{üretilmiş}}$ sırası ağdaki ayırıcı öznitelik katman sayısı, gerçek görüntü, üretilmiş görüntü olmak üzere bu kayıp değeri aşağıdaki eşitlikle hesaplanır.

$$\text{Öznitelik Eşleştirme Kayıp}\ (L2) = \sum_{i=1}^{T} ||Y_{gerçek} - Y_{\text{üretilmiş}}|| \tag{4.7}$$

Son kayıp fonksiyonu algısal kayıp değerini belirtir. VGG transfer derin öğrenme ağının öznitelik çıkarım katmanlarından tahmin olarak elde edilen özniteliklerle gerçek öznitelikler arasında uzaklığı belirtir. Aşağıdaki eşitlikle bu kayıp hesaplanmaktadır. VGG transfer derin öğrenme yöntemi gerçek ve üretilmiş görüntülerde farklı katmanlardan öznitelik çıkarımı için kullanılmıştır.

$$Vgg\ Kayıp\ (L3) = \sum_{i=1}^{T} ||Y_{VGGgerçek} - Y_{VGGüretilmiş}|| \quad (4.8)$$

Üretici ağın genel kaybı yukarıdaki 3 kayıp değerin ağırlıklı değeri olarak hesaplanmaktadır. λ1, λ2 ve λ3 ağırlık faktörleri olup genel kayıp aşağıdaki eşitlikle hesaplanmaktadır.

$$Generator\ Kayıp\ = L1 * \lambda_1 + L2 * \lambda_2 + L3 * \lambda_3 \quad (4.9)$$

Ayırıcı Ağ için Kayıplar

Ayırıcı ağın amacı, gerçek görüntüler ve oluşturulan görüntüler arasında doğru bir şekilde ayrım yapmaktır. Ayırıcı ağın kayıp değeri iki bileşenin toplamıdır. Birincisi bir (1) vektörü ile gerçek görüntülerde ayırıcı ağın tahminleri arasındaki farkın karesidir. İkincisi sıfır (0) vektörü ile tahmin edilen görüntülerde (üretilen görüntülerde) ayırıcı ağın tahminleri arasındaki farkın karesidir. Bu kayıp fonksiyonu aşağıdaki eşitlikte gösterilmiştir.

$$Ayırıcı\ Kayıp\ (L1) = \left(0 - \hat{Y}_{üretilmiş}\right)^2 + (1 - Y_{Gerçek})^2 \quad (4.10)$$

5.VGG-19 ağı ve Özellikler

VGG transfer derin öğrenme yöntemi gerçek ve üretilmiş görüntülerde farklı katmanlardan öznitelik çıkarımı için kullanılmıştır. Hem üretici hem de ayırıcı ağ için benzer ayarlardan faydalanılmıştır. Adam optimizasyon yöntemi, eğitim için 60 iterasyon, 0.002 öğrenme oranı, mini-batch boyutu ise bir (1) olarak ayarlanmıştır.

6.Ağın Eğitimi

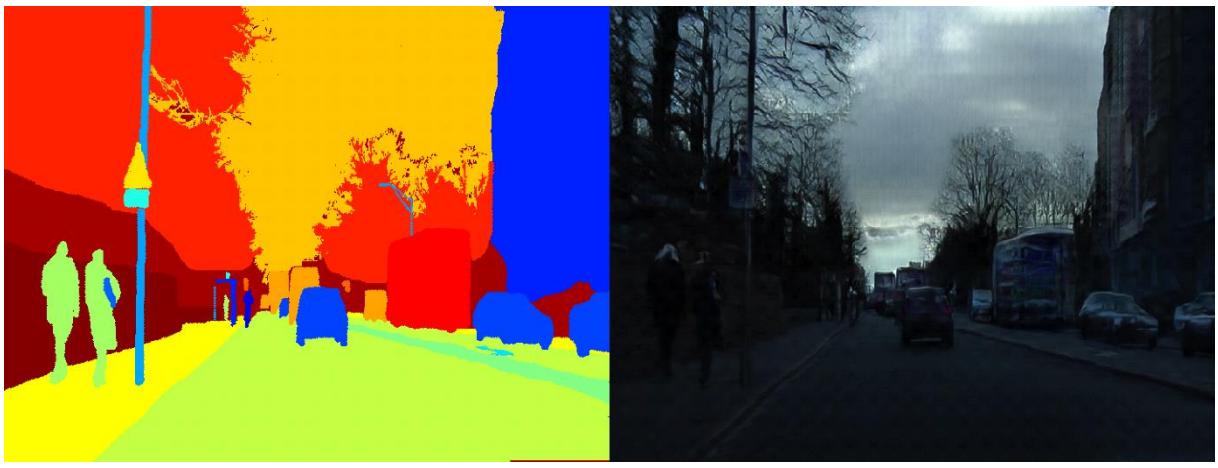
Ağın eğitimi CPU ortamında gerçekleştirilmiştir. AMD Ryzen 7 4800H with Radeon Graphics 2.90 GHz işlemcili, 8 GM RAM'e sahip bir dizüstü bilgisayarda gerçekleştirilmiştir.

