

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

**3B NOKTA BULUTLARININ
DERİN ÖĞRENME YÖNTEMİYLE
SEMANTİK SEGMENTASYONU**

YÜKSEK LİSANS TEZİ

Mert BEŞİKTEPE

Geomatik Mühendisliği Anabilim Dalı

Geomatik Mühendisliği Programı

ŞUBAT 2022

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

**3B Nokta Bulutlarının
Derin Öğrenme Yöntemiyle
Semantik Segmentasyonu**

YÜKSEK LİSANS TEZİ

**Mert BEŞİKTEPE
(501181657)**

Geomatik Mühendisliği Anabilim Dalı

Geomatik Mühendisliği Programı

Tez Danışmanı: Doç. Dr. Caner GÜNEY

ŞUBAT 2022

İTÜ, Lisansüstü Eğitim Enstitüsü’nün 501181657 numaralı Yüksek Lisans Öğrencisi Mert BEŞİKTEPE, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı “3B Nokta Bulutlarının Derin Öğrenme Yöntemiyle Semantik Segmentasyonu” başlıklı tezini aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

Tez Danışmanı : **Doç. Dr. Caner GÜNEY**
İstanbul Teknik Üniversitesi

Jüri Üyeleri : **Prof. Dr. Orhan AKYILMAZ**
İstanbul Teknik Üniversitesi

Dr. Öğr. Üyesi Umut AYDAR
Çanakkale Onsekiz Mart Üniversitesi

Teslim Tarihi : **14.01.2022**
Savunma Tarihi : **11.02.2022**





Aileme,



ÖNSÖZ

Bu yüksek lisans tezinin konusunun belirlenmesi, genel çerçevesinin oluşturulması ve tamamlanmasında destek ve rehberliğini esirgemeyen danışmanım Sayın Doç. Dr. Caner GÜNEY'e emekleri ve teşvikleri için teşekkür ederim.

Ayrıca bugüne dek her türlü zorluğa göğüs gererek eğitim öğretim hayatımın ve kariyerimin bugünlerine ulaşmamı sağlayan Aileme sonsuz şükranlarımı sunarım.

Pandemi koşulları nedeniyle zamana yayılan tezimin tamamlanması sürecinde sabır ve anlayış gösteren Sayın Müdürüm Yasemin Meriç Hanım'a ve rehberlerim olan Sayın Kayhan Köken ve Sayın Kadriye Cansu Berber başta olmak üzere tüm çalışma arkadaşlarıma teşekkür ederim.

Yüksek Lisans tezimi hazırladığım süre boyunca arkadaşlığı ve destekleri için Sayın Vahdeddin Aslan'a teşekkür ederim.

Son olarak küresel pandemi dönemine denk gelen bu yüksek lisans tezinin hazırlanması sürecinde her durumda sonsuz sabrı, anlayışı, desteği, yardımları ve teşviği ile katkı koyarak tamamlanabilmesinde en büyük manevi katkıyı sağlayan Sayın Gamze Zeytin'e teşekkürlerimi ve minnetlerimi sunarım.

Ocak 2022

Mert BEŞİKTEPE
(Harita Mühendisi)



İÇİNDEKİLER

Sayfa

ÖNSÖZ	vii
İÇİNDEKİLER	ix
KISALTMALAR	xi
SEMBOLLER	xiv
ÇİZELGE LİSTESİ.....	xvi
ŞEKİL LİSTESİ.....	xviii
ÖZET	xxi
SUMMARY	xxiii
1. GİRİŞ	1
2. LİTERATÜR ARAŞTIRMASI	5
2.1 Nokta Bulutları.....	5
2.2 Yapay Öğrenme.....	7
2.2.1 Yapay zeka	7
2.2.2 Makine öğrenmesi.....	8
2.2.3 Temsil öğrenme.....	9
2.2.4 Yapay sinir ağları	11
2.2.5 Derin öğrenme.....	14
3. 3B NOKTA BULUTLARI VE DERİN ÖĞRENME	21
3.1 Nokta Bulutlarından Derin Öğrenme Çalışmalarında Karşılaşılan Zorluklar ..	22
3.1.1 Üretim koşullarından doğan zorluklar	22
3.1.2 Veri tipinin karakteristik özelliklerinden doğan zorluklar	23
3.2 Nokta Bulutlarından Derin Öğrenme Sonuçlarını Değerlendirmede Kullanılan Ölçütler	24
3.3 3B Nokta Bulutlarından Derin Öğrenme İçin Geliştirilen Çözümler	27
3.3.1 Projektif yöntemler	27
3.3.2 Volumetrik yöntemler	30
3.3.3 Nokta bulutu temelli yöntemler	31
3.3.3.1 Çok katmanlı algılayıcı temelli yöntemler	32
3.3.3.2 3B nokta evrişim metodları	34
3.3.3.3 Diğer yöntemler	38
3.4 Derin Öğrenme İle Semantik Segmentasyon 3B Nokta Bulutu Veri Setleri ...	39
3.5 3B Nokta Bulutu Semantik Segmentasyonunda Son Teknoloji Çözümler	47
3.6 Kernel Point Convolution (KPConv) Modelinin İncelenmesi	53
4. 3B NOKTA BULUTLARININ SEMANTİK SEGMENTASYONU UYGULAMASI.....	57
4.1 KP-FCNN Modelinin Bulut Ortamında Eğitilmesi.....	58
4.1.1 Model eğitimi için bulut geliştirme ortamının düzenlemesi	58
4.1.2 Model eğitimi	62
4.1.3 Eğitim sonuçları ve modelin test edilmesi	65
5. SONUÇLAR VE BULGULAR	71

KAYNAKLAR.....	75
ÖZGEÇMİŞ.....	85



KISALTMALAR

AWS	: Amazon Web Services
ANN	: Artificial Neural Network
CAD	: Computer Aided Design
CNN	: Convolutional Neural Network
DCNN	: Deep Convolutional Neural Network
DL	: Deep Learning
HLT	: Hava Lazer Tarayıcı
FN	: False Negative
FP	: False Positive
EC2	: Elastic Compute Cloud
EoC	: Error of Comission
EoO	: Error of Omission
GCP	: Google Cloud Platform
IaaS	: Infrastructure as a Service
IoU	: Intersection over Union
ILSVRC	: ImageNet Large Scale Visual Recognition Challenge
kNN	: k Nearest Neighbor
KP	: Kernel Point
KPFCNN	: Kernel Point Fully Convolutional Neural Network
LSTM	: Long Short Term Memory
MLP	: Multi Layer Perceptron
MLT	: Mobil Lazer Tarayıcı
mAcc	: mean per-class Accuracy
OA	: Overall Accurcy
PLY	: Polygon File Format
RGB	: Red Green Blue
RNN	: Recurrent Neural Network
SFTP	: Secure File Transfer Protocol
SSH	: Secure Shell
TN	: True Negative

TP : True Positive
URL : Uniform Resource Locator
USD : United States Dollar
VRAM : Video Random Access Memory
YLT : Yersel Lazer Tarayıcı





SEMBOLLER

Ψ : Psi
F1 : F1 Skoru





ÇİZELGE LİSTESİ

Sayfa

Çizelge 2.1 : Çeşitli nokta bulutu üretim yöntemlerine yönelik bilgiler.	6
Çizelge 3.1 : Seçilen veri setlerine ilişkin genel bilgiler.	48
Çizelge 3.2 : Hessigheim3D veri seti için raporlanan performans sonuçları.....	49
Çizelge 3.3 : DALES veri seti için raporlanan performans sonuçları.	50
Çizelge 3.4 : SensatUrban veri seti için eğitim sonuçları.	52
Çizelge 4.1 : SageMaker ve EC2 tarafından derin öğrenme için sunulan makine fiyatları.	60
Çizelge 4.2 : Test seti üzerinde elde edilen sonuçlar.	65
Çizelge 4.3 : Test seti üzerinden elde edilen doğruluk ölçütleri.....	67



ŞEKİL LİSTESİ

Sayfa

Şekil 2.1 : Yapay Zeka, Makine Öğrenmesi ve Derin Öğrenme kavramlarının birbirleri ile olan ilişkisini ifade eden Venn şeması.	7
Şekil 2.2 : Makine öğrenmesi ve temsil öğrenme kavramlarını açıklamaya yönelik örnek.	10
Şekil 2.3 : Örnek bir fonksiyon üzerinde meyil azalımlı optimizasyon algoritmasında gradyenler.	12
Şekil 2.4 : Perceptron ve modern tek katmanlı algılayıcı modellerinin gösterimi. ...	14
Şekil 2.5 : 2012 ILSVRC yarışmasındaki en başarılı 6 modelin, her bir örneğe ait ürettikleri 5 tahmin sonucunda hesaplanan ortalama hata oranları.	15
Şekil 2.6 : 2011-2022 yılları arasında ILSVRC kazananları.	16
Şekil 2.7 : Temsiller hiyerarşisi kavramını ifade etmek üzere düzenlenmiş görsel. .	19
Şekil 3.1 : 3B nokta bulutlarında veri üretim koşullarından kaynaklı noksanlıklar. .	23
Şekil 3.2 : Nokta bulutu veri tipine dair yapısal özelliklerden bazılarının görselleştirilmesi.	24
Şekil 3.3 : Kesinlik ve duyarlılık kavramlarını ifade etmek için hazırlanmış görsel.	26
Şekil 3.4 : Kesişim-Birleşim oranı (IoU) değerini bir küme problemi olarak ifade etmek için hazırlanmış görsel.	26
Şekil 3.5 : 3B nokta bulutlarından derin öğrenme için geliştirilen yöntemlerin veriyi ele alış biçimlerine göre hazırlanmış bir taksonomisi.	28
Şekil 3.6 : MVCNN mimarisi.	28
Şekil 3.7 : MV3D mimarisi.	29
Şekil 3.8 : Örnek bir obje için (A) basit uniform vokselizasyon yapısı ve (B) octree tabanlı volumetrik temsile ait kesitler.	31
Şekil 3.9 : PointNet Mimarisi.	33
Şekil 3.10 : Simetrik fonksiyonlara örnekler.	33
Şekil 3.11 : PoinNet++ mimarisi, yerel özellik vektörlerinin hiyerarşik biçimde elde edilmesi.	34
Şekil 3.12 : 3B nokta evrişim işlemini uygulamak için genellikle kullanılan yaklaşımların görselleştirmesi.	35
Şekil 3.13 : Sürekli fonksiyon olarak tanımlanmış 3B nokta evrişim işlemine bir örnek.	36
Şekil 3.14 : Graf evrişim temelli yöntemlerin örnek işlem akışı şeması.	38
Şekil 3.15 : Özyinelemeli mimari temelli yöntemlere örnek işlem akış şeması.	39
Şekil 3.16 : Oakland 3D Point Cloud Dataset veri örneği.	40
Şekil 3.17 : Semantic3D veri setinden bir örnek.	42
Şekil 3.18 : DALES veri setinden bir örnek.	43
Şekil 3.19 : Nokta bulutu semantik segmentasyonu öncesi ve sonrasını temsil eden görsel.	47
Şekil 3.20 : KPConv katmanın işlem prensibinin 2B üzerinde görselleştirilmesi.	53
Şekil 3.21 : Ağ boyunca altörnekleme ile nokta sayısının azaltılması ve yerel komşuluk alanlarının genişletilmesi.	54

Şekil 3.22 : Semantik segmentasyon için tasarlanan KP-FCNN modeli.....	55
Şekil 4.1 : KP-FCNN modelinin bulut ortamında eğitilmesi için kurulan yazılım geliştirme ortamı bileşenlerinin ilişkisini gösteren şema.....	62
Şekil 4.2 : 3 Epok boyunca eğitim seti üzerinde hesaplanan OA değerleri.....	64
Şekil 4.3 : Eğitim boyunca validasyon seti üzerinde hesaplanan ortalama IoU değerlerinin değişimi.	64
Şekil 4.4 : Eğitim boyunca validasyon seti üzerinde hesaplanan sınıflar ortalaması olarak hesaplanan IoU değeri.	65
Şekil 4.5 : DALES veri setindeki sınıflara ait nokta sayılarının dağılımı.	67
Şekil 4.6 : Modelin test edilmesi sonucunda üretilen hata matrisi.	69
Şekil 4.7 : Model tarafından üretilen sonuçlar.	70





3B NOKTA BULUTLARININ DERİN ÖĞRENME YÖNTEMLERİ İLE SEMANTİK SEGMENTASYONU

ÖZET

Üç boyutlu (3B) kent modelleri, kentsel alanlarda yaşanabilecek değişimlerin 3B olarak izlenmesi, kent çevre ilişkisinin modellenmesi ve kentsel simülasyon uygulamaları için altyapı sağlayarak mekansal planlama çalışmaları için veri temelli karar verme süreçlerinin geliştirilmesine olanak tanımaktadır. Söz konusu modellerin geliştirilmesi için; detay düzeyi uygulama amaçlarına bağlı olarak değişen, yüksek doğruluklu, güncel ve büyük miktarda 3B mekansal bilgiye gereksinim duyulmaktadır. Lazer tarama, görüntüleme ve yer gözlem uyduları gibi veri elde etme teknikleri ile fotogrametri ve bilgisayarlı görü alanlarındaki veri değerlendirme yöntemlerine ilişkin yaşanan gelişmeler 3B mekansal bilginin nokta bulutu biçiminde doğrudan ölçme ile ya da başka veri kaynaklarından türetme yolu ile üretilmesine olanak sağlamıştır. 3B nokta bulutları günümüzde klasik kullanım alanları olan ölçme ve planlama çalışmaları haricinde robotik, otonom mobilite, 3B modelleme ve dijital ikiz oluşturma gibi yüksek teknolojili ve katma değerli uygulamalara olanak tanımaktadır.

Kent ölçeğinde ve yüksek çözünürlüklü bu veri kaynağının klasik yaklaşımlarla yönetilmesi ve işlenmesi pek mümkün olamamaktadır. Öte yandan derin öğrenme yaklaşımları büyük veriden modern hesaplama donanımları kullanılarak derin yapay sinir ağlarının eğitilmesini sağlayacak bir paradigma değişimi getirmiştir. Derin öğrenme sayesinde sınıflandırma, nesne yakalama ve segmentasyon görevlerinin veri işlem süreçlerinin otomasyonu amacıyla yerine getirilmesi için geçmişe nazaran çok daha güçlü çözümler elde edilebilir hale gelmiştir.

Bu çalışmada, 3B nokta bulutlarından derin öğrenme alanının gelişimi incelenmiş, bu modelleri kıyaslamada kullanılabilecek değerlendirme ölçütleri, araştırma grupları tarafından kabul görmüş açık veri setleri, 3B nokta bulutlarını işlemek için mevcut açık kaynaklı derin öğrenme yazılım kütüphanesi, performans yarışmaları tanıtılmıştır. 3B nokta bulutlarının semantik segmentasyonu için yararlanılabilecek modern bulut bilişim olanakları araştırılmıştır. Eğitim sürecini yürütmek için gereksinimler belirlenmiş ve çözüm önerileri geliştirilmiştir. Bulut bilişim imkanları kullanılarak derin öğrenme için oluşturulan bir geliştirme ortamında bir segmentasyon modelinin eğitim ve test süreci uygulama olarak gerçekleştirilmiştir.

Yapılan literatür taraması ve yürütülen derin öğrenme ağının eğitimi süreci sonucunda 3B nokta bulutlarından derin öğrenme çalışmalarında 3B şehir modelleme ve diğer pek çok farklı uygulama alanının 3B uzamsal veri gereksinimini gerçek zamanlı olarak dahi karşılayabilecek çözümler geliştirildiği görülmüştür. Bu çözümlerden yararlanmak için ise yüksek performanslı hesaplama donanımlarına erişime ve açık

kaynaklı derin öğrenme yazılım kütüphaneleri ile 3B nokta bulutundan derin öğrenme çalışmalarında geliştirilen yazılımlardan oluşan bir dizi teknolojiyi uyumlu biçimde bir araya getirme becerisine sahip olunması gerektiği görülmüştür.



SEMANTIC SEGMENTATION OF 3D POINT CLOUDS USING DEEP LEARNING METHODS

SUMMARY

3D city models play an integral role for enabling data driven decision making processes in spatial planning field by providing an infrastructure for urban monitoring and simulation applications such as 3D change detection, utility planning, crisis and emergency management, estimation of solar irradiation, energy demand estimation. Generation of such models require large scale, accurate and up-to-date 3D data at different levels of detail. Advancements on acquisition techniques like laser scanning, imaging, earth observation and data processing fields of photogrammetry and computer vision facilitated collection of large amounts of accurate 3D data rapidly in the form of point clouds by direct acquisition as well as generation from other resources like stereo images and alike. Today 3D points clouds, while enhancing their well-known use cases of surveying and planning, are enabling, enabling and accelerating more value added high tech sectors like autonomous mobility, robotics, 3D modeling and digital twin building. This kind of large scale and high resolution data arise challenges for both manual and automated analysis using classical approaches. On the other hand advent of deep learning introduced a paradigm shift in terms of handling and using of large scale data by utilizing them with increased computing power of modern hardware to train deep neural networks for obtaining much more potent descriptors than prior handcrafted counterparts for classifying, detecting and segmenting inputs to automate data analysis pipelines.

In this study, a historical review of recently developing field of deep learning for 3D point cloud analysis has been presented, evaluation metrics, popular open datasets, available open source deep learning libraries for processing 3D point clouds and benchmarks have been introduced. Modern cloud computing capabilities that can be utilized for training 3D point cloud deep learning models have been inspected. Requirements for establishing a healthy training process have been detected and possible solutions have been suggested. Practical application of training a state-of-the art solution concerning 3D semantic segmentation have been conducted by utilizing cloud development environment that have been architected.

Conducted literature review and executed deep learning training process showed that although there exists deep learning solutions that may accomodate for 3D spatial information requirements for 3D city modeling and other application areas, even in real time. Utilizing from those came out to be heavily dependent upon being able to access high performance computing hardware and ability to formulate a solution regarding integration of different software and tech stacks concerning open source deep learning frameworks and specialized software components.



1. GİRİŞ

Günümüzde artan nüfus ve kentleşme oranı, (United Nations, 2019); karar vericilere, şehirlerin insanoğlunun karşılaşmakta olduğu iklim değişikliği, kirlilik, salgın hastalıklar gibi çevresel problemlere karşı dayanaklı (resilience, resilient cities) ve 1987 Burtland raporunda tanımlandığı üzere “bugünün gereksinimlerini yarının kaynaklarını tüketmeden karşılama” (Al-Athel ve diğ, 1987) anlayışına uygun olarak sürdürülebilir biçimde planlanması ödevini yüklemektedir. Mekânsal planlamanın bu anlayışa uygun olarak gerçekleştirilebilmesinde akıllı şehir ve veri temelli karar alma süreçleri; ihtiyaç ve hedeflerin tespiti, öneri ve teorilerin test edilmesi açısından katkı sağlayabilir (van Maarseveen ve diğ, 2018). Fakat tüm bu fonksiyonların yerine getirilebilmesi, gerekli bilgilerin mekân ile ilişkilendirilerek sorgulanıp sunulabilir biçimde yönetilmesini zorunlu kılmaktadır. Bu ihtiyacın karşılanması amacı ile şehir modellerinin geliştirilmesi önemli bir çalışma alanı olarak ön plana çıkmıştır (Billen ve diğ, 2014). Günümüzde 3B şehir modelleri; görselleştirme amaçlarının ötesinde; topografya, yapı geometrileri, ulaşım ve altyapı hatları, şehir mobilyaları, sosyo-ekonomik veriler, çevresel bilgiler gibi geometrik ve geometrik olmayan veri bileşenlerini entegre edebilen ve kent-çevre ilişkisine dair analizler yürütülmesine olanak sağlayan bilgi sistemleri olarak tasarlanmakta ve uygulanmaktadır (Biljecki ve diğ, 2015).

Çağdaş şehirlerin yatay ve düşeydeki hızlı değişimleri, kent-çevre ilişkisini modellemede kullanılan 3B şehir modellerinin asıllarına uygun biçimde oluşturulması ve güncel halde tutulması için geniş çapta, yüksek çözünürlükte ve doğrulukta 3B mekansal bilgi gerekmektedir. Lazer tarama, görüntüleme, yer gözlem uyduları gibi uzaktan algılama tekniklerinin yaygınlaşması ile fotogrametri ve bilgisayarlı görüş alanlarında yaşanan ilerlemeler sayesinde kent ölçeğinde büyük alanlar için gereksinim duyulan nitelikte 3B uzamsal veri nokta bulutu formatında doğrudan ölçme yolu ile ya da başka kaynaklardan türetilerek hızlıca üretilebilmektedir (Ying ve diğ, 2020). Nokta bulutları, ilgi nesnelere ait geometrileri yüksek doğruluk, çözünürlük ve aslına uygunlukla modellemeye olanak tanısa da aynı zamanda karakteristik

özelliklerinden dolayı veri yönetimi ve işleme bakımından klasik yöntemlere göre bazı zorluklar ortaya çıkarmaktadır (Park ve Guldman, 2019).

3B nokta bulutu elde etme yöntemlerinin yaygınlaşmasıyla bu veri türü daha kolay erişilebilir hale gelmiş (Ivan ve diğ, 2017), özel sektör ve kamu kurumları elinde, manuel veya klasik otomatik yöntemlerle etkin biçimde kullanılması mümkün görülmeyen boyutta veri kaynağı oluşmaya başlamıştır (Büyüksalih ve diğ, 2019). Büyük miktarlarda bulunan bu uzamsal veri kaynaklarından modern yöntemler ile katma değer üretilmesi için bulut bilişim imkanlarının araştırılması ve otomatik veri işlem süreçlerinin geliştirilmesi önemli bir çalışma alanı olarak ön plana çıkmıştır (Shekhar ve diğ, 2012).

Nokta bulutları, çevresel/mekansal farkındalık gerektiren uygulama alanları için değerli bir bilgi kaynağı olarak değerlendirilmektedir. Bu veri tipinden hızlı biçimde katma değerli uygulamalarda faydalanmak için veri işlem süreçlerinin otomasyonu önemli bir çalışma alanı olarak öne çıkmıştır. Klasik yaklaşımların yanı sıra (Vosselman, 2013) güncel bir makine öğrenmesi tekniği olan derin öğrenmenin görüntü ve zaman serisi tipindeki veriler üzerine başarısı, araştırmacıları bu yeni teknikten 3B nokta bulutları ile alakalı veri işlem süreçlerinde de faydalanmak için çözümler geliştirmek üzere cesaretlendirmiştir (Li ve diğ, 2021). Yalnızca birkaç yıl içerisinde bu alandaki çalışmaların 2B görüntü verileri için geliştirilen çözümlerin basit birer uzantısı olmaktan sıyrılıp 3B nokta bulutlarının karakteristik özelliklerine uygun özelleşmiş ürünlere doğru olgunlaştığı görülmüştür (Guo ve diğ, 2021).

Kent ölçeğindeki alanlar için 3B nokta bulutları hızlı üretilebilse de, ham verinin analize hazır hale getirilmesi için uygulanması gereken modelleme adımları, verinin boyutu ve karakteristik özellikleri nedeniyle veri işlem sürecinde bir darboğaz oluşturmaktadır (Zhou ve Di, 2020). Özellikle nesne geometrilerini elde etmeye yönelik modelleme adımlarının otomasyonunun emek, zaman ve maliyet bakımından fayda sağlayacağı düşünülmüştür. Bu amaç doğrultusunda büyük çapta mekansal veriyi işlemede kullanılabilecek modern veri işlem olanakları ve makine öğrenmesi çözümlerinin 3B nokta bulutları için uygulanabilirliği araştırılmıştır. 3B nokta bulutlarından derin öğrenme alanında yaşanan gelişmeler özetlenmiş, popüler açık veri kaynakları tanıtılmış, geliştirilen derin öğrenme çözümlerini kıyaslamada ve değerlendirmede kullanılabilecek ölçütler aktarılmış, konuya ilişkin uygulama ve başarımlar standartları belirleyen çalışmalar incelenmiş, açıklanmış, bulut bilişim

olanaklarının kullanıbilirliğini göstermek ve test etmek amacıyla 3B nokta ulutları için derin öğrenme çözümlerini konu alan özel iş süreçleri geliştirilmiş, çözümlerin başarımı yapılan uygulama kapsamında denenmiştir.



2. LİTERATÜR ARAŞTIRMASI

2.1 Nokta Bulutları

Nokta bulutları günümüzde; klasik olarak ölçme, mühendislik, mimarlık, arkeoloji ve şehir planlama gibi dalların modelleme, planlama ve kontrol faaliyetleri için gerekli 3B geometrik altlığın oluşturulmasında kullanılmaktadır. Daha yeni teknoloji uygulama alanları olan robotik ve otonom mobilite çalışmalarına çevresel farkındalık çözümleri sağlamak için 3B nokta bulutlarından faydalanılmaktadır (Li ve diğ., 2021).

Nokta bulutu; nesneleri, dış yüzeyleri üzerinden örneklenmiş çok sayıda nokta ile geometrisini temsil etmeyi sağlayan bir veri tipi olarak tanımlanabilir (Linsen, 2001). Günümüzde gelişen teknoloji, yukarıda bahsedilen uygulama alanları ile bu çalışmanın teması olan 3B şehir modellerini oluşturmak ve güncellemek için gerekli mekânsal bilginin toplanmasını kolaylaştıran; lazer tarayıcılar, sayısal görüntüleme donanımları, yer gözlem uyduları verilerinin erişilebilirliğini arttırmıştır. Nokta bulutları doğrudan ölçme yöntemleri ile ya da başka kaynaklardan türetme yöntemi ile üretilabilmektedir (Xie ve diğ., 2020). Bu geniş olanaklar sayesinde hızlı ve uygun maliyetli biçimde yüksek çözünürlük ve doğruluğa sahip nokta bulutu verisine ulaşılabilmektedir (Wei, 2008). Nokta bulutları; 3B koordinatlar haricinde, radyometrik bilgileri de içerebilir. Dahası günümüzde sayısal kamera ve lazer tarayıcı entegrasyonundan oluşan sistemler de bulunmaktadır. Bu sistemler 3B koordinat bilgisine ek özellikler nokta bulutu verilerinden obje geometrilerinin yanı sıra doku gibi özelliklerin de çıkarılabilmesini sağlamakta, otomatik sınıflandırma ve segmentasyon işlemleri için referans miktarını artırmaktadır (Pirotti ve diğ., 2013).

Doğrudan nokta bulutu üretmek için en yaygın kullanılan ölçme aracı, LIDAR (Light Detection and Ranging) tekniğinden faydalanan lazer tarayıcılardır. Farklı ölçme amaçları için çeşitli platformalara entegre lazer tarayıcı çeşitleri ortaya çıkmıştır. Çizelge 2.1 de yaygın olarak kullanılan nokta bulutu üretim yöntemleri ile ilgili güçlü ve zayıf yönleriyle beraber kullanım alanlarına örnekler de sunulmuştur.

Yersel lazer tarayıcılar (Terrestrial Laser Scanner, YLT) ölçme esnasında statik olarak kullanılır. Mobil lazer tarayıcılar (MLT) özellikle yol çevresinin modellenmesinde tercih edilir ve otomobil gibi hareketli platformlarda taşınırlar. El ile ve sırt çantasında taşınabilen modeller de bulunmaktadır. Hava Lazer Tarayıcılar (Aerial Laser Scanner,

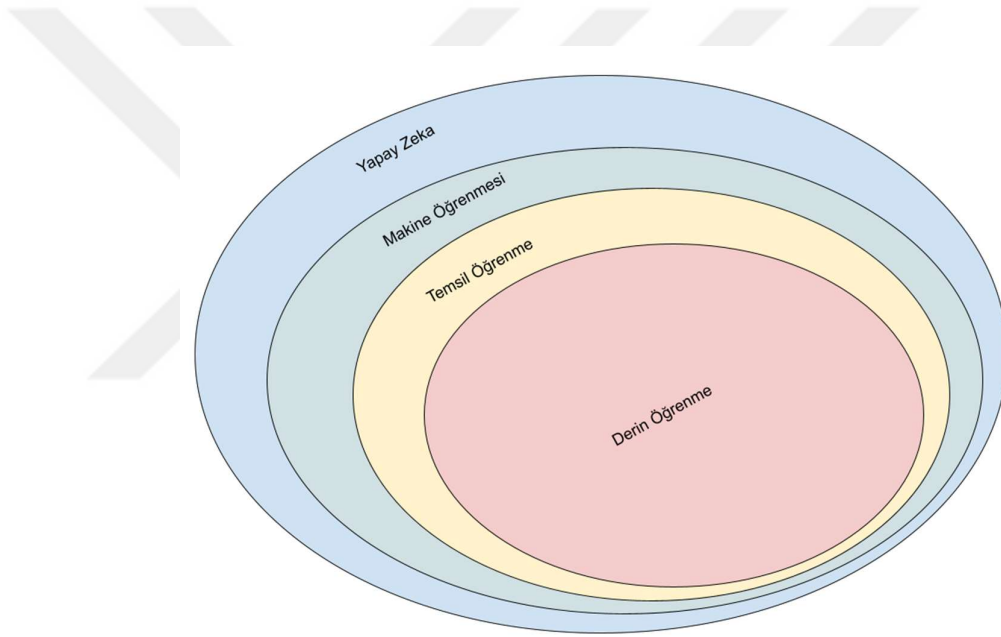
HLT), çok büyük alanların hızlıca 3B biçimde ölçülmesi için tercih edilen ve insansız hava aracı, helikopter uçak gibi çok farklı türdeki platformlarda taşınabilen bir ölçme aracı olarak ön plana çıkmaktadır.

Çizelge 2.1 : Çeşitli nokta bulutu üretim yöntemlerine yönelik bilgiler. (Xie ve diğ, 2020)’dan uyarlanmıştır.

	Avantajlar	Dezavantajlar	Uygulama Alanları
Görüntü kaynaklı	Hem geometrik , hem multispektral bilgi sağlayabilir. Havadan ve uzaydan geniş alanlar için elde edilebilir. Otomatik biçimde üretilebilir.	Veri kalitesi; hava şartlarına, yöneltme elemanlarının kalitesine, İlgili yüzeylerin karakteristiğine, stereo geometri oluşturmak için kullanılan algoritmaların kısıtlılıklarına bağlı.	Vejetasyonun izlenmesi, kentsel alanların izlenmesi, 3B modelleme, Robotik ve diğerleri.
LIDAR			
YLT	Yüksek çözünürlük ve doğruluk. (~mm)	Büyük alanların ölçülmesi için çok sayıda alım noktasında ölçmeye ihtiyaç duyulur. Maliyetli bir üretim tekniğidir.	Yüksek çözünürlüklü 3B modelleme. Arkeoloji, mimari, endüstri ve diğerleri.
MLT	Hareketli biçimde yüksek çözünürlüklü ve doğruluklu veri üretebilir. (~cm)	Veri kalitesi, algılayıcı sistemin kalibrasyonuna bağlı. Maliyetli bir üretim tekniği.	Kentsel modelleme, Yol durumunun takip edilmesi, Otomom mobilite ve diğerleri
HLT	Büyük alanların hızlı biçimde yüksek doğrulukla ölçülebilmesini sağlar (~10 cm)	Çözünürlük alım mesafesine bağlıdır. Maliyetli üretim tekniği.	Kentsel alanların izlenmesi, Ormancılık, vejetasyonun izlenmesi ve diğerleri
RGB-D	Diğerlerine göre daha ucuz algılayıcı imkanı	Diğer yöntemlere nazaran daha kısa menzil	İç plan 3B modelleme ve robotik. İnsan poz kesitimi
InSAR TomoSAR	Global kapsama alanı, zamansal çözünürlük, hava durumundan bağımsız alım	Uzmanlık isteyen ön işleme adımı gerektirir.	Kentsel alanların izlenmesi, orman alanlarının izlenmesi, vejetasyonun izlenmesi.

2.2 Yapay Öğrenme

Son yıllarda derin öğrenme çalışmalarının; bilgisayarlı görü ve doğal dil işleme alanlarında insan yeteneklerine yakın çözümler üretebilecek seviyeye ulaşmasıyla başlayan popülerite neticesinde hem akademide hem medyada yapay zeka (AI), makine öğrenmesi ve derin öğrenme kavramları birer yaygınlıkla olarak kullanılmakta; vaatler, beklentiler, endişeler ve spekülasyonlar yayınları doldurmaktadır (Chollet, 2017). Bu bölümde yukarıda bahsedilen 3 kavramın birbirleri ile ilişkileri kısaca açıklanmaya çalışılmış ve derin öğrenme ile ilgili bilgi verilmiştir.



Şekil 2.1 : Yapay Zeka, Makine Öğrenmesi ve Derin Öğrenme kavramlarının birbirleri ile olan ilişkisini ifade eden Venn şeması.

2.2.1 Yapay zeka

19. yüzyılda Charles Babbage isimli İngiliz matematikçi, matematiksel işlemlerde yardımcı olması için “Analytical Engine” adı ile bilinen ve ilk bilgisayar tasarımlarından biri olan eseri üzerinde çalışmalarına devam ederken aynı dönemin matematikçilerinden ve ilk bilgisayar programcılarının biri olarak gösterilen arkadaşı Ada Lovelace, bu makine için notlarında “...ona nasıl yapılacağını emretmeyi bildiğimiz her şeyi yapabilir. Görevi zaten aşına olduğumuz işlerde bize yardımcı olmaktır.” yazarak icadın sınırlarını tespit etmiştir (Goodfellow ve diğ, 2015).

Alan Turing 1950 tarihli ünlü Turing Testi'ni dünyaya kazandırdığı ve eğitilebilir bir bilgisayarı tarif ettiği “Computing Machinery and Intelligence” çalışmasında Ada Lovelace'in yukarıda bahsedilen alıntısına atıfla bilgisayarların bildiklerimizi yansıtmının ve bize yardımcı olmanın ötesinde insan becerilerini öğrenmeye, orijinallığe ve insanları şaşırtma kabiliyetine kavuşup kavuşamayacağını sorgulamıştır (Turing ve Haugeland, 1950).

Yapay Zeka çalışmaları 1940'lı yıllarda bilgisayarların “insanlar tarafından gerçekleştirilen entelektüel görevleri yerine getirmek için kullanılabilirliğini sağlama” amacı ile yürütölen araştırmalar ile başlamıştır. Bu tanım bilgisayarların insan becerilerini öğrenmesini ön gördüğünden yapay zeka kavramı Şekil 1.1'de de görölebileceğı gibi makine öğrenmesi ve derin öğrenme başlıklarını kapsayan ana alan olmaktadır (Chollet, 2017).