7.Modelin Test Edilmesi

Modeli test etmek için farklı etiket görüntüler kullanılmıştır. Bir kısmı test veri setinden rastgele seçilirken birkaç deneme bizim tarafımızca etiketler çizilip test edilmiştir. İlk denemeye ait grafikler aşağıda verilmiştir (Bkz: Şekil 4.11).

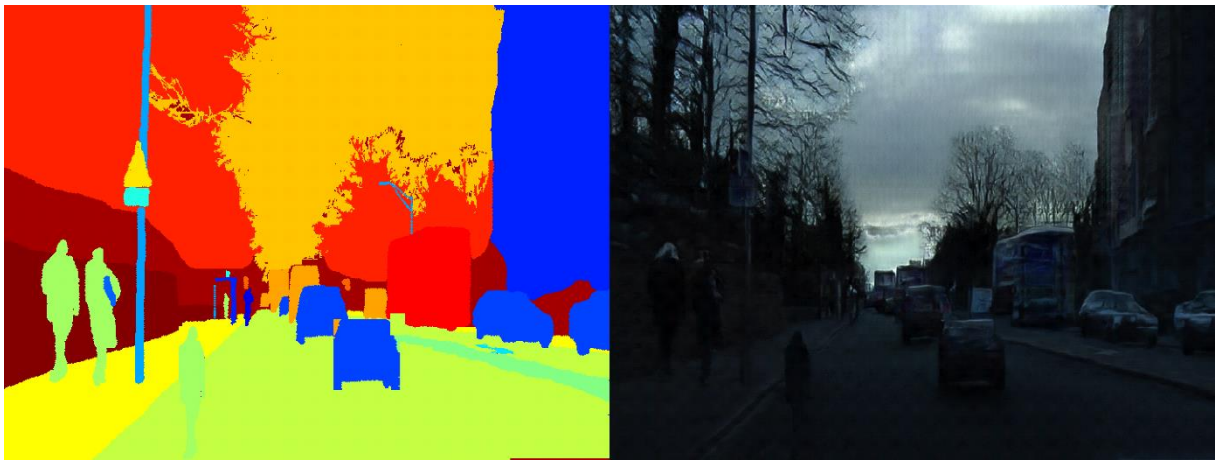
Şekil 4. 11. (A) Piksel görüntü, (B) üretilmiş ve (C) gerçek görüntü

İlk örnek için etiket piksel görüntü için üretilmiş görüntü gerçek imgeler içermektedir. Sonraki denemelerde aynı piksel görüntü üzerinde değişiklikler yapılarak görüntüler üretilmiştir. Piksel görüntü ile üretilmiş görüntü Şekil 4.12'de verilmiştir.



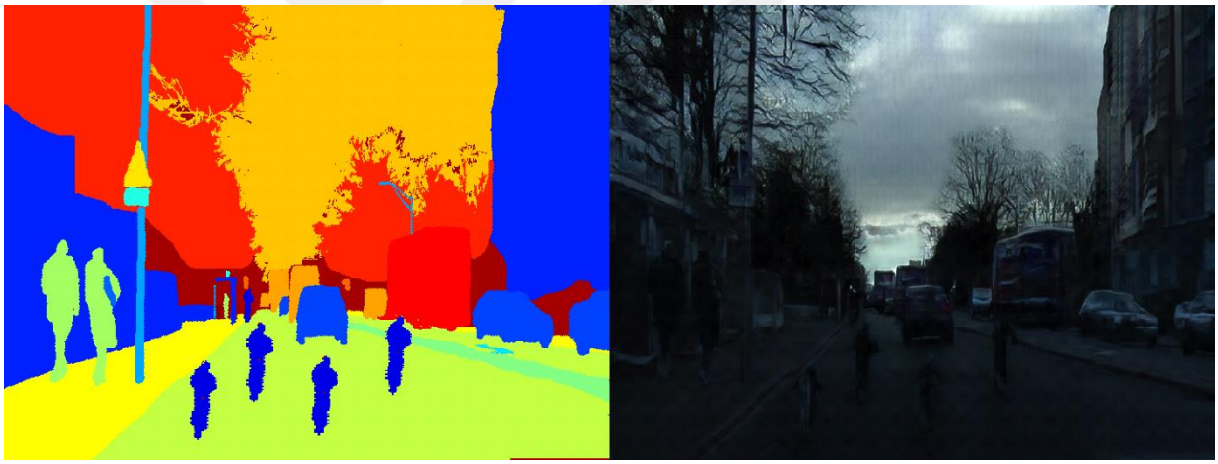
Şekil 4. 12. Aynı piksellere sahip görüntü üzerinde deneme 1

Şekil 4.12'deki piksel görüntü üzerinde yola yeni bir araba ve kişi eklenmiştir. Üretilen görüntüde kişi ve araba görüntülerin düzgün bir şekilde üretildiği Şekil 4.13'ten görülmektedir.



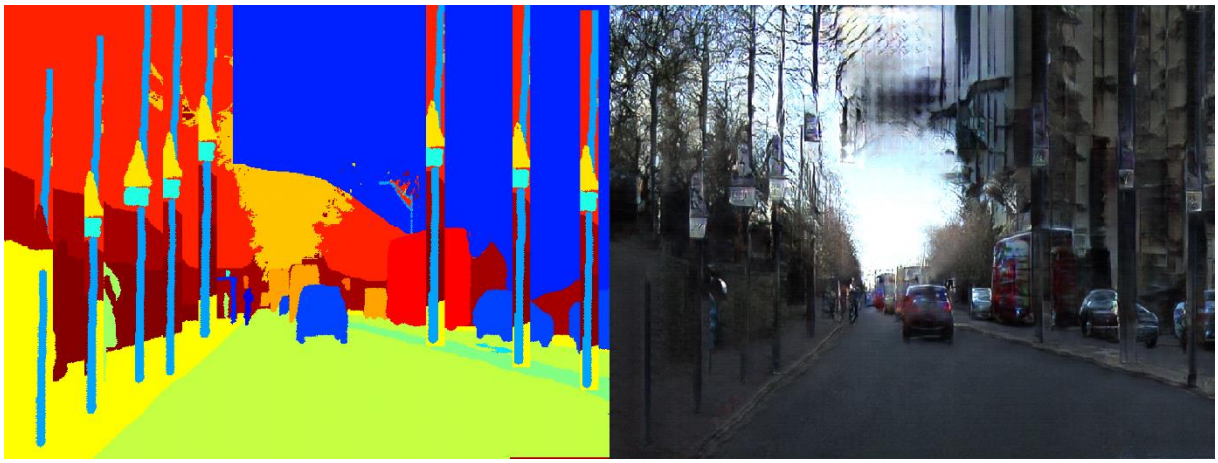
Şekil 4. 13. Aynı piksellere sahip görüntü üzerinde deneme 2

Aynı piksel görüntü üzerinde yola bisikletli kişiler eklenmiştir. Bu kişilerin gerçek olarak üretildiği Şekil 4.14'te görülmektedir.