İnsanlar günlük yaşamlarında karşılaştıkları görevleri bilgi birikimlerini kullanarak çözerler. Bu sebeple başlangıçta araştırmacılar yeterince büyük bilgi bankaları ve bu bilgiyi manipöle edebilecek kural takımlarını oluşturmayı hedefleyen çalışmalar yürötmüşlerdir. 1950'lerden 1980'lerin sonuna kadar süren dönemde bu yaklaşımla ortaya konan çalışmalar “sembolik yapay zeka” olarak adlandırılmıştır. Bu anlayışla tasarlanan sistemler satranç gibi kuralları detaylı biçimde belirli ve mantıksal olarak açıkça ifade edilebilir problemlerin çözümünde kullanışlı olmuşlardır. Bu dönemde geliştirilen sistemler insanlar tarafından kolayca gerçekleştirilen fakat matematik-mantık kuralları çerçevesinde kolayca tarif edilemeyen görevlerde yeterli başarıyı gösterememişlerdir. Bu durum sonucunda problemlerin çözümünde yeni bir anlayışa gerek duyulmaya başlanmıştır (Goodfellow ve diğ, 2015).

2.2.2 Makine öğrenmesi

Makine öğrenmesi; bilgisayarların, kendilerine sunulan görevleri yerine getirmede insanlar tarafından titizlikle hazırlanıp formalize edilerek sunulan bilgi bankalarının ötesinde bu görevlerin nasıl gerçekleştirileceğini öğrenmesinin hedeflendiğı bir çalışma alanıdır.

Makine öğrenmesi kavramının ortaya çıkmasıyla problemlerin çözümünde kullanılacak sistemlerin programlanarak değil bir referans veri seti kullanılarak yürötülen eğitim süreçleri ile geliştirilmesine çalışılmıştır.

İdeal olarak, bir makine öğrenmesi sistemi eğitildikten sonra elde edilen kural setinin kullanım aşamasında yeni karşılaşılan veriye karşılık doğru çıktıları ortaya koyması ile iş süreçlerinde otomatizasyon hedeflenmiştir (Chollet, 2017).

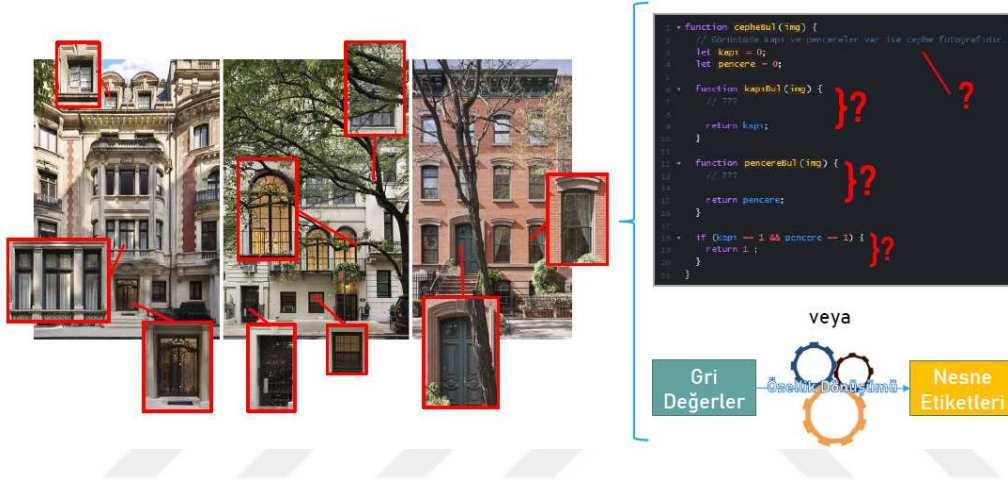
Makine öğrenmesi çalışmalarında, sunulan girdilere göre beklenen çıktıları meydana getirecek ideal kural setlerinin otomatik biçimde elde edilmesi hedeflenir. Bilgisayar ortamında girdi ve çıktılar arasındaki ilişkiyi ölçüp ifade edebilmek için sayısal büyüklüklere ihtiyaç duyulmaktadır. Bu sebeple makine öğrenmesi çalışmalarında girdi ve çıktı arasındaki ilişkiyi matematiksel ve istatistiksel yöntemler ile ortaya çıkarabilecek modeller tasarlanmaya çalışılmıştır. Makine öğrenmesi modelleri genel olarak girdi veriyi beklenen çıktılarına dönüştürmeyi başaracak fonksiyon ve ilişkili parametrelerden oluşmaktadır. Bu fonksiyon ve parametre setlerinin seçilmesi, ilgili parametrelerin otomatik biçimde öğrenilmesinin sağlanması ana araştırma alanları olarak ön plana çıkmıştır (Goodfellow ve diğ, 2015).

2.2.3 Temsil öğrenme

Bilgisayar ortamında bilgiler kullanım amaçlarına uygun biçimde standartlaştırılmış veri yapılarında tutulur ve temsil (representation) edilirler. Örneğin görüntüler RGB ya da HSV gibi formatlarda, nokta bulutları 3 boyutlu koordinatları ile, kurum ve kişi gibi gerçek dünya varlıklarını ifade eden kayıtlar veri tabanlarında öznetelikleri ile bulundurulup temsil edilirler. Tercih edilen temsil biçimi hedeflenen kullanım amaçları için kolaylık ya da zorluk sağlayabilmektedir. Örneğin piksellere dair mavi bant gri değerleri RGB formatında daha kolay elde edilirken doygunluk değeri HSV formatında daha kolay sorgulanır. Benzer biçimde makine öğrenmesi modelleri, girdi verilerin mevcut temsillerinden, amaçlanan hedefler için daha uygun temsiller elde etmeye çalışırlar. Bu temsiller genellikle sayısal olarak ifade edilebilen özellik vektörleri olmaktadır.

Makine öğrenmesi teknikleri, bilgisayar ortamında tutulan verilerin otomatik olarak değerlendirilmesini sağlayacak sistemlerin geliştirilmesinde kullanılmaktadırlar. Bu çalışmalara; görüntülerin sınıflandırılması ve görüntülerde bulunan objelerin tespit edilmesi, konuşma tanıma ile sesten yazıya aktarma, metin çevirme, kredi başvurularını değerlendirmede kişiye göre risk analizi, 3B nokta bulutlarından nesne geometrilerini elde etme gibi veri temelli örnekler gösterilebilir. Bu uygulamaların hepsinde son kararı vermede kullanılan model bileşenine en uygun temsil sunulmaya

çalışılır (Goodfellow ve diğ, 2015). Modeller eğitim boyunca girdi temsillerden beklenen çıktı temsilleri elde etmeyi sağlayacak dönüşümleri otomatik olarak gerçekleştirebilecek parametreleri ve/veya dönüşüm fonksiyonlarını kestirmeye çalışırlar. Bu dönüşümler; koordinat dönüşümleri, lineer veya non-lineer fonksiyonlar, projeksiyonlar ve daha pek çok çeşitte olabilirler. Modellerin bu süreçteki arayışları boyunca da rehberleri doğruluk ölçütleri ve geri besleme mekanizmaları olmaktadır (Chollet, 2017).



Şekil 2.2 : Makine öğrenmesi ve temsil öğrenme kavramlarını açıklamaya yönelik örnek Url-1’den uyarlanmıştır.

Makine öğrenmesi amaçları için uygun temsilleri elde etme yolunda farklı makine öğrenmesi yöntemleri geliştirilmiştir. Bunlar arasında Naive Bayes, lojistik regresyon gibi olasılıksal yöntemler, yapay sinir ağları, destek vektör makineleri (SVM), karar ağaçları, rastgele orman algoritmaları ve gradient boost makineleri bulunmaktadır. Sembolik yapay zeka çalışmalarında olduğu gibi el ile dizayn edilmiş temsiller gereksinim ve kuralları iyi anlaşılmış ve açıkça ifade edilebilen ödevlerde faydalı sonuçlar üretebilmektedir. Fakat obje yakalama ya da konuşma algılama gibi daha soyut ve matematiksel açıdan ifade edilmesi güç çalışma alanlarında kullanılabilir özellik vektörlerini tasarlamak zordur (Goodfellow ve diğ, 2015). Bu konuya ilişkin olarak şekil 2.2’de gösterildiği gibi bir sokak görüntüsü üzerinde evleri otomatik biçimde seçebilen bir program geliştirmeye çalışıldığı örnekleme yapılabilir. Kapı ve pencere gibi detayların varlığı ev nesnesinin varlığı ile ilişkilendirilebilir fakat bu detayları gri değer cinsinden ölçek, dönüklük ve öteleme ya da örtülme, eksiklik gibi durumlardan bağımsız biçimde ifade edebilecek evrensel özellik vektörlerini ve kural setlerini manuel olarak meydana getirmek mümkün değildir. Bu durum girdi veri

üzerinden hem özellikleri hem de özelliklerin hiyerarşini oluşturabilecek makine öğrenmesi tekniklerine gereksinimi gözler önüne sermiş ve yeni çalışma alanlarına kapı aralamıştır.

2.2.4 Yapay sinir ağları

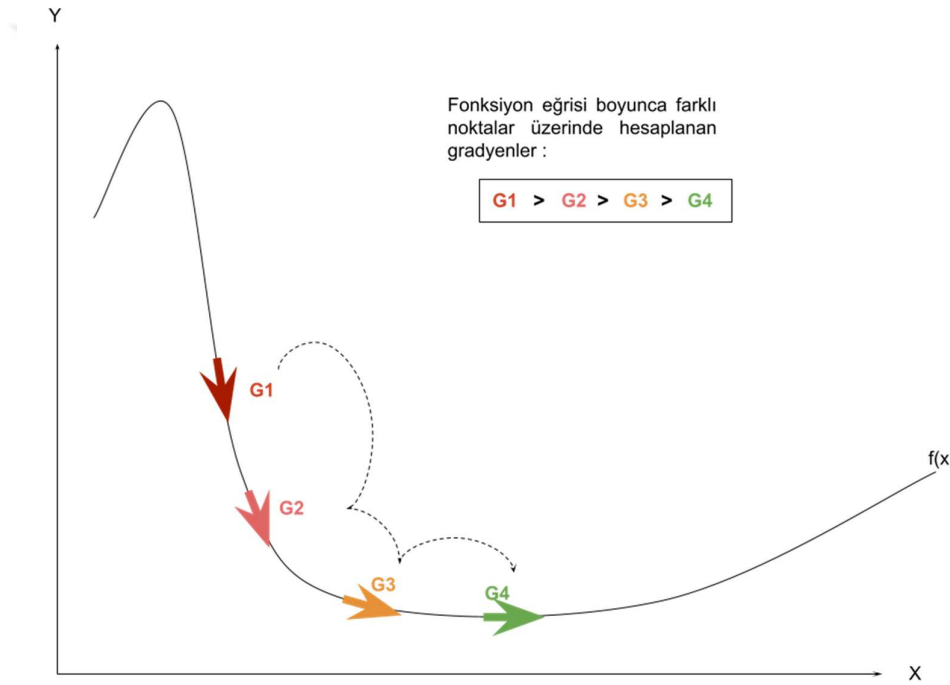
Yapay zeka araştırmacıları için beyin ve işlem birimi olan nöronların çalışma prensibi merak konusu olmuş, makine öğrenmesi amaçları için nöronları taklit edebilecek sistemler geliştirilmeye çalışılmıştır (Kothari ve Oh, 1993).

Biyolojik nöronlar temel olarak; dendrit adı verilen uzantıları ile komşu nöronların akson adı verilen uzantılarından gelen elektrik sinyallerini sinaps adı verilen biyokimyasal olay ile alıp işleyerek aksonları ile diğer nöronlara aktarırlar. Bu mekanizmayı, makine öğrenmesi amaçları için taklit etmeyi hedefleyen çalışmalar 1940’larda başlamıştır. 1949 yılında, nöronlar arası iletişimde kullanılan bağlantıların güçlendiğinin ortaya konması ile model nöronların birlikte çalışarak eğitilmesine dayanan yapay sinir ağları (YSA) araştırmaları hızlanmıştır (Goodfellow ve diğ, 2015).

İlk bilinen eğitilebilir yapay sinir ağı, 1950’lerde Rosenblatt tarafından geliştirilen ve ‘Perceptron’ olarak adlandırılan model nöron elemanından oluşan bir tek katmanlı algılayıcı olarak ortaya çıkmıştır. ‘Perceptron’ modeli olumlu ve olumsuz olarak ayırt etmesi gereken örnekleri lineer bir sınıflandırıcı ile ayırt etmek üzere parametreleri güncellenerek eğitilebilmekteydi (Rosenblatt, 1958). Aynı yıllarda, Widrow ve Hoff tarafından geliştirilen ‘Adaptive Linear Element’ (ADALINE) tek katmanlı algılayıcısı ve ADALINE birimlerinin arka arkaya eklenmesinden oluşan ‘Multi Layer ADALINE’ (MADALINE) modeli modern meyil azalımlı öğrenme tekniği ile eğitilebilen yapay sinir ağlarının ilk örnekleri olarak ortaya çıkmıştır (Goodfellow ve diğ, 2015).

Meyil/eğim azalım (gradient descending) yöntemi, tasarlanmış bir hedef fonksiyonun iteratif biçimde optimize edilmesi için geliştirilmiş bir tekniktir. Hedef fonksiyonun bir maksimum ya da minimum noktasına ulaşmasını sağlamak için, fonksiyonu oluşturan parametreler kısmi türevleri oranında güncellenir. Bir fonksiyon optimum değerine yakınsamaya başladıkça yönlü türevi diğer bir deyişle gradyeni de sıfıra yakınsayacaktır. Bu durumun sonucunda Şekil 2.3’de de görülebileceği gibi hedef fonksiyonun eğrisi üzerinde gittikçe küçülen gradyenler elde edilir (Rumelhart ve diğ,

1986). Stokastik meyil azalımı (stochastic gradient descent) yöntemi ise aynı yaklaşımın tüm veri seti yerine rastgele seçilen bir alt kümesi kullanılarak uygulanmasına verilen isimdir. Stokastik meyil azalımı yönteminin geliştirilmesindeki amaç verisetinin, yeterli büyüklükte rastgele seçilmiş bir örnekleminin benzer dağılım özellikleri göstererek tümünü temsil edebileceği düşüncesinden hareketle işlem yükünü azaltmaktır (Bottou, 1998). Benzer yaklaşımlar makine öğrenmesi alanında hata fonksiyonunu minimize etmek için en yaygın kullanılan optimizasyon yöntemleri arasındadır.



Şekil 2.3 : Örnek bir fonksiyon üzerinde meyil azalımlı optimizasyon algoritmasında gradyenler.

Yapay sinir ağları ile girdi temsillerden, eğitim boyunca öğrenilen non-lineer dönüşümler vasıtasıyla sınıflandırma gibi ödevler için daha kullanışlı temsilleri ifade eden özellik vektörlerinin otomatik biçimde elde edilmesi hedeflenmektedir. Söz konusu özellik vektörlerini oluşturabilmek için ağın hareketli parçaları olan katsayıların uygun değerlere ulaşması gerekmektedir. Bu değerler eğitim boyunca deneme yanılma yöntemi ile, hesaplanan ve beklenen çıktılar arasındaki fark ölçülüp bu farkı küçültücü yönde her bir ağırlık kendi anlık değeri uyarınca güncellenerek otomatik biçimde hesaplanır.

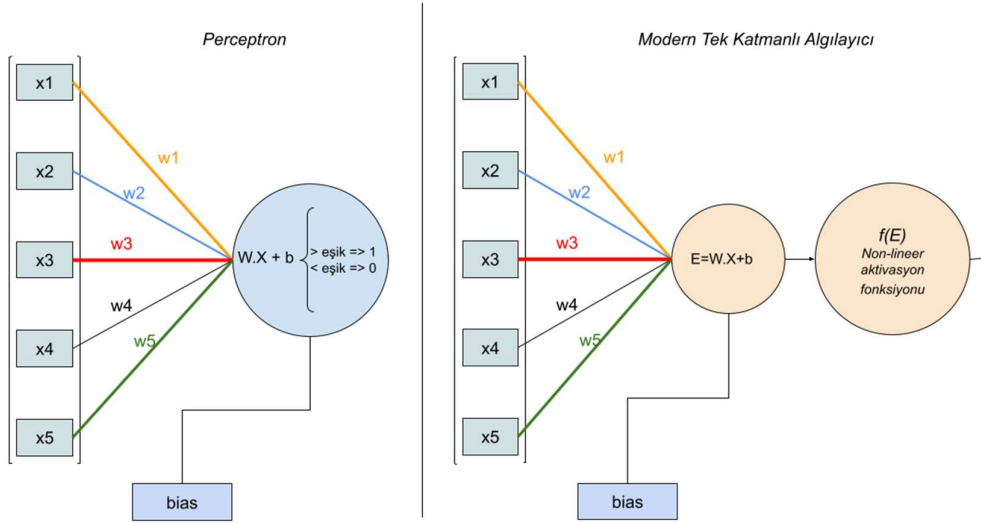
Ağırlıkları uygun değerlere getirebilmek için uygulanması gereken düzeltmelerin büyüklükleri ve yönleri türevden faydalanarak hesaplanır. Bir fonksiyonun, parametrelerinin diferansiyel anlamdaki değişmelerine duyarlılığı söz konusu parametreye göre kısmi türevi kullanılarak hesaplanabilir.