Şekil 4. 14. Aynı piksellere sahip görüntü üzerinde deneme 3

Son denemede piksel görüntü üzerinde direkler eklenmiştir. Direkler düzgün eklenmemesine rağmen üretilen görüntüde bu direklerin üretildiği Şekil 4.15'te görülmektedir.



Şekil 4. 15. Aynı piksellere sahip görüntü üzerinde deneme 4

5. TARTIŞMA VE SONUÇ

Araştırma iki bölümden oluşmaktadır. *Birinci bölümde;* anlamsal bölütleme işlemi gerçekleştirilmiştir. *İkinci bölümde*; piksel görüntülerden gerçek sentetik görüntüler üretilmektir. Birinci bölümde gerçek bir veri seti üzerinde anlamsal bölütleme işlemleri gerçekleştirilmiştir. Derin öğrenme, temelde üç veya daha fazla katmana sahip yeni nesil yapay sinir ağlarıdır. Derin öğrenme metotları makine öğreniminin bir alt kümesidir. Yapay sinir ağları insan beyin yapısını simüle eden metotlardır. Yapay sinir ağları tek bir sinir ağı yaklaşık tahminler yapılırken, ek gizli katmanlarla daha yüksek doğruluk sonuçlarına ulaşılabilinmektedir. Derin öğrenme metotları analitik ve fiziksel görevleri insan müdahalesi olmadan gerçekleştirilebilen yapay zekâ uygulamalarıdır. Derin öğrenme metotları veri setinin boyutu ve öğrenme metotları ile klasik makine öğrenmesinden ayrılır. Bugün derin öğrenme metotları çok farklı alanlarda kullanılabilinmektedir. Finansal servislerde, mühendislik uygulamalarında, sağlık, ziraat, eğitim, dil modelleme gibi neredeyse her alanda uygulamalarına rastlanmaktadır. Derin öğrenme metotlarının bir uygulaması başarılı bir şekilde anlamsal bölütleme işlemlerinde kullanılabilinmesidir.

Anlamsal bölütleme, bilgisayarla görü uygulamalarında önemli bir araştırma konusudur. Görüntülerdeki nesnelerin ve alanların anlamlı sınıflandırmasını gerçekleştirir. Görüntüdeki her piksel için bir sınıflandırma etiketi oluşturmayı amaçlar. Örneğin, görüntü içerisindeki araba, insan, bina gibi nesnelerin daha önceden belirlenmiş piksel gruplarına ait olduğunu tespit edebilmektedir. Bazı uygulamalarda bu sınıflandırma son derece hassas olmaktadır. Özellikle otonom araçlarda anlamsal bölütleme başarısı çok önemlidir (Hongshan *vd.,* 2018). Görüntü, birbirinden farklı birçok renkten piksellerin bütünleştiği bir yapıdır ve bilgisayarların görüntüyü işleyebilmesi için, anlamsal bölütleme işlemine ihtiyaç duyulmaktadır. Anlamsal bölütleme işlemi, temel olarak iki adımda gerçekleşir. İlk aşamada görüntü içindeki pikseller benzerliklerine göre bölütlenmektedir. Birbirine yakın pikseller piksel bulutları şeklinde birbirinden ayrışmaktadır. *İkinci aşamada* ise her piksel bulutuna etiketleme gerçekleştirilmektedir. Örneğin; bir görüntüdeki insanlar ve köpekler ayrı şekilde piksel gruplarına bölündükten sonra hangi piksel grupların insan, hangilerinin köpek olduğunun etiketlenmesi anlamsal bölütleme işlemidir. Görüntüdeki her insan veya her köpek ayrı ayrı etiketlenmelidir. Derin öğrenme metotları anlamsal bölütleme işlemleri içinde kullanılabilir. DeepLab v3+, U-Net gibi evrişimli sinir ağlarından anlamsal bölütleme

amaçlı da faydalanılabilmektedir. Çok çeşitli derin sinir ağları (DNN) vardır. Derin evrişimli sinir ağları (CNN veya DCNN), görüntülerdeki ve videodaki nesneleri tanımlamak için en sık kullanılan türdür. DCNN'ler, hayvanların görsel korteksinden ilham alan üç boyutlu bir sinir modeli kullanan geleneksel yapay sinir ağlarından evrimleşmiştir. Derin evrişimli sinir ağları, esas olarak nesne algılama, görüntü sınıflandırma, öneri sistemleri gibi uygulamalarda kullanılmaktadır. Bir evrişimli sinir ağı makine öğreniminin bir alt kümesidir. Farklı uygulamalar ve veri türleri için kullanılan çeşitli yapay sinir ağlarından biridir. Bir CNN, derin öğrenme algoritmaları için bir tür ağ mimarisidir ve özellikle görüntü tanıma ve piksel verilerinin işlenmesini içeren görevler için kullanılır.

CNN ağlar büyük boyutlarda veri setleri ile çalışır. Veri setlerin yeterli olmaması durumunda farklı mimarilerde transfer CNN yöntemleri kullanılır. Makine öğrenimi için transfer öğrenimi, mevcut öğrenmiş modellerin yeni bir problemde veya uygulamada yeniden kullanılmasıdır. Transfer öğrenimi, farklı bir makine öğrenimi algoritması türü değildir, bunun yerine modelleri eğitirken kullanılan bir teknik veya yöntemdir. Önceki eğitimden geliştirilen bilgi, yeni bir görevin gerçekleştirilmesine yardımcı olmak için geri dönüştürülür. Orijinal eğitilmiş model, genellikle yeni görünmeyen verilere uyum sağlamak için yüksek düzeyde bir genelleme gerektirir. Transfer öğrenimi, her yeni görev için eğitimin sıfırdan yeniden başlatılması gerekmediği anlamına gelir. Yeni makine öğrenimi modellerini eğitmek hem kaynak yetersiz hem de zaman alıcı olabilir, bu nedenle transfer öğrenimi hem kaynaklardan hem de zamandan tasarruf sağlar (Aswathy *vd.,* 2021). Anlamsal bölütleme için Resnet-18 transfer derin öğrenme mimarisi kullanılarak oluşturulan DeepLab V3+ ağı kullanılmıştır. Derin öğrenme metotları sınıflandırma problemlerinde genel olarak başarılıdır. Ancak farklı nedenlerden dolayı ağların eğitimi zor olmaktadır. Birincisi patlayan (yok olan) gradyanlar olarak isimlendirilen, bir nöronun eğitim sürecinde ölür ve aktivasyon fonksiyonuna bağlı olarak asla geri gelmemesi durumudur. İkincisi ise transfer yöntemlerin parametrelerinin optimizasyon problemidir. Ağın derinliği arttıkça bu parametrelerin fazla olması yüzünden eğitim zorlaşmaktadır (Srivastava, 2015).

Önerilen bu problemleri çözmenin etkili bir yolu Resnet Ağlar' dır (He *vd.,* 2015). ResNet' teki temel fark, normal evrişimli katmanlarına paralel kısayol bağlantılarına sahip olmalarıdır. Evrişim katmanlarının aksine, bu kısayol bağlantıları her zaman canlıdır ve hatayı kolayca geri yayılabilmektedir. ResNet18, ResNet34, ResNet50, ResNet101ve ResNet152 farklı mimarilerde ResNet ağlar önerilmiştir. Anlamsal bölütleme deneyleri

için Cambridge Üniversitesi'nden CamVid veri seti kullanılmıştır. Bu veri seti, sürüş sırasında elde edilen sokak görüntülerinden oluşmaktadır. Veri seti araba, yaya ve yol dâhil olmak üzere 32 anlamsal sınıf için piksel düzeyinde etiketlenmeler yapılmıştır. Veri setinde 32 sınıf bilgisi mevcuttur. Ancak mevcut çalışmada bir kısım etiketler birleştirilerek tek etiket ile ifade edilmiştir. Örneğin; tren, otobüs, araba ve diğer hareket eden taşıtlar araba etiketi ile belirtilmiştir. Bu 11 etiket {gökyüzü, bina, direk, yol, kaldırım, ağaç, işaret/ sembol, çit, araba, yaya ve bisiklet} olarak belirtilmiştir. Performans ölçütleri olarak Jaccard indeksi, Sorensen-Dice ve BF skore ölçütleri kullanılmıştır. Gerçekleştirilen anlamsal bölütleme deneylerinde 11 etiket {gökyüzü, bina, direk, yol, kaldırım, ağaç, işaret/ sembol, çit, araba, yaya ve bisiklet} için yüksek başarılı tahminler gözlenmiştir (Bkz: Tablo 4.3., Tablo 4.4.).