Modern geri yayılım algoritmaları; yapay sinir ağlarını, katmanlardan oluşan bileşke fonksiyonlar, ağırlıkları ise bu fonksiyonların parametreleri olarak kabul ederek her iterasyonda düzeltmeleri zincir kuralından faydalanarak geriye doğru yayma yöntemini esas alırlar. Gerekli kısmi türevlerin hesaplanabilmesi için ilk olarak kullanılan hata fonksiyonun türevi alınır ve zincir kuralı ile ağ boyunca işlem devam eder. Meyil Azalımlı Eğitim adı verilen bu yaklaşımda amaç hata fonksiyonun türevi sıfıra yakınsayan bir noktaya getirilmesini sağlayarak minimum hata değerine ulaşmaktır.

Bir optimizasyon yöntemi olan Meyil Azalımlı Eğitimde düzeltme büyüklükleri; hata fonksiyonunun anlık türevi, öğrenme hızı adı verilen bir katsayı ile ölçeklenerek ağda yayılır. Bu sayede öğrenme hızı kontrol edilerek ağırlıkların daha ince ayarla düzeltme almasını sağlamak mümkün olur. Bu eğitim türünde katmanlarda kullanılan aktivasyon fonksiyonuna ait türevin, fonksiyonun farklı değerlerindeki davranışı önemlidir. Özellikle belirli değerlerde küçülen türev, zincir kuralı boyunca çarpma işlemlerinde kullanıldığından ilk katmanlara ulaşan düzeltme değeri etkisiz denebilecek büyüklüklerde takılı kalabilmektedir. “Kaybolan Gradyen Problemi” ismi verilen bu durumun önüne geçebilmek için aktivasyon fonksiyonlarının dikkatle tasarlanması gerekmektedir.

Açıklanan mimarideki sistemlere ileri beslemeli geri yayımlı yapay sinir ağı adı verilmektedir. Bu açıdan bakıldığında bir yapay sinir ağı eğitim işinin veri ve hesaplama gücü yoğun bir süreç olduğu kolayca anlaşılabilir. Yapay sinir ağı kavramının ortaya çıktığı 1940’lardan beri ilerleyen teknoloji, bu tekniği zamanla daha uygulanabilir hale getirmiştir.

Yapay sinir ağı çalışmalarında tarih boyunca, nöron modelleri, öğrenme yöntemleri ve eğitim algoritmaları açılarından çeşitli algoritmik ilerlemeler yaşanmıştır. Şekil 2.4’te ‘Perceptron’ modeli ve 1970 ve 1980’lerde yapılan çalışmalar neticesinde şekillenen modern yapay sinir ağlarında işlem birimi olarak kullanılan nöron elemanlarının karşılaştırmalı bir görseli sunulmuştur.



Şekil 2.4 : Perceptron ve modern tek katmanlı algılayıcı modellerinin gösterimi.

1980'lerle birlikte yapılan çalışmalar neticesinde yapay sinir ağları gerçek dünya problemlerinde uygulanmaya başlamıştır. İlk pratik örnek Yann LeCun tarafından 1989 yılında el yazısı ile yazılmış rakamları görüntüler üzerinden sınıflandırmaya yönelik geliştirilen LeNet'in, ABD Posta Servisi tarafından 1990'larda kullanılmaya başlanması ile ortaya çıkmıştır (LeCun ve diğ, 1989). Modern yapay sinir ağı modelleri, girdi özellik vektörü elemanlarını ölçeklemede kullanılan parametreler olan ağırlıklar, bias değişkeni ve girdi verinin non-lineer bir dönüşümünü sağlamak için kullanılan aktivasyon fonksiyonundan meydana gelen nöron birimlerinden oluşmaktadır.

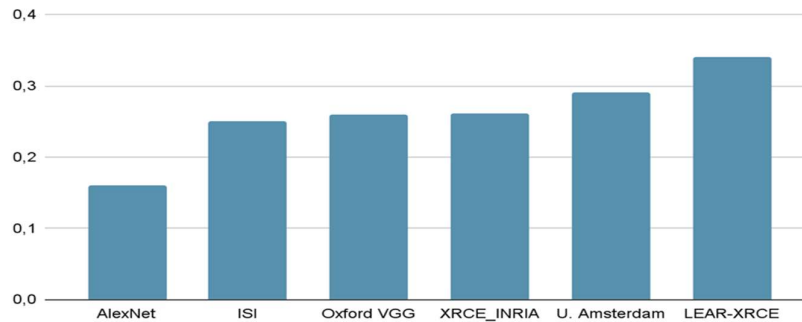
2000'li ve 2010'lu yıllarda internet ve işlemci teknolojilerinde yaşanan gelişmeler yapay sinir ağlarının gereksinim duydukları veri ve hesaplama yeteneklerini ulaşılabilir kılmıştır. Bu verimli ortam sayesinde yapay sinir ağları gitgide artan sayıda katman sayıları, eğitim verisi miktarı, iyileştirilen geri yayılım algoritmaları ve aktivasyon fonksiyonlarına kavuşmuşlardır. Önceki versiyonlarına göre daha başarılı bu çok katmanlı mimariler, derin öğrenme ismi ile anılmaya başlamıştır (Goodfellow ve diğ, 2015).

2.2.5 Derin öğrenme

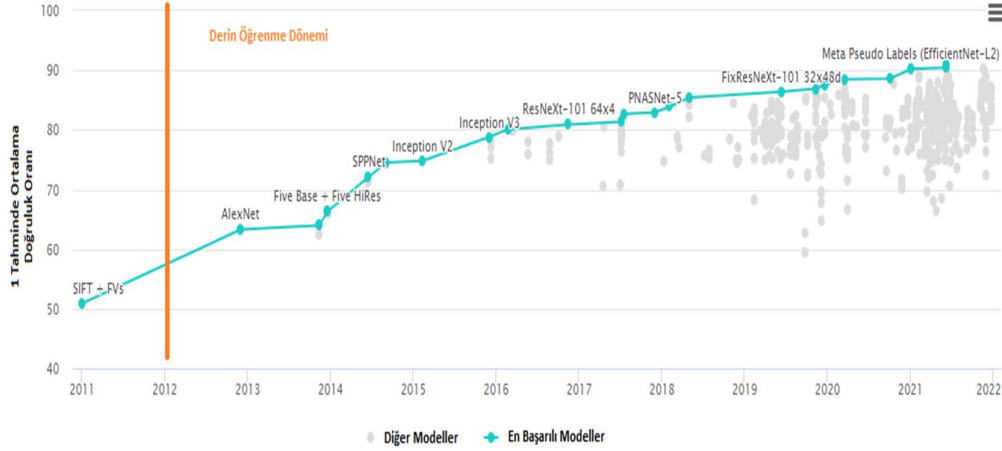
Derin öğrenme temellerini yapay sinir ağlarından alan bir makine öğrenmesi tekniğidir. Kavramın isminde geçen derinlik terimi modellerin içerdiği katman sayısına bir referans olarak kullanılmaktadır. Bu yaklaşımda amaç çok sayıda sıralı

katmandan oluşan modeller kullanarak sınıflandırma, segmentasyon ve obje yakalama gibi veri işleme görevleri için daha kullanışlı olduğu düşünülen bir temsiller hiyerarşisi elde etmektir.

Derin öğrenme çalışmaları, özellikle 2010'lu yıllarda, çok katmanlı yapay sinir ağları alanındaki algoritmik gelişmeler, farklı veri türleri için özelleşmiş mimari ve ağ topolojilerinin ortaya çıkması, eğitim verisi olanaklarının artması ve bu ağları eğitmek için gerekli hesaplama gücünü sağlayabilecek modern genel kullanıma uygun grafik hesaplama birimlerinin (ekran kartı) erişilebilir hale gelmesi ile büyük oranda hızlanmıştır (Chollet, 2017). 2011'de Dan Ciresan'ın modern ekran kartları kullanarak eğittiği derin yapay sinir ağları akademik görüntü sınıflandırma yarışmalarını kazanmaya başlamıştır (Ciresan ve diğ., 2011). Derin öğrenmenin popülerliğini başlatan esas kırılma noktası ise Krizhevsky ve ekibinin geliştirdiği derin evrimsel yapay sinir ağının (Deep Convolutional Neural Network, DCNN) (Krizhevsky ve diğ., 2012) 2012 yılında, o zamanlar başarılı olmanın oldukça zor kabul edildiği, ImageNet Large Scale Visual Recognition Challenge (ILSVRC) yarışmasını Şekil 2.5'de de görülebileceği gibi rakiplerine göre büyük oranda farklı bir performans göstererek kazanması olmuştur. ImageNet binlerce sınıftan sayısı milyonları bulan ve farklı çözünürlüklerde etiketli görüntüyü içeren açık kaynaklı bir veri setidir (Russakovsky ve diğ., 2015). AlexNet ya da SuperVision adı ile anılan bu model, makine öğrenmesi ve yapay sinir ağları ile görüntü işleme alanında önceden geliştirilmiş yöntemleri, eğitim aşamasını kolaylaştırıcı fikir ve tekniklerle harmanlayıp döneminin veri zenginliği ve modern ekran kartlarının hesaplama gücünden faydalanarak başarı elde etmiştir. Şekil 2.6'da da görülebileceği gibi ilk ismi SuperVision olan AlexNet'in sahneye çıktığı 2012 tarihinden sonraki her ILSVRC yarışmasını farklı bir derin evrimsel yapay sinir ağı kazanmıştır.



Şekil 2.5 : 2012 ILSVRC yarışmasındaki en başarılı 6 modelin, her bir örneğe ait ürettikleri 5 tahmin sonucunda hesaplanan ortalama hata oranları.



Şekil 2.6 : 2011-2022 yılları arasında ILSVRC kazananları, Url-2'den alınmıştır.

AlexNet mimarisi temellerini, görüntü işleme alanında 1980'de tarihli Neocognitron (Fukushima, 1980) ve 1989'da tarihli LeNet çalışmalarından almaktadır. Fakat bu tekniklerin ILSVRC ve benzeri yarışmalarda; sınıflandırmadan, obje yakalama ve segmentasyona kontrollü laboratuvar ortamlarında üretilmiş basit verilerin dışında geniş çaplı bilgisayarlı görü problemlerinin çözümünde kullanılabilir hale gelmesi için hem bilgisayarların işlem gücünde hem eğitim verisi miktarında hem de eğitim sürecinde kullanılan algoritmalarda uzun yıllar süren gelişmeler yaşanmasına ihtiyaç duyulmuştur. Donanımsal ihtiyaçların karşılanmasında yüksek performanslı modern grafik işlem birimlerinin ortaya çıkması etkili olmuştur. Bu işlemciler saniye başına işlem sayısı ve hafızalarıyla yeterli derinlikte modellerin kabul edilebilir hızlarda eğitilebilmelerine imkan tanımıştır (Wang ve diğ, 2019). Veri ihtiyaçlarının karşılanmasında depolama donanımlarında yaşanan ilerlemeler ve internetin gelişip yaygınlaşmasıyla büyük veri setlerini toplayıp paylaşmanın kolaylaşması etkili olmuştur (Goodfellow ve diğ, 2015). Bu gelişmelerin yanı sıra daha verimli eğitim süreçlerini mümkün kılacak algoritmik ilerlemeler derin öğrenme ağlarının başarısını büyük oranda arttırmıştır.

Modern grafik işlem birimlerinin ortaya çıkışı, özellikle bilgisayarlı görü alanında standart çok katmanlı algılayıcılar yerine girdi veride bulunan yerel komşuluk alanlarının ilişkisel özelliklerini ortaya çıkarabilen temsilleri elde etmede faydalı evrimsel katmanların geçmişe nazaran artan derinliklerde kullanılmasına imkan sağlamıştır. Artan hesaplama kabiliyetinden, zaman serisi verilerini işlemede

kullanılan özyinelemeli yapay sinir ağıları (RNN) ve Long Short Term Memory (LSTM) (Hochreiter ve Schmidhuber, 1997) mimarisine sahip modellerde de faydalanılmıştır.

Eğitim sürecinin verimliliğini düşüren algoritmik engeller de geçmişte yapay sinir ağı çalışmalarını yavaşlatıcı etki göstermiştir. Özellikle derin yapay sinir ağlarının verimli biçimde eğitimini zorlaştıran temel algoritmik engellerden biri “Kaybolan Gradyen” problemi olmaktadır (Chollet, 2017). Bu probleme çözüm olarak geri yayılım adımıyla kısmi türev büyüklüklerini giderek zayıflamasına engel olan karakterde aktivasyon fonksiyonları kullanılmaya başlanmıştır.

“Rectified Linear Unit” (ReLU) (Glorot ve diğ, 2011) ve benzerleri bu amaçla kullanılan aktivasyon fonksiyonları arasında yer almıştır. Bu fonksiyonlar; türevleri bir, sıfır ya da uygun bir değeri üretecek biçimde konfigüre edilerek zincir kuralı boyunca çarpma işlemine tabi tutulan gradyenlerin sağlıklı büyüklüklerde kalması sağlanabilmektedir.

Yapay sinir ağlarının eğitiminde matematiksel çözümler, eğitim süreci hata fonksiyonun optimizasyonu problemi şeklinde ele alınarak geliştirilmektedir. Ağdaki ağırlıkların girdileri doğru çıktı temsillere dönüştürebilecek uygun değerlere hassas biçimde güncellenebilmesi için, daha önce bahsedilen stokastik meyl azalımı tekniği uyarınca, eğitim süresince sağlıklı düzeltmeler hesaplayabilecek optimizasyon algoritmalarının geliştirilmesi önemli bir çalışma alanı olmuştur (Goodfellow ve diğ, 2015). Bu amaçla Momentum (Qian, 1999), Adagrad (Duchi ve diğ, 2011) ve ADAM (Kingma ve Ba, 2015) gibi öğrenme hızını model bütününde ve/veya parametreler özelinde düzenleyebilen algoritmalar geliştirilmiştir. Bu alanda çalışmalar; gittikçe daha hızlı yakınsayabilen, hesaplama ve hafıza gereksinimleri yönünden verimli, çok parametrelili problemlerin çözümüne uygun (Ruder, 2016) optimizasyon tekniklerinin geliştirilmesi yönünde devam etmektedir.

Makine öğrenmesi çalışmalarında ana amaç eğitim verisinden öğrenilen kural setinin uygulama adımında yeni verilerle karşılaştığında da kullanılabilir olmasını sağlamaktır. Model gereksinimlerine yeter büyüklük ve çeşitlilikte eğitim verisi kullanılmadığı durumlarda “ezberleme” ismi verilen olumsuzluk ortaya çıkabilmektedir. Bu sebeple geliştirilen modellerin nihai başarısı eğitim süreci boyunca hiç gösterilmeyen test verileri kullanılarak ölçülür. Modellerin

genelleştirilebilirliği olarak ifade edilen yeni veri karşısındaki başarı oranını arttırmak için regularizasyon adı verilen teknikler geliştirilmiştir. Bu teknikler model parametrelerinin göreceli büyüklükleri ve katman çıktılarının dağılımları gibi bileşenleri düzenleyerek ağı genelleştirilebilirliğini, eğitim sırasında hesaplanan doğruluklarını düşürmek pahasına arttırmayı hedefler (Kukacka ve diğ., 2017).

Derin öğrenme modelleri için geliştirilen dropout (Srivastava ve diğ., 2014) ve batch normalizasyonu (Ioffe ve Szegedy, 2015) gibi algoritmalar test verisinde doğruluğu artırarak katkıda bulunmuştur. Derin öğrenme alanında algoritmik açıdan; başlangıç ağırlıklarını belirleme yöntemleri (Kumar, 2017) ve farklı mimariler geliştirme süreçleri yönlerinde çalışmalar da yapılmaktadır.

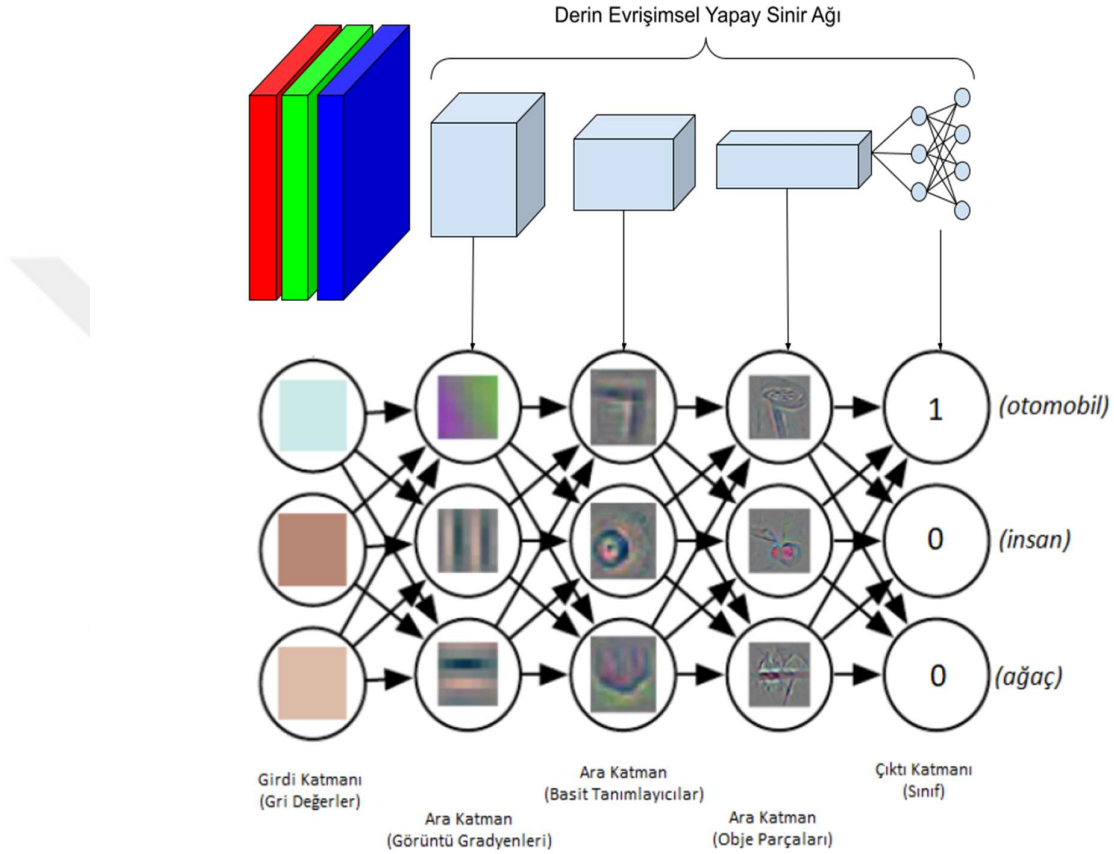
Derin öğrenme çalışmalarına motivasyon ve referans kaynağı olan fıkırsel temeller ve uygulamalar yapay sinir ağları üzerine araştırmaların arttığı 1970'lerde başlayan bağlantıcılık (connectionism) akımı döneminde önemli ilerleme kaydetmiştir (Goodfellow ve diğ., 2015).