Tezin ikinci aşamasında derin öğrenme metotları piksel görüntülerden sentetik gerçekçi görüntüler oluşturulmaya çalışılmıştır. Piksel bazlı görüntüler, genellikle düşük çözünürlüklü ve sınırlı renk uzayına sahiptir. Ancak, gerçek görüntüler daha yüksek çözünürlüklü ve genellikle daha fazla renk seçeneği sunar. Bu nedenle, piksel bazlı görüntüler gerçek görüntülere dönüştürülerek daha gerçekçi ve ayrıntılı görüntüler elde edilebilir. Piksel bazlı görüntülerin gerçek görüntülere dönüştürülme işlemi, oyun, tarım, sağlık gibi birçok alanda kullanılmaktadır (Güzel ve Bilge, 2020). Sentetik görüntüler üretmek için GAN modellerine ihtiyaç vardır. Örneğin; moda tasarımında yeni kıyafetlerin tasarlanması, mimaride yeni binaların tasarlanması, sanat eserleri üretimi gibi alanlarda kullanılmaktadır. Ayrıca, oyun geliştiricileri de GAN modellerini kullanarak oyunlarda yeni nesneler üretebilmektedir. Bununla birlikte, GAN modelleri ile üretilen nesnelerin gerçekliği konusunda bazı sınırlamalar ve zorluklarda vardır.

Tez çalışmamızda gerçek görüntüler oluşturmak için Px2PxHD GAN modeli kullanılmıştır. Pix2PixHD, yüksek çözünürlüklü görüntülerin düşük çözünürlüklü eşlemelerinden gerçekçi ve ayrıntılı görüntüler üretmek için kullanılan bir görüntü çeviri yöntemidir. Bu yöntemin temelinde, derin öğrenme ve özellikle evrişimli sinir ağları vardır. Pix2PixHD GAN yönteminde CNN ağı olarak VGG19 transfer derin öğrenme metodu kullanılmıştır.

VGG modeller ağın derinliğini artıran modeller olup buna bağlı olarak performansı da artıran modellerdir. Basit modül, küçük bir evrişim çekirdeği, küçük havuzlama çekirdeği ve ReLU'dan oluşur. 5 evrişimli katman, 3 tam bağlantılı katman ve bir softmax çıktı katmanı vardır. Katmanları ayırmak için maksimum havuzlama kullanılır ve tüm gizli katmanların etkinleştirme birimleri için ReLU aktivasyon

fonksiyonu kullanılır (Kong ve Cheng,2022). Bu nedenle VGG ağlarının en büyük avantajlarından biri, sinir ağlarının yapısını basitleştirmeleridir. Elde edilen 7 × 7 × 512 öznitelik haritası tam bağlantılı katmana ve ardından softmax aktivasyonu kullanılır.

Modeli test etmek için farklı piksel görüntüler kullanılmıştır. Bir kısmı test veri setinden rastgele seçilirken birkaç deneme bizim tarafımızca piksel görüntüler çizilip test edilmiştir (Bkz: Şekil 4.11., Şekil 4.12., Şekil 4.13., Şekil 4.14. ve Şekil 4.15.). Sonuç olarak; piksel görüntülerden gerçeğe yakın görüntülerin elde edildiği görülmüştür.

KAYNAKÇA

Alo, Z. (2018). *Improved Deep Convolutional Neural Networks (DCNN) Approaches for Computer Vision and Bio-Medical Imaging*, (Ph.D. Dissertation). Dept. Electrical and Computer Engineering, University of Dayton, Dayton, USA.

Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., ... ve Asari, V. K. (2018). The history began from alexnet: A comprehensive survey on deep learning approaches. arXiv preprint arXiv:1803.01164.

Asano, S., Maruyama, T., ve Yamaguchi, Y., 2009. Performance comparison of FPGA, GPU and CPU in image processing, *International Conference on Field Programmable Logic and Applications 2009*, IEEE.

Aswathy, A. L., Hareendran, A., ve SS, V. C. (2021). COVID-19 diagnosis and severity detection from CT images using transfer learning and back propagation neural network. *Journal of Infection and Public Health, 14*(10), 1435-1445.

Brostow, G. J., Fauqueur, J., ve Cipolla, R. (2009). Semantic object classes in video: A high-definition ground truth database. *Pattern Recognition Letters, 30*(2), 88-97.

Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., ve Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 801-818).

Çelik, G., ve Talu, M. F. (2019). Çekişmeli üretken ağ modellerinin görüntü üretme performanslarının incelenmesi. *Balıkesir Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 22*(1), 181-192.

Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., ve Garcia-Rodriguez, J. (2017). A review on deep learning techniques applied to semantic segmentation. arXiv preprint arXiv:1704.06857.

Goodfellow, I., Bengio, Y., Courville, A.,ve Bengio, Y., (2016). *Deep Learning*, MIT press Cambridge.

He, K., Zhang, X., Ren, S., ve Sun, J. (201). Identity mappings in deep residual networks. In: *European Conference on Computer Vision* (pp. 630-645). Springer, Cham.

Hu, S., Bonardi, F., Bouchafa, S., ve Sidibé, D. (2023). *Multi-Modal Unsupervised Domain Adaptation for Semantic Image Segmentation.* Pattern Recognition, 109299.

Huang, G.-B., Zhou, H., Ding, X., ve Zhang, R. (2012). Extreme Learning machine for regression and multiclass classi cation, *IEEE Trans. Syst. Man Cybern. Part B Cybern, 42*(2), 513-529.

Isola, P., Zhu, J. Y., Zhou, T., ve Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (ss: 1125-1134).

Jiao, L., Hu, C., Huo, L., & Tang, P. (2021). Guided-Pix2Pix: End-to-end inference and refinement network for image dehazing. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 3052-3069.

Jaworek-Korjakowska, J., Kleczek, P., & Gorgon, M. (2019). Melanoma thickness prediction based on convolutional neural network with VGG-19 model transfer learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 0-0).

Kim, K.G. (2016). Deep learning book review, *Healthcare Informatics Research, 22*(4), 351–354.

Kong, L., ve Cheng, J. (2022). Classification and detection of COVID-19 X-Ray images based on DenseNet and VGG16 feature fusion. *Biomedical Signal Processing and Control, 77*, 103772.

Lateef, F., ve Ruichek, Y. (2019). Survey on semantic segmentation using deep learning techniques. *Neurocomputing, 338*, 321-348.

Li, R., Pan, J., Li, Z., ve Tang, J. (2018). Koşullu üretici hasım ağı aracılığıyla tek görüntü bulanıklaştırma. *Bilgisayarla Görme ve Örüntü Tanıma Üzerine IEEE Konferansı*. (ss. 8202-8211).

Liu, W., Sun, J., Li, W., Hu, T., ve Wang, P. (2019). Deep learning on point clouds and its application: A survey. *Sensors, 19*(19), 5.

Li, Y., Fu, R., Meng, X., Jin, W., ve Shao, F. (2020). A SAR-to-optical image translation method based on conditional generation adversarial network (cGAN). *IEEE Access, 8*, 60338-60343.

Mirza, M. ve Osindero, S., (2014). Conditional Generative Adversarial Nets, *CoRR*, 1–7.

Min, H., Zhang, Y., Zhao, Y., Jia, W., Lei, Y., ve Fan, C. (2023). Hybrid Feature Enhancement Network for Few-Shot Semantic Segmentation. *Pattern Recognition*, 109291.

Özkan, İ. N. İ. K., ve Ülker, E. (2017). Derin öğrenme ve görüntü analizinde kullanılan derin öğrenme modelleri. *Gaziosmanpaşa Bilimsel Araştırma Dergisi, 6*(3), 85-104.