Bağlantıcılık felsefesine göre tekil nöronlar zekadan yoksun olsa da sistematik şekilde beraber çalışan nöronlardan oluşan modeller zeki davranışlar gösterebilirler. Bu anlayıştan hareketle zaman içinde modern çok katmanlı yapay sinir ağlarından derin öğrenme ağlarına varan ürünler ortaya çıkmıştır. Bu süreçte ilerleyen karmaşıklıkta görevlere dair doğruluk oranlarının iyileşmesinde en önemli faktör, algoritmik ve teknolojik gelişmeler sayesinde, parametre sayısı ve ağ derinliklerinin artarak daha büyük modellerin üretilmesi ile daha kompleks temsillerin elde edilebilir hale gelmesidir.

Derin öğrenme, temsili öğrenme alanındaki önemli eksikliklerden birini gidermede etkili olmuştur. Artan parametre sayısı ve model derinliği, her katmanda girdi temsillerin kombinasyonu biçiminde daha kompleks temsillerin elde edilebilmesini mümkün kılmıştır. Bu sayede ağ boyunca artan derinlikte bir temsiller hiyerarşisinin elde edildiği ve daha önceleri soyut ve sezgisel bulunarak formülize edilemeyen problemlerin yüksek doğrulukla çözülebilmesine imkan tanındığı ileri sürülmüştür (Chollet, 2017).

Şekil 2.7'de görülebileceği gibi derin öğrenme ağlarının oluşturabildiği temsiller hiyerarşisi sistemi sayesinde önceleri veri bilimcileri tarafından emek ve zaman yoğun iş süreçleri sarf edilmesini gerektiren temsil tasarlama adımı terk edilerek çok temel

ön dönüşüm işlemleri ile standartlaştırılan ham veriyi işleyip, veri içerisinde bulunan özellikleri temelden karmaşığa biçimde elde etmeyi sağlayabilecek mimariler ortaya çıkmış, veri işleme ile ilgili iş ve süreçlerin tam otomasyonu yolunda önemli bir adım atılmıştır.



Şekil 2.7 : Temsiller hiyerarşisi kavramını ifade etmek üzere düzenlenmiş görsel (Goodfellow ve diğ, 2015).

Derin öğrenme teknolojisinin çeşitli araştırma alanlarından farklı seviyede çalışmacılara ulaştırılması çok daha geniş insan gruplarına katma değerli son ürünler ve hizmetler sunulmasını sağlayarak kalkınma yolunda faydalı olacaktır. Bu nedenle derin öğrenme teknolojisine erişimi, daha demokratik hale getirmek için çeşitli açık kaynaklı yazılım kütüphaneleri (Tensorflow (Abadi ve diğ, 2016), PyTorch (Paszke ve diğ, 2019), Keras (Chollet, 2018) vb.), bulut tabanlı hesaplama platformları (Google Colab, Google Cloud Platform, Amazon Web Services) ve eğitim programları geliştirilmesi çalışmaları akademi, büyük teknoloji firmaları, sivil toplum kuruluşları ve resmi kurumlar tarafından yürütülmekte ve desteklenmektedir.

Derin öğrenme çözümlerinin başarısı ile ön plana çıkan bir diğer konu etik sorunlar olmaktadır. Kullanılan eğitim veri setlerinin objektiflik derecesi, uygulama geliştirme amaçları, özel hayatın gizliliğin korunması gerekliliği, geliştirilen otonom ajanların cezai ehliyetleri gibi konular makine öğrenmesi çalışmalarında etik tartışmalarının yaşanmasına neden olmaktadır (Etzioni ve Etzioni, 2017). Mekansal bilgi dünyası alanında da geliştirilen uygulamaların ve onların dayanağı olan verilerin elde ediliş biçimi bakımından özel hayatın ve kişisel verilerin gizliliği gibi başlıklar da etik açıdan tartışılmaktadır. Literatüre bakıldığında etik yapay zeka konusunun küresel çapta genel olarak beş başlık üzerinde tartışıldığı görülmektedir. Şeffaflık, adillik, zararlı amaçlardan kaçınma, sorumluluk ve gizlilik başlıklarında tartışılan etik prensiplerde henüz uygulama açısından bir konsensüs bulunmamaktadır (Jobin ve diğ, 2019). Söz konusu durum neticesinde insanlığa katkı sağlama potansiyeli yüksek olan yapay zeka teknolojisinin etik sınırlar içinde kullanılması sorumluluğu şimdilik bu alanda çalışanlara düşmektedir.

3. 3B NOKTA BULUTLARI VE DERİN ÖĞRENME

Nokta bulutları, yaşadığımız çevre ile ilgili 3B mekânsal bilgi elde etmede önemli bir veri kaynağıdır. Günümüzde artarak daha erişilebilir hale gelen nokta bulutu verisi çevremizdeki nesnelerin geometrileri ile ilgili yüksek çözünürlükte bilgi elde edilebilmesini sağlamaktadır. Nokta bulutları, en küçük birimleri olan noktalar bazında ilişkili oldukları kavramsal sınıf ile etiketlenerek çok yüksek çözünürlüklü mekânsal modeller elde edilebilir. Bu yüksek çözünürlüklü bilgi, çevresel farkındalığa ihtiyaç duyan mobilite, robotik, 3B kentsel modelleme gibi alanlarda iş süreçlerinin otomatize edilmesinde kullanılabilir. Bu çalışmada, girdi veriden mümkün olan en yüksek çözünürlükte tematik bilgi edinilmesini sağlayan semantik segmentasyon işleminin, 3B nokta bulutları için derin öğrenme çözümleri ile gerçekleştirilebilirliği araştırılmıştır.

Nokta bulutları için semantik segmentasyon; noktaların, yer yüzünde rastladıkları doğru kavramsal sınıflar; örneğin yapı, yol, ağaç, zemin otomobil, şehir mobilyası, ile ilişkilendirilmesi işlemi anlamına gelir. Semantik segmentasyon, ham nokta bulutlarından katma değerli ürünler elde edebilmek için önemli bir çalışma alanı olmaktadır (Wang ve diğ, 2018). Bu bölümde ilk olarak nokta bulutlarının derin öğrenme yöntemleri ile işlenmesi ile ilgili zorluk ve engeller açıklanmaya çalışılmıştır. Devamında genel olarak nokta bulutlarının otomatik olarak analizi ve işlenmesi için geliştirilmiş ve kilometre taşı olarak kabul edilebilecek derin öğrenme modellerinden örnekler verilerek alanın tarihsel gelişimi gösterilmeye çalışılmıştır. 3B nokta bulutlarından derin öğrenme ile semantik segmentasyon amacıyla derin öğrenme modellerini eğitmek için kullanılabilecek nokta bulutu veri setlerinden örnekler sunulmuş, nokta bulutları ile ilgili makine öğrenmesi çalışmalarının performanslarını ölçmekte kullanılabilecek metrikler tanımlanmıştır. Veri ve değerlendirme ile ilgili bu temel konuların ardından literatürde 3B nokta bulutlarının semantik segmentasyonu için geliştirilmiş ve son teknoloji standartlarını belirleyen derin öğrenme modelleri veriyi ele alış biçimleri bağlamında sınıflandırılarak incelenmiştir.

3.1 Nokta Bulutlarından Derin Öğrenme Çalışmalarında Karşılaşılan Zorluklar

3B nokta bulutları, görüntü ve metin gibi veri türlerine nazaran derin öğrenme bağlamında daha yeni değerlendirilmeye başlamış bir veri yapısıdır. Nokta bulutları, 3B geometrik bilginin ifade edilmesinde kullanılabilecek en sade veri türlerinden biridir. Fakat derin öğrenme çözümlerinin bu veri türüne uyarlanması basit bir görev değildir.

Nokta bulutlarının derin öğrenme yöntemi ile işlenmesi ve analiz edilmesine yönelik çalışmalar, özellikle bilgisayarlı görü alanında yaşanan büyük gelişmelerden alınan motivasyonla başlamıştır. Bu çalışmalarda sınıflandırma, nesne yakalama ve semantik segmentasyon görevlerine yönelik derin evrimsel sinir ağlarının kullanılmasına dayanan çözümler geliştirilmiştir. Derin öğrenme literatüründe 3B nesne sınıflandırma (object recognition); nesne veya nesne parçalarını ifade eden nokta bulutlarının bütün olarak sınıflar ile ilişkilendirilmesini ifade eder. 3B nesne tespit yakalama/konumlama (object detection); nokta bulutunda bulunan ilgi sınıflara ait objelerin veri üzerinde sınırları ile belirlenmesi görevidir. 3B semantik segmentasyon ise verinin nokta bazında ilgili olduğu sınıflar ile ilişkilendirilmesi işlemidir.

Derin öğrenme çözümlerinin, yukarıda bahsedilen görevlerde etkin ve verimli biçimde kullanılabilmesi için nokta bulutlarının hem üretim yöntem ve koşullarından kaynaklı hem de veri türü olarak doğal yapılarının etkisiyle ortaya çıkan zorlukların aşılması gerekmektedir (Guo ve diğ, 2021).

3.1.1 Üretim koşullarından doğan zorluklar

Nokta bulutu verisinin üretim yöntem ve koşullarından kaynaklı verinin içeriği ile ilgili zorluklar çoğunlukla geometrik, spektral ve tematik özelliklerine ilişkindir (Guo ve diğ, 2021). Nokta bulutunun bilgi içeriği üretildiği kaynağa göre değişkenlik gösterdiğinden bu bölümde geometrik/uzamsal özellikler üzerine eğilinmiştir.

Değişken nokta yoğunluğu, nokta bulutu çözünürlüğü üretildiği andaki ölçme geometrisine bağlı olarak veri kümesi boyunca değişkenlik gösterir. Bunun yanı sıra farklı amaçlar için üretilmiş nokta bulutları örnekleme gereksinimlerine göre farklı mekânsal çözünürlük karakterlerine sahip olurlar. Bu durum belirli platformlardan elde edilmiş veriler ile eğitilen ağların farklı koşulları taşıyan veri üstünde kullanılabilirliğini etkilemektedir.

Gürültü, tüm algılayıcılar standart ve kalitelerine oranla değişen gürültüye sahiptir. Verideki gürültü ölçülen nesnelere ait özelliklerin olduklarından farklı biçimde gözükmesine sebep olabilir. Bu durum eğitim esnasında ağın oluşturduğu temsillerin performansına ve genelleştirme başarısına etki edebilir.

Eksiklik, algılama geometrisine ve ölçülmekte olan çevrenin fiziki koşullarına bağlı olarak ortaya çıkan örtüşmeler ile yüzeylere dair yansıma özelliklerine bağlı olarak farklı büyüklüklerde veri kayıpları meydana gelebilir. Bu durum özellikle otonom araçlar tarafından dinamik çevrelerde gerçek zamanlı olarak üretilen HD-map gibi verilerde ciddi biçimde ortaya çıkmaktadır. Bu nedenle geliştirilen modellerin ilgi objelerdeki kısmi eksikliklere karşı güçlü olması gerekmektedir.

Sınıf bulanıklığı, farklı sınıflara dahil olup, geometrik, spektral ya da diğer mevcut özniteliklere göre benzerlik gösteren objeler modeller açısından zorlayıcı örnekler olmaktadır.



Şekil 3.1 : 3B nokta bulutlarında veri üretim koşullarından kaynaklı noksanlıklar.

Şekil 3.1’de solda algılama geometrisine bağlı olarak örtülmelerin etkisi gösterilmiştir. Ortada ve sağda algılama geometrisine bağlı olarak veri yoğunluğunun eksenler boyunca değişkenlik gösterebileceği ifade edilmiştir.

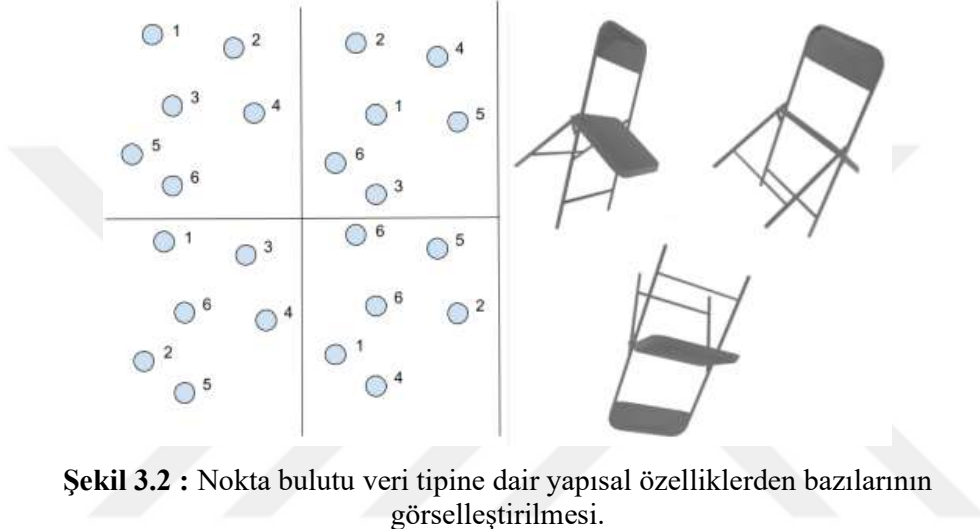
3.1.2 Veri tipinin karakteristik özelliklerinden doğan zorluklar

Geliştirilen derin öğrenme modelleri nokta bulutlarının üretim koşullarından doğan zayıflıklarının yanı sıra veri tipinin yapısal özelliklerini göz önüne almalıdır.

Permütasyon invaryansı, nokta bulutları temel olarak ilgi objeleri düzensiz biçimde örnekleyen 3B noktalardan oluşan kümelerdir. Bahsedilen kümeleri oluşturan noktalar ham hallerinde belirli bir uzamsal ölçüte göre sıralı biçimde tutulmazlar. Bu sebeple nokta bulutlarından derin öğrenme amaçları için geliştirilen modeller, Şekil 3.2’de görülebileceği gibi ‘N’ sayıda noktadan oluşan bir girdinin ‘N!’ farklı permütasyonu

için de aynı sonucu üretebilmeli veya nokta bulutunu oluşturan noktaları veri işleme amaçları bakımından anlamlı bir sıralamaya tabi tutmaya yarayacak çözümler içermelidir.

Rijit dönüşümlere karşı duyarsızlık. İlgi objeler nokta bulutlarında dönüklük, öteleme ya da ölçekleme gibi rastlantısal rijit dönüşümler altında bulunabilirler. Derin öğrenme modelleri bu etkilere karşı duyarsız olmalıdırlar.



Şekil 3.2 : Nokta bulutu veri tipine dair yapısal özelliklerden bazılarının görselleştirilmesi.

Solda nokta bulutlarının, elemanları rastlantısal biçimde sıralanmış kümeler olmalarından dolayı indislerinin mekansal bir bağlamı olmadığı ifade edilmiştir. Sağda aynı objenin 3B uzayda farklı dönüşümler altında bulunabilirliği görselleştirilmiştir.

Verimli çalışılabilirlik, nokta bulutu veri setleri milyonlarca noktadan oluşabilmektedir. Bu ölçekte bir verinin eğitim, test ve çıkarım süreçlerinde kullanılabilmesi için iyi bir depolama stratejisine ihtiyaç duyulur. Klasik konvolüsyonel mimari hesaplama açısından büyük oranda optimize edilmiştir. 3B nokta bulutları için doğruluk ve işlem yükü açısından verimli modellerin geliştirilmesi özellikle gerçek zamanlı çözüm hedeflenen görevler için kritiktir (Li ve diğ, 2021).

3.2 Nokta Bulutlarından Derin Öğrenme Sonuçlarını Değerlendirmede

Kullanılan Ölçütler

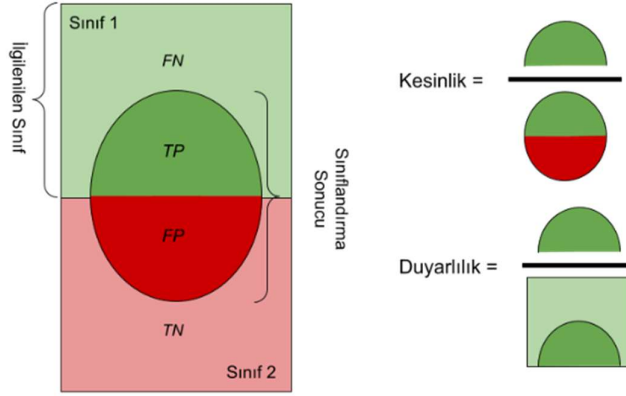
Makine öğrenmesi çalışmalarında eğitilen modellerin başarısını ifade etmek ve tartışmak amacıyla çeşitli istatistiksel ifadelerden faydalanılmaktadır. Bu bölümde özellikle semantik segmentasyon görevleri için geliştirilen modelleri kıyaslamada yaygınlıkla kullanılan değerlendirme ölçütlerinden bahsedilmiştir.

Genel doğruluk, (overall accuracy, OA) tüm makine öğrenmesi çalışmalarında kullanılabilecek en temel değerlendirme ölçütü olarak görülmektedir. Herhangi bir girdi veri setinde yürütülecek bir test için OA değeri doğru sonuçlandırılan test sonuçlarının tüm test sonuçlarına oranı şeklinde hesaplanabilir. Sınıf bazında ortalama doğruluk (mean class accuracy, mAcc) bir diğer temel değerlendirme ölçütüdür. Yürütülen bir test için mAcc değeri, girdi veri setinde sınıf bazında hesaplanan doğruluk oranlarının sınıf sayısına göre ortalaması ile bulunur. Modellerin farklı sınıftan test örneklerine karşı başarısı farklı farklı olabildiğinden bütünsel bir değerlendirme için sıklıkla tercih edilen bir metriktir (Guo ve diğ, 2021).

Makine öğrenmesi çalışmalarında model performansını daha detaylı değerlendirmeye imkan tanıyan metriklere örnek olarak kesinlik ve duyarlılık verilebilir. Bu kavramlar genel olarak bir sınıflandırma testi için gerçekleşmesi olası senaryolar ile Şekil 3.3'teki gibi görsel olarak özetlenebilir. Kesinlik; bir sınıfa atanan örnekler arasında doğru olanların sayısının, bu sınıfa atanan tüm örnekler oranı olarak ifade edilebilir. Duyarlılık ise belirli bir sınıfın örnek uzayda fark edilebilirliğinin bir ölçütü olarak; doğru sınıflandırılan örnek sayısının bu sınıfa ait tüm örnekler oranı şeklinde hesaplanabilir (Ghorbanzadeh ve diğ, 2019).

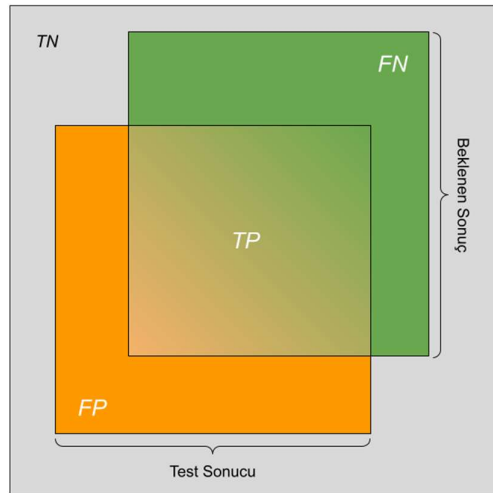
Model başarısını raporlama amacı ile kesinlik ve duyarlılığı bir arada değerlendirme imkanı sunmak için geliştirilmiş F1-skoru popüler bir metrik olarak kullanılmıştır. Kesinlik ve duyarlılık değerleri oranlar ile ifade edildiklerinden F1-skoru kesinlik ve duyarlılık değerlerinin, oranların ortalamasını elde etmek için kullanılan (örneğin ortalama hız), harmonik ortalaması şeklinde hesaplanmaktadır. Hesaplama yöntemi olarak harmonik ortalama fonksiyonu kullanması neticesinde kesinlik ya da duyarlılık değerlerinden birinin diğerine göre çok düşük olması durumunda F1-skorunun düşük olana yakın biçimde ortaya çıkacağı görülmüştür. F1-skoru kesinlik ve duyarlılık metriklerinin birbirlerine göre dengeli olup olmadığını da gösterebilen bir metrik olarak ön plana çıkmıştır.

Segmentasyon ve obje yakalama görevleri için geliştirilen modeller tarafından üretilen sonuçların kalitesi büyük oranda girdi örneklerin geometrik şekillerini yansıtmaya becerisine bağlıdır. Modellerin başarısını ürettikleri sonuçların hedef nesne geometrilerine uyum oranı bakımından değerlendirmeye ve kıyaslamaya imkan tanıyan kesişim-birleşim oranı (intersection over union, IoU) ya da diğer ismi ile



Şekil 3.3 : Kesinlik ve duyarlılık kavramlarını ifade etmek için hazırlanmış görsel. TP: doğru pozitif sınıflandırma durumu, FP: yanlış sınıflandırma durumu, FN: yanlışlıkla sınıf dışında bırakma durumu, TN: doğru biçimde sınıf dışında bırakma durumunu ifade etmektedir.

Jaccard Skoru yaygın biçimde kullanılan bir metrik olarak ön plana çıkmaktadır (Guo, ve diğerleri, 2021). IoU değerinin hesaplanması, test sonuçları ve beklenen değerler tarafından oluşturulan kümeler üzerinde tanımlı bir küme problemi olarak ifade edilebilir. Şekil 3.4'te de görülebileceği gibi test sonucu ve gerçek değer kümelerinin oluşturduğu kesişim ve birleşimler ilgili örnek sayıları ile oranlanarak, sonuçların gerçek ile örtüşmesini ifade eden bir değer elde edilir. IoU değerinin derin öğrenme çalışmalarında sınıf bazında ortalama, ve tüm sınıflar ortalaması şeklinde hesaplanabilen ortalama kesişim-birleşim oranı anlamına gelen (mean IoU) mIoU kısaltması ile raporlandığı görülmüştür. IoU hesaplama yöntemi Denklem 3.1'de sunulmuştur.



Şekil 3.4 : Kesişim-Birleşim oranı (IoU) değerini bir küme problemi olarak ifade etmek için hazırlanmış görsel.

Şekil 3.4’te, TP: doğru pozitif sınıflandırma durumu, FP: yanlış sınıflandırma durumu, FN: yanlışlıkla sınıf dışında bırakma durumu, TN: doğru biçimde sınıf dışında bırakma durumunu ifade etmektedir.

$$IoU = TP / (FP + TP + FN) \quad (3.1)$$

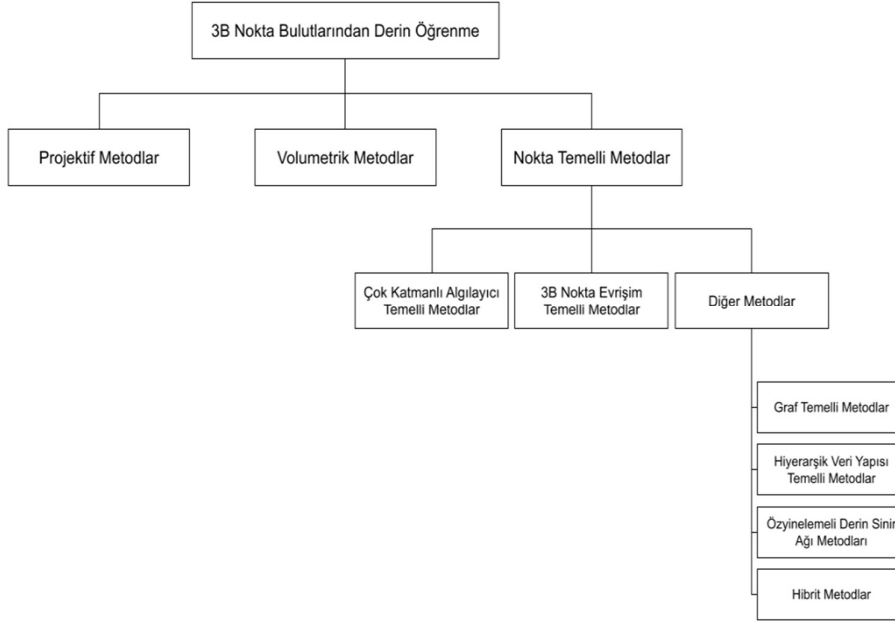
Eğitim için kullanılan veri setlerinde tanımlanan sınıflar birbirleri ile orantısız biçimde temsil edilebilmektedirler. Bu sebeple; veriden öğrenme yaklaşımı üzerine kurgulanmış olan makine öğrenmesi çalışmalarında az sayıda temsil edilen sınıflar aleyhine bir durum ortaya çıkabilmektedir. Veri setinin büyük çoğunluğunun orantısız biçimde az sayıda sınıfa ait örnekten oluştuğu durumlarda OA değeri model performansını ifade etmekte yetersiz kalabilmektedir. Bu nedenle sınıf bazında değerlendirme imkanı sunan ölçütlerin model performansını analiz etmede ve kıyaslamada daha kullanışlı olduğu görülmüştür. Ayrıca segmentasyon ve obje yakalama görevlerinde kullanılan modeller tarafında üretilen sonuçların hedef nesne geometrilerinin aslına uygunluğunu ölçmede IoU faydalı bir metrik olarak tespit edilmiştir.

3.3 3B Nokta Bulutlarından Derin Öğrenme İçin Geliştirilen Çözümler

3B nokta bulutlarından derin öğrenme alanındaki çalışmalar yakın geçmişte 2B görüntü verisi üzerine geliştirilen çözüm ve yaklaşımların birer uzantısı olarak ortaya çıkmış fakat birkaç yıllık süre içerisinde özelleşmiş çözümlere doğru olgunlaşmıştır (Ahmed ve diğ, 2018). Alanın gelişimi boyunca nokta bulutları tarafından sunulan detaylı 3B bilgiden faydalanabilmek için çeşitli yaklaşımlar uygulanmıştır. Genel olarak bakıldığında çalışmaların veriyi ele alış biçimi temelinde ayrılarak evrildiği ve geliştiği (Şekil 3.5) görülmüştür.

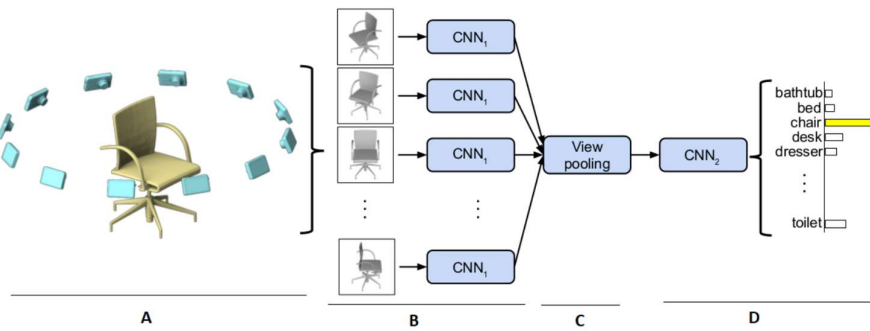
3.3.1 Projektif yöntemler

Projektif yöntemler modern evrişimsel sinir ağı mimarisinin görüntü verisindeki başarılı uygulamalarından hareketle geliştirilmişlerdir. Bu çalışmalarda nokta bulutlarından elde edilen çoklu bakış görüntüleri kullanılmıştır. Görüntülerin derin öğrenme ağı ile işlenmesi sonucunda elde edilen temsillerin bir araya getirilerek global özellik vektörünün oluşturulmasında izlenecek yol ve yöntemler çalışmalarda ön plana çıkmıştır (Ahmed ve diğ, 2018).



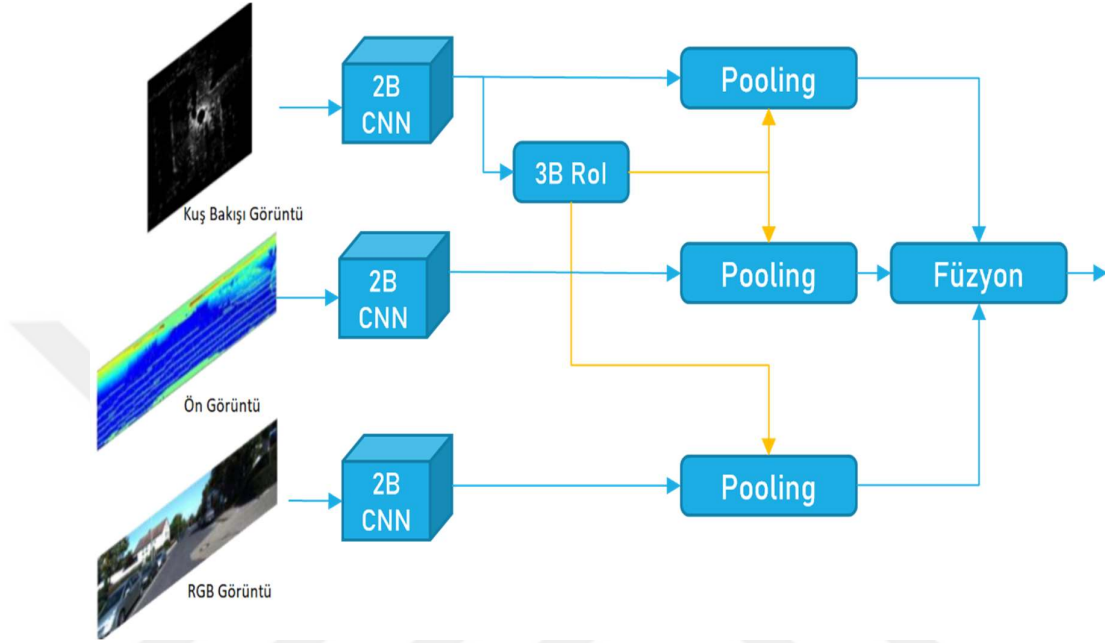
Şekil 3.5 : 3B nokta bulutlarından derin öğrenme için geliştirilen yöntemlerin veriyi ele alış biçimlerine göre hazırlanmış bir taksonomisi.

Projektif metodların kullandığı bir çalışma olan MVCNN’de (Su ve diğ, 2015) Şekil 3.6’da da görülebileceği gibi, bir nesneye ait farklı perspektif noktalardan elde edilen görüntüler kullanılmıştır. Uygulamada 3B nesne modellerine dair farklı perspektiflerden elde edilmiş sanal görüntülerden 2B evrişimsel sinir ağları kullanılarak temsiller elde edilmiştir. Bu temsilleri optimum biçimde kombine edecek tanımlayıcıların elde edilmesinde maksimum pooling katmanı kullanılmıştır. İlerleyen çalışmalarda çoklu görüntüler arasındaki ilişkileri göz önüne alan temsillerin elde edilmesi imkanları ve görüntülerden elde edilen temsillerin bir araya getirilmesinde veri temelli ve esnek yöntemlerin kullanılması imkanları araştırılmıştır (Yang ve Wang, 2019).



Şekil 3.6 : MVCNN mimarisi, Su ve diğ. (2015)’den uyarlanmıştır.

Şekil 3.6’da MVCNN mimarisi sunulmuştur. A, görüntülerin elde edilmesi. B, görüntülerden 2B evrişimsel sinir ağı modelleri ile özellik vektörlerinin oluşturulması. C, global özellik vektörünün elde edilmesi. D, sınıflandırıcı uçta işlem.



Şekil 3.7 : MV3D mimarisi, Guo ve diğ. (2021)’den uyarlanmıştır.

MV3D; şekil 3.7, lidar ve RGB kamera algılayıcıları tarafından elde edilen verilerin füzyonu için de örnek teşkil eden bir mimari olmuştur. Lidar verisinin ‘kuş bakışı görüntüsü’ ve ‘ön görüntüsü’nden; ayrıca RGB kamera tarafından elde edilen görüntülerden 2B DCNN’ler ile özellik çıkarımı yapılmıştır. ‘Kuş bakışı görüntü’den elde edilen ilgi alanları (RoI)’ ayrıca bu özellik vektörleri üzerine projekte edilmiştir. Son olarak tüm görüntülerden elde edilen özellik vektörleri dikkat mekanizmaları ve pooling operasyonları ile bir araya getirilmiştir (Chen ve diğ, 2017).

Projektif metodlar sınıflandırma, obje yakalama ve semantik segmentasyon görevleri için geliştirilen modellerde kullanılmıştır. Özellikle otonom sürüş uygulamalarını konu alan çalışmalarda sayısal kamera ile desteklenmiş lazer tarama verilerini içeren veriler üzerinde başarılı sonuçlar elde edilmiştir. Projektif metodlar perspektif noktalarını belirleme sürecinin uçtan uca eğitilebilir bir mimari kurmayı zorlaştırması, en uygun perspektif noktalarını belirlemenin zor olması, 3B verinin içerdiği şekilsel ve mekânsal ilişkilerden faydalanma imkanını zayıflatmaları öne sürülerek eleştirilmiştir (Guo ve diğ, 2021).

3.3.2 Volumetrik yöntemler

Bu yöntemler modern CNN mimarisi çözümlerini 3B veriye uyarlamak amacıyla geliştirilmiştir. Uygulamada nokta bulutları vokselizasyon işlemi gibi sürekli verinin kesikli bir temsili elde etmeye yarayan algoritmalar ile düzenli volumetrik bir temsile dönüştürülüp 3B evrişimsel sinir ağları ile işlenir. Evrişimsel sinir ağı mimarisi girdi verideki yerel komşuluk alanlarının ilişkisel özelliklerini değerlendirme konusunda başarılıdır. Bu beceri 3B modellerin sunduğu bilgi içeriği açısından değerli olmaktadır. Fakat nokta bulutları veri tipi olarak klasik evrişim işlemleri ile işlenmeye uygun olmadığından düzenli bir temsil elde etmek için Şekil 3.8’de gösterildiği gibi işlemler gerçekleştirilir (Li ve diğ., 2021).

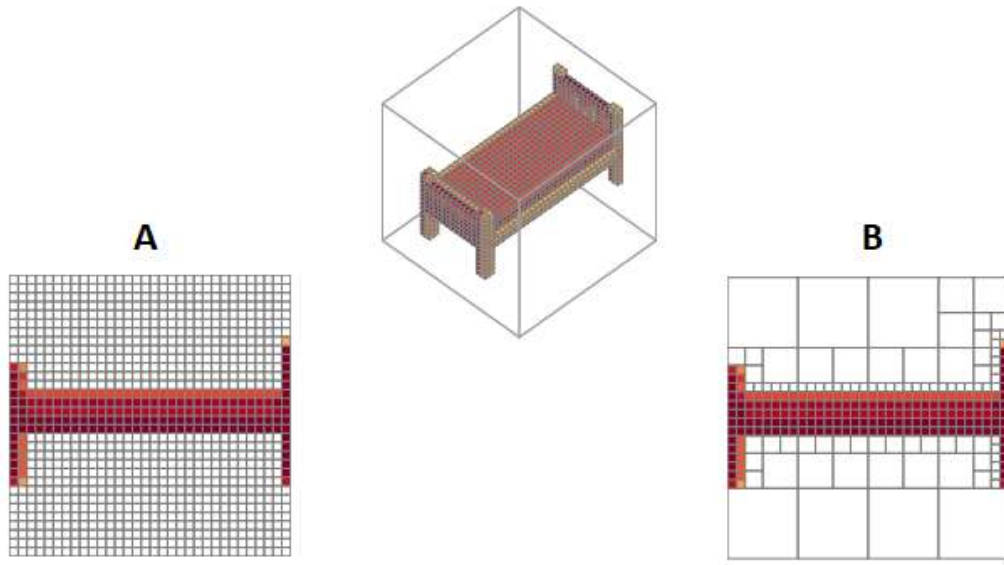
VoxNet, nokta bulutlarını vokselizasyon işlemi ile 3B grid temsillere çevirmiştir. Daha sonra bu temsiller, 3B evrişim elemanları ile filtrelenerek işlenmiştir. Bilgisayarlı görü görevlerinde kullanılan 2B CNN mimarisinin tekdüze biçimde 3B veriye uyarlanmasına imkan tanımıştır. Çeşitli nesnelere ait 3B CAD modellerini içeren bir veri seti olan ModelNet40 veri setinden yeniden örneklenmiş nokta bulutları için, VoxNet %83 ortalama sınıflandırma doğruluğuna ulaşmıştır (Maturana ve Scherer, 2015).

İlgi nesneyi ifade eden 3B gridin çözünürlüğü arttıkça, ince detaylar daha çok ortaya çıkar fakat buna karşılık veriyi saklamak için gerekli hafıza gereksinimi kübik biçimde artar. Aynı zamanda girdi birimi olarak kullanılan volumetrik temsilin doluluk oranı azalır. Sonuç olarak evrişim işlemi, git gide daha seyrek bir veri üstüne verimsiz biçimde uygulanmak zorunda kalır. Bu olumsuz duruma karşılık, devam eden çalışmalarda çeşitli çözümler geliştirilmiştir.

Gerçek dünya koşullarında üretilmiş yoğun nokta bulutlarının volumetrik yöntemler ile verimli biçimde işlenebilmesi için doğrudan vokselizasyon yerine, depolama ve hesaplama açısından daha verimli dönüşümler ve veri yapıları araştırılmıştır.

OctNet çalışmasında octree adında hiyerarşik bir ağaç yapısından faydalanılmıştır. Uygulamada volumetrik temsil, grid çözünürlüğü alışıldık üzere veri bütününde uniform kabul edilmemiş, veri sıkıştırma ve mekansal indisleme algoritmalarına benzer şekilde, veriyi oluşturan noktaların mekânsal dağılımına göre dinamik biçimde belirlenmiştir. Gridler, ayrıklaştırma adımıyla veri yeniden örneklenirken iteratif biçimde bir veriye rastlayıp rastlamamalarına göre alt gridlere ayrılmıştır. Sonuçta

dolu voksellerin daha yüksek çözünürlüklü biçimde temsil edildiği hiyerarşik bir ağaç elde edilmiştir. Bu ağaç veri yapısı ile nokta bulutu, yaprak ve düğümlerini, boş ve dolu olma durumuna göre 0 ve 1 değerleri ile ifade eden bit-dizeleri ile verimli biçimde temsil edilebilmiştir. Bahsedilen bit-dize temsiller, grafik işlem birimlerinde yürütülecek matematiksel işlemlerin verimli biçimde gerçekleştirilmesi için özel olarak dizayn edilmiş bir evrişim algoritması ile ağa beslenmiştir. Bu sayede hem hesaplama hem depolama yönünden kazanımlar elde edilmiştir. Ek olarak geliştirilen evrişim işleminin hücre boyutlarına göre adaptif biçimde uyarlanması sağlanarak hesaplama yönünden verim elde edilmiştir (Riegler ve diğ., 2017).



Şekil 3.8 : Örnek bir obje için (A) basit uniform vokselizasyon yapısı ve (B) octree tabanlı volumetrik temsile ait kesitler, Riegler ve diğ. (2017)’den uyarlanmıştır.

Volumetrik yöntemler doğrudan 3B veri üstünde çalışarak cesaret verici sonuçlar elde etmiş olsalar da hesaplama ve depolama gereksinimleri yönünden ham nokta bulutlarına göre kübik bir artışa sebep oldukları için eleştirilmişlerdir.

3.3.3 Nokta bulutu temelli yöntemler

Bu yöntemler, projektif ve volumetrik yöntemlerin aksine şekil bilgisi kaybına sebep olmamaları, verinin farklı ön dönüşümlerini zorunlu kılmamaları, uçtan uca eğitilebilir mimarilerin kurgulanmasını kolaylaştırmaları ile git gide popülerleşmektedir (Guo ve diğ., 2021). Doğrudan nokta bulutları üzerinde çalışmak amacıyla araştırmacılar; permütasyon invaryansı ve 3B pozdan bağımsızlık şartlarına uyum sağlayabilecek farklı derin öğrenme modelleri, katmanlar ve veri işlem adımları geliştirmiştir.

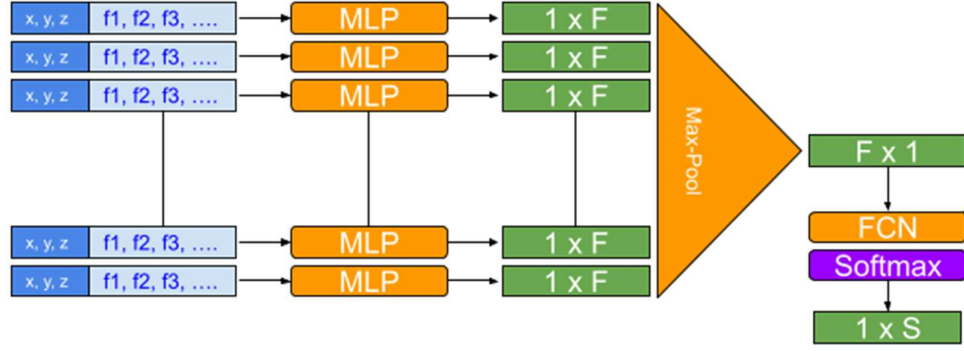
PointNet, Qi ve diğ. (2017), doğrudan nokta bulutu temelli metodlar açısından öncül olmuştur ve çoğu çalışma için temel kabul edilmiştir (Xie ve diğ, 2020). Bu bölümde geliştirilen yaklaşımlar mimari yönünden incelenmiş ve çok katmanlı algılayıcı, 3B nokta konvolüsyon, diğer metodlar olarak gruplandırılmıştır.

3.3.3.1 Çok katmanlı algılayıcı temelli yöntemler

Nokta bulutlarının otomatik olarak işlenmesi amacıyla özellik çıkarımı için derin çok katmanlı algılayıcılar kullanılması hedefi açısından PointNet (Şekil 3.9) çalışması önemli bir kilometre taşı olmuştur.

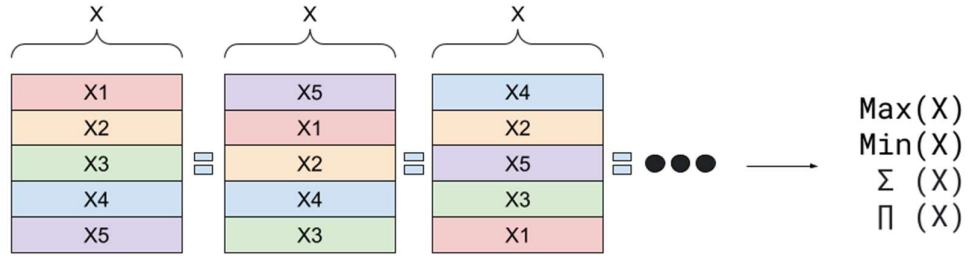
Bu çalışmada, girdi olarak kullanılan nokta bulutları, elemanları rastlantısal olarak sıralanmış kümeler olarak ele alınarak bir mimari tasarlanmıştır. Girdi olarak kullanılan tüm noktalar, aynı çok katmanlı algılayıcıya beslenerek her biri için özellik vektörleri çıkartılmış. bu özellik vektörleri tüm girdiyi temsil etmesi için global bir özellik vektörüne dönüştürülürken, permütasyon invaryansı sağlamak için simetrik (Şekil 3.10), yani elemanlarına göre değişme özelliğine sahip, bir fonksiyon olan maksimum pooling katmanları kullanılmıştır. Elde edilen global özellik vektörü segmentasyon ya da sınıflandırmada kullanılmak amacıyla uygun mimaride değerlendirilmiştir. Ek olarak 3B uzaydaki pozdan bağımsız biçimde değerlendirme yapabilmek için girdi nokta bulutları eğitilebilir adaptif bir ortogonal dönüşüme tabi tutulmuştur. Koordinat dönüşümü katmanının doğruluk açısından pozitif etki yarattığı görüldüğünden benzer bir strateji yapay sinir ağı katmanları tarafından üretilen temsiller için de kullanılmak üzere eğitilebilir adaptif özellik dönüşümü katmanları eklenmiştir. PointNet mimarisi ile ModelNet40 veri setinde %89 ortalama doğruluk elde edilmiştir. Doğrudan nokta bulutlarını girdi kabul ederek uçtan uca eğitilebilir bir mimarinin dizayn edilebileceğini gösterdiğinden PointNet ardından gelen çalışmalara ilham olmuştur (Joseph-Rivlin ve diğ, 2019).

Bir diğer örnek olan DeepSets, Zaheer ve diğ. (2017) çalışmasında ise permütasyon invaryansı, maksimum pooling gibi aynı şekilde simetrik bir fonksiyon olan toplam fonksiyonu kullanılarak elde edilmiştir. Bu çalışmada global özellik vektörü, noktalar için üretilmiş tekil özellik vektörlerinin toplamının adaptif lineer bir dönüşümü kullanılarak elde edilmiştir.



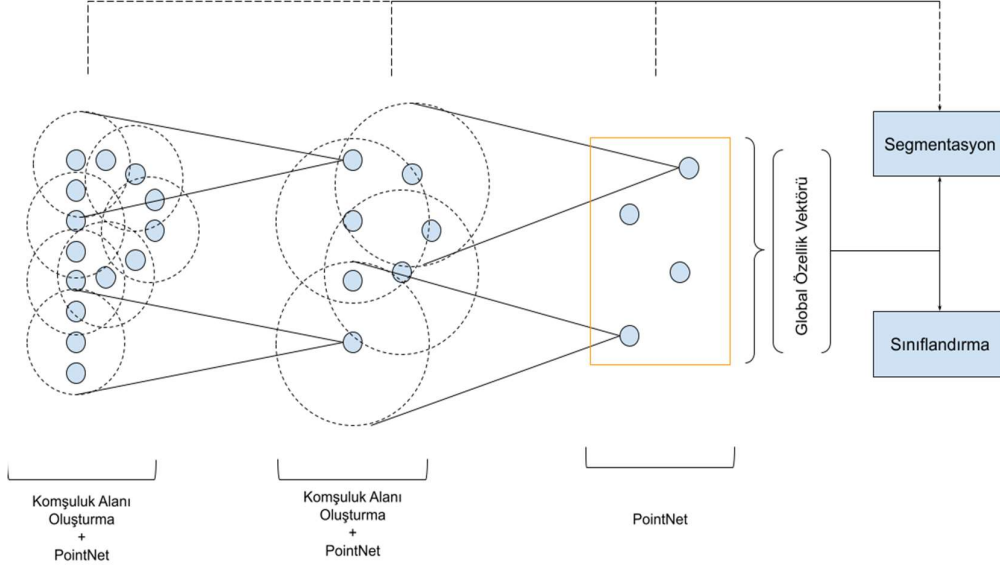
Şekil 3.9 : PointNet Mimarisi.

PointNet ve benzeri mimariler, nokta bulutlarının otomatik olarak işlenmesi ile ilgili çok temel problemlere basit ve adaptif çözümler geliştirmiş olsalar da girdi noktaları tekil biçimde işleme almaları nedeniyle yerel komşuluk özelliklerini dolayısı ile de şekil bilgisini göz ardı ettikleri gerekçeleriyle eleştirilmişlerdir (Li ve diğ, 2021).



Şekil 3.10 : Simetrik fonksiyonlara örnekler.

PointNet++, Qi ve diğ. (2017) çalışmasında girdi nokta bulutu, merkezleri nokta örnekleme algoritmaları ile tespit edilen komşuluk alanlarına bölünmüştür. Sonrasında komşuluk alanı kümelerine dahil edilecek noktalar yakınlık ölçütlerine göre belirlenmiştir. Elde edilen komşuluk alanları PointNet ağı ile işlenerek her komşuluk alanı için özellik vektörleri elde edilmiş ve sonraki adımda komşuluk alanını temsil etmek üzere seçilen noktanın özellik vektörüne eklenmiştir. Şekil 3.11’de PointNet++ mimarisi özetlenmiştir. Elde edilen yeni noktalar da kendi aralarında komşuluk alanlarına ayrılmış ve bir önceki adımdaki işlemlere tabi tutulmuştur. Bu döngü belirli sayıda temsili noktaya ulaşıncaya kadar devam ettirilerek nokta bulutundaki şekil bilgisini hiyerarşik biçimde yakalamayı hedefleyen bir mimari geliştirilmiştir. Elde edilen özellik vektörleri sınıflandırmada ya da üst örnekleme adımları sonrasında segmentasyonda kullanılmıştır.



Şekil 3.11 : PoinNet++ mimarisi, yerel özellik vektörlerinin hiyerarşik biçimde elde edilmesi.

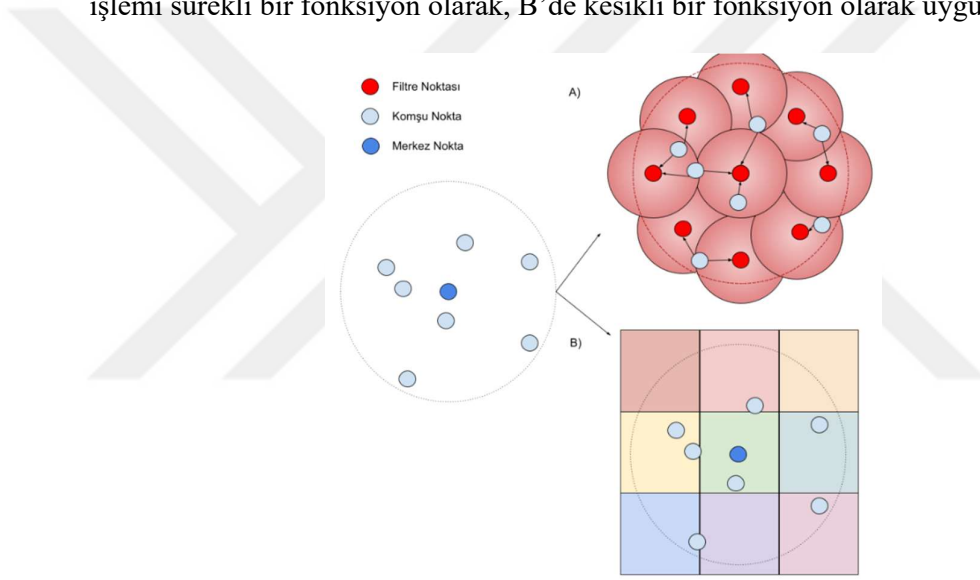
Doğrudan nokta bulutları üzerinden derin çok katmanlı algılayıcı mimarisi ile öğrenme çalışmaları, nokta bulutları tarafından sunulan detaylı 3B bilgiden faydalanma imkanlarına kapı aralamıştır. Bu yaklaşımla yürütülen çalışmalarda farklı komşuluk alanı oluşturma yöntemleri (Yang ve diğ, 2019), komşuluk alanlarındaki noktaların birbirlerine göre konumlarını ifade edebilecek özelliklerin tasarlanması (Zhao ve diğ, 2019), farklı ölçeklerde elde edilen komşuluk alanlarına ait özellik vektörlerinin kombine edilmesi (Yan ve diğ, 2020) gibi mekansal bağlamdan faydalanmayı hedefleyen çalışmalar yürütülmüştür.

3.3.3.2 3B nokta evrişim metodları

Nokta bulutları ile derin öğrenme için, yerel nokta komşuluk alanlarının tarif ettiği mekansal ilişkilerden faydalanmaya yönelik çalışmalarda, söz konusu görevde başarısını görüntü verisi üzerinde ispat etmiş evrişimsel mimarinin uygulanabilirliği araştırılmıştır.

PointNet++ gibi derin çok katmanlı algılayıcı çözümleri; özellik çıkartımına yönelik operasyonlarını yerel komşuluk alanları üzerinde yürütmek üzere geliştirilmiş olsalar da işlem birimi olan noktalar, aynı çok katmanlı algılayıcıya tekil biçimde beslendiklerinden bu çalışmalarda noktaların birbirleri ile olan mekansal ilişkilerinden faydalanma becerisi kısıtlı olarak gerçekleşmiştir. Bunun üzerine çözüm olarak 3B nokta evrişim modelleri önerilmiştir.

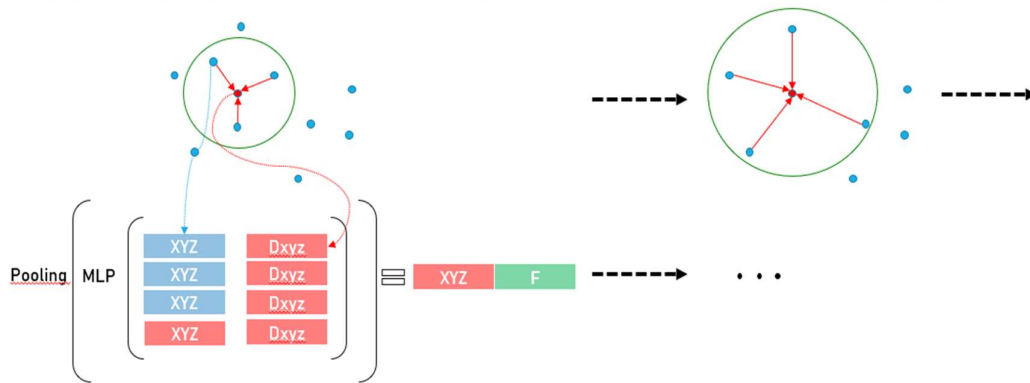
3B evriřim iřlemine volumetrik temsiller yerine dođrudan nokta bulutu zerinde uygulanabilir kılmak iin eřitli mler geliřtirilmiřtir. Farklı rneklemeye ve yakınlık analizi algoritmaları ile meydana getirilen yerel komřuluk alanlarında yer alan noktaların evriřim iřlemine tabi tutulması iin genel olarak ye noktaların ve komřuluk alanı merkezi arasındaki geometrik iliřkiye gre ifade eden zellik vektrlerinden faydalanılmıřtır. Evriřim filtrelerinin komřuluk alanlarının merkezlerine konumlandırıldıđı kabul edilmiřtir. Noktalar ve filtre elemanları arasındaki etkileřim řekil 3.12’de gsterildiđi zere aralarındaki geometrik iliřkiye bađlı biimde farklı yntemlerle kurgulanmıřtır (Guo ve diđ, 2021). A’da evriřim iřlemi srekli bir fonksiyon olarak, B’de kesikli bir fonksiyon olarak uygulanmıřtır.



řekil 3.12 : 3B nokta evriřim iřlemine uygulamak iin genellikle kullanılan yaklařımların grselleřtirmesi.

Evriřim iřlemine srekli bir fonksiyon olarak ele alan alıřmalar arasında RS-CNN, Liu ve diđ. (2019) nc olmuřtur. Bu alıřmada yerel komřuluk alanlarından objeye iliřkin řekil bilgisinin hiyerarřık biimde elde edilebilmesi iin zel olarak tasarlanmıř bir 3B evriřim iřleminden yararlanılmıřtır. Evriřim iřlemi adaptif ve eđitilebilir bir biimde ok katmanlı algılayıcı (Multi Layer Perceptron, MLP) kullanılarak kurgulanmıřtır. İlk olarak yerel komřuluk alanları nokta rneklemeye ve yakınlık analizi algoritmaları ile oluřturulmuřtur. Komřuluk alanında bulunan her bir nokta iin, merkezde bulunan rneklemeye noktasıyla olan geometrik iliřkisini temsil etmek amacıyla, iki nokta arasındaki koordinat farkları ve 3B klid mesafesinden oluřan zellik vektr tanımlanmıřtır. Bu zellik vektr her nokta iin aynı ok katmanlı algılayıcı ile iřlenip zellik dnřm sađlandıktan evriřim iřlemi yine bir ok

katmanlı algılayıcı ile gerçekleştirilmiştir. Daha sonra nokta bazında elde edilen özellik vektörleri simetrik bir fonksiyon olan toplama işlemine tabi tutularak komşuluk alanını temsil eden bir özellik vektörü hesaplanmıştır. Bu sayede de permütasyon invaryansı sağlanmış, yerel komşuluk alanındaki geometrik ilişkileri ifade edebilen bir özellik vektörü elde edilmiştir. Bu işlem, örnekleme noktaları bir sonraki adıma elde edilen özellik vektörü ile birlikte taşınarak bir temsiller hiyerarşisi elde edilmiştir. Elde edilen temsiller sınıflandırma ya da segmentasyon işleminde kullanılmıştır. RS-CNN modeli, ModelNet40 veri setinde sınıflandırma görevinde %93.6 gibi önemli bir genel doğruluk performansına ulaşmıştır.



Şekil 3.13 : Sürekli fonksiyon olarak tanımlanmış 3B nokta evrişim işlemine bir örnek.

Şekil 3.13’de evrişim operasyonunun sürekli bir fonksiyon olarak tanımlanmasına örnek bir mimaride işlem akışı özetlenmiştir. Örnekleme yöntemleri ile komşuluk alanı merkezleri belirlenir. Komşuluk alanları en yakın komşuluk algoritmaları ile tespit edilir. Komşuluk alanındaki noktalar ve merkez noktalar arasındaki geometrik ilişkiyi ifade eden özellik vektörleri oluşturulur. Bu özellik vektörleri, noktalar ve evrişim filtresi elemanları arasındaki mekansal ilişkiye göre ağırlıklandırılarak çok katmanlı algılayıcılar ile işleme tabi tutulur ve komşuluk alanı temsilleri simetrik fonksiyonlardan faydalanılarak nokta özellik vektörlerinin bir araya getirilmesi ile elde edilir (Guo ve diğ, 2021).

3B evrişimsel mimari yaklaşımını temel alan diğer çalışmalarda, işleme tabi tutulacak komşuluk alanını temsil edebilecek uygun girdi özellik vektörlerinin tasarlanması (Boulch, 2019), evrişim işleminin noktalara ait özellik vektörlerinin mekansal bileşenleri ve diğer bileşenleri için ayrı ayrı yürütülmesi (Boulch, 2020), farklı komşuluk alanı boyutlarından elde edilen özellik vektörlerinin uygun biçimde

birleştirilmesi (Liu ve diğ, 2019), evrişim işleminin nokta bulutu kartezyen koordinatları küresel bir yüzeye projekte edildikten sonra harmonik fonksiyonlardan yararlanılarak uygulanması (Esteves ve diğ, 2018) gibi yöntemler geliştirilmiştir.

Evrişimsel yöntemin kullanıldığı çalışmaların diğer ana grubunu; evrişim filtresi, yerel komşuluk alanı üzerinde kesikli biçimde yeniden örneklenerek tasarlanan çalışmalar oluşturmıştır. Bu çalışmalarda evrişim filtreleri komşuluk alanlarının merkezlerinde 3B grid ya da küre gibi düzenli veri yapıları biçiminde tanımlanmış ve noktalar rastladıkları hücreye ilişkin filtre parametresi ile işleme alınmıştır (Guo ve diğ, 2021).

Hua ve diğ. (2018)'de, nokta bulutunun her bir noktasında en yakın komşular elde edilip örnekleme noktasını merkez alan uniform 3x3 boyutunda 3B gridler ile evrişim filtreleri oluşturulmuştur. Hücrelere rastlayan noktalara ait özellik vektörlerinin ağırlıklı ortalamaları alınmış ve hücre ile ilişkili ağırlık ile işleme sokularak yerel özellik vektörleri elde edilmiştir. Bu işlem birimleri arka arkaya tekrar edilerek temsiller hiyerarşisi oluşturulmuş, sınıflandırma ve semantik segmentasyon görevlerinde PointNet'e yakın performans elde edilebilmiştir. Lei ve diğ. (2019)'da evrişim filtreleri, merkezdeki nokta etrafında belirli bir yarıçap ile küresel biçimde, küresel koordinatlarına göre hacimsel birimlere ayrılmış halde tanımlanmıştır. Araştırmacılar, bu küresel filtrelerin 3B nokta bulutları için doğal evrişim işlem birimi olması gerektiğini savunmuşlardır. Katmanlar arası örtüşmeyi ve hesaplama verimliliğini artırmak amacıyla girdi nokta bulutlarını octree yapıları ile indisleyip işlem yönünü yaprakları teşkil eden noktalardan düğümler ve kök noktaya doğru tanımlamışlardır bu sayede temsiller hiyerarşisi elde edilmiştir. Oluşturulan mimari, Ψ -CNN (Psi-CNN) sınıflandırma ve segmentasyon görevlerinde başarılı sonuçlar elde etmiştir. GeoCNN, Lan ve diğ, (2019) çalışmasında, küresel komşuluk alanlarındaki noktaların, merkez noktaya aralarındaki öklid mesafeleri 3B bileşenlerine, ana eksenler ile yaptıkları açılar kullanılarak projekte edilmiştir. Elde edilen bu özellik vektörü çok katmanlı algılayıcı kullanılarak, işlenerek yerel geometrik özellikler ortaya çıkartılmaya çalışılmıştır. PointCNN, Li ve diğ. (2018) çalışmasında girdi nokta bulutu çok katmanlı algılayıcı yardımıyla öğrenilen ve χ -conv dönüşümü adı verilen çok boyutlu özellik çıkarımı ve permutasyon işlemine tabi tutularak adaptif biçimde sıraya sokulmuştur. Öğrenelebilen bir sıralamaya tabi tutulan noktaların meydana getirdiği nokta bulutu zaman serisi verilerine benzer biçimde evrişim işlemine tabi tutulmuştur.

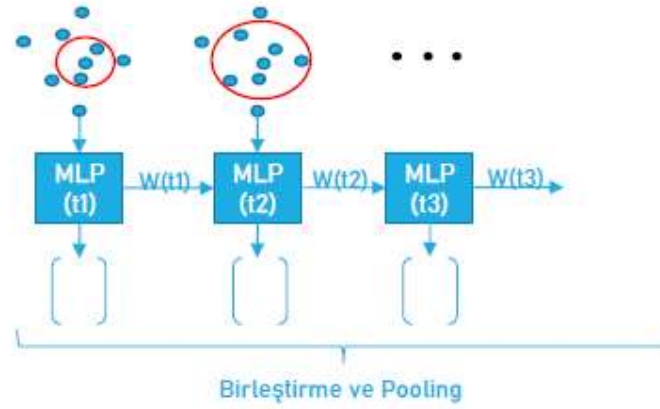
3.3.3.3 Diğer yöntemler

Graf temelli yöntemlerde şekil 3.14; girdi noktalar düğüm olarak kabul edilmiş, komşuluk alanları ile sınırlı ya da nokta bulutunun tamamını içeren graf veri yapıları oluşturulmuştur. Düğüm görevi gören noktalar özellik vektörleri ile, düğümleri birbirine bağlayan kenarlar ise komşu özellik vektörlerinin geometrik ilişkilerine bağımlı fonksiyonlar şeklinde tanımlanmıştır. Graf veri yapısı elde edildikten sonra özelleşmiş graf evrişim işlemi uygulanmış ve şekil bilgisine dair temsiller elde edilmesi hedeflenmiştir (Guo ve diğ, 2021).



Şekil 3.14 : Graf evrişim temelli yöntemlerin örnek işlem akışı şeması.

Hiyerarşik veri yapısı temelli metodlarda, girdi nokta bulutu, octree ya da KD-tree gibi ağaç veri yapılarına dönüştürülmüştür. Elde edilen ağaçların yaprak ve düğümleri komşuluk alanlarını ifade edecek biçimde ayarlanmıştır. Daha sonra yapraklarda konumlandırılmış noktalardan başlayarak düğümler ve kök noktaya doğru önceki kısımlarda anlatılan çok katmanlı algılayıcı ya da 3B nokta evrişim elemanları sayesinde temsiller hiyerarşisi elde edilmiştir (Klokov ve Lempitsky, 2017). Özyinelemeli mimari metodlarından özellikle otonom sürüş senaryoları için obje yakalama ve takip etme görevlerini yerine getirmek amacı ile eğitilen modellerde ya da semantik segmentasyon çalışmalarında faydalanılmıştır. Özyinelemeli yapay sinir ağlarından yararlanılan bir çalışmada girdi nokta bulutları birden çok ölçekte yeniden örneklenilip her ölçek için özellik vektörleri elde edildikten sonra bu özellik vektörleri kombine edilmek amacıyla özyinelemeli sinir ağlarına beslenmiş ve yerelden globale bir temsil hiyerarşisi oluşturulmuştur (Engelmann ve diğ, 2017). Şekil 3.15’de örnek bir işlem akışı şematik olarak sunulmuştur.



Şekil 3.15 : Özyinelemeli mimari temelli yöntemlere örnek işlem akış şeması.

Hibrit metodlarda, projektif, volumetrik ve nokta bulutu temelli metodlarda uygulanan yöntemlerin bir veya birkaçı birlikte kullanılmaya çalışılmıştır. Farklı yöntemlerle elde edilen temsillerin eldeki nokta bulutunun tamamını temsil edebilecek global özellik vektörlerine dönüştürülmesi bu metodlar için ana araştırma konusu olmuştur (Guo ve diğ., 2021). Bu yaklaşımla geliştirilen modeller farklı algılayıcılar tarafından elde edilen mekansal verinin birlikte değerlendirilmesi uygulamaları için de başarılı birer örnek olmuştur (Xu ve diğ., 2018).

Nokta bulutlarından derin öğrenme alanının kısa tarihsel gelişimi incelendiğinde güncel çalışmalarda doğrudan nokta bulutu verisini girdi olarak kullanma yaklaşımının yaygınlaştığı görülmüştür. Bunun temel sebepleri olarak:

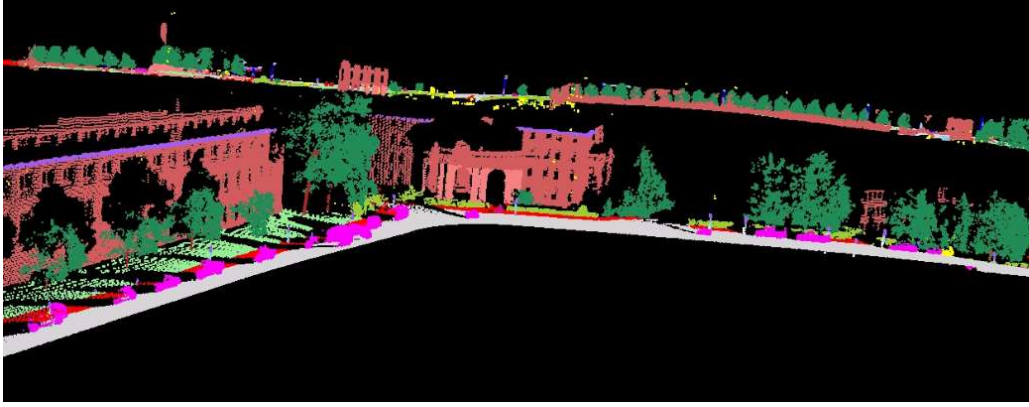
- Nokta bulutları daha düzenli bir kesikli ara temsil formatına dönüştürülürken yaşanan veri kayıpları,
- Nokta bulutlarının ara bir temsil elde etmek için veri dönüşüm işlemlerine ihtiyaç duyulmadan işlenebilmesi için çözümler geliştirilmiş olması,
- Nokta bulutu temelli metodların depolama ve hesaplama yükü bakımından ekstra yükümlülükler getirmemesi,
- Noktalar arasındaki mekansal ilişkilerden verinin en ham hali ile faydalanılmasına imkan tanınması görülmüştür.

3.4 Derin Öğrenme İle Semantik Segmentasyon 3B Nokta Bulutu Veri Setleri

Kontrollü derin öğrenme çalışmalarında genelleştirme başarısını arttırmak ve modellerin ezberlemesinin önüne geçmek için büyük veri setlerine ihtiyaç duyulur. Nokta bulutlarının semantik segmentasyonunu hedefleyen çalışmalar için yeterli büyüklükte nokta bazında etiketlenmiş veri gereklidir. İçerdiği nokta sayısı milyarlara,

depolama açısından boyutları onlarca, yüzlerce gigabyte'a (GB) bazen de terabyte'a (TB) ulaşabilecek eğitim verilerinin toplanması, hazırlanması ve sunulması aşamaları emek, zaman ve maliyet yoğun süreçler