Qin, Z., Liu, Z., Zhu, P., ve Ling, W. (2022). Style transfer in conditional GANs for cross-modality synthesis of brain magnetic resonance images. *Computers in Biology and Medicine, 148*, 8.

Rakhimol, V., ve Maheswari, P. U. (2022). Restoration of ancient temple murals using cGAN and PConv networks. *Computers & Graphics, 109*, 100-110.

Salehi, P., ve Chalechale, A. (2020). Pix2pix-based stain-to-stain translation: A solution for robust stain normalization in histopathology images analysis. In: *IEEE.2020*

International Conference on Machine Vision and Image Processing (MVIP) (ss: 1-7).

Seker, A., Diri, B., ve Balık, H.H. (2017). derin öğrenme yöntemleri ve uygulamaları hakkında bir ˙inceleme, *Gazi Mühendislik Bilimleri Dergisi*, *3*(3), 47–64.

Shen, D., Wu, G., ve Suk, H. I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering, 19*, 221-248.

Turhan, C. G., ve Bilge, H. Ş. (2020). Çekişmeli üretici ağ ile ölçeklenebilir görüntü oluşturma ve süper çözünürlük. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, 35*(2), 953-966.

Wang, P., Chen, P., Yuan, Y., Liu, D., Huang, Z., Hou, X., ve Cottrell, G. (2018, March). Understanding convolution for semantic segmentation. In: *2018 IEEE winter conference on applications of computer vision (WACV)* (ss: 1451-1460).

Wang, S., Liu, Y., Wang, L., Sun, Y., ve Yin, B. (2023). PASIFTNet: Scale-and-Directional-Aware Semantic Segmentation of Point Clouds. *Computer-Aided Design, 156*, 8.

Woldesellasse, H., ve Tesfamariam, S. (2022). Prediction of lateral spreading displacement using conditional Generative Adversarial Network (cGAN). *Soil Dynamics and Earthquake Engineering, 156*, 107214.

Xiao, J., Wang, J., Cao, S. ve Li, B. (2020). Maske takan işçilerin tespitinde yeni ve geliştirilmiş bir VGG-19 ağının uygulanması. *Journal of Physics: Conference Series'de. 1518*(1) 012041. GİB Yayıncılık.

Yang, X., Zhao, J., Wei, Z., Wang, N., ve Gao, X. (2022). SAR-to-optical image translation based on improved CGAN. *Pattern Recognition, 121*, 9.

Yu, H., Yang, Z., Tan, L., Wang, Y., Sun, W., Sun, M., ve Tang, Y. (2018). Methods and datasets on semantic segmentation: A review. *Neurocomputing*, *304*, 82-103.

Yuan, K., Schaefer, G., Lai, Y. K., Wang, Y., Liu, X., Guan, L., ve Fang, H. (2022). *MuSCLe: A Multi-Strategy Contrastive Learning Framework for Weakly Supervised Semantic Segmentation*. arXiv preprint arXiv:2201.07021.

Zhang, H., Dana, K., Shi, J., Zhang, Z., Wang, X., Tyagi, A., & Agrawal, A. (2018). Context encoding for semantic segmentation. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (pp. 7151-7160).

Zhou, T., Lu, H., Yang, Z., Qiu, S., Huo, B., ve Dong, Y. (2021). The ensemble deep learning model for novel COVID-19 on CT images. *Applied Soft Computing, 98*,12.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı **:** Emre ARICA
Uyruğu **:** T.C.

EĞİTİM

Derece		**Adı, İlçe, İl**	**Bitirme Yılı**
Üniversite	:	**Siirt Üniversitesi**	**2017**
Yüksek Lisans	:		
Doktora	:		

İŞ DENEYİMLERİ

Yıl	**Kurum**	**Görevi**
2019	**Milli Eğitim Bakanlığı**	**Öğretmen**

UZMANLIK ALANI

YABANCI DİLLER

YAYINLAR

Arıca, E., Kaya, Y. (2023). ResNET-18 Transfer Derin Öğrenme ile Anlamsal Bölütleme. Uluslararası Bilişim Kongresi. Mart-Batman, 2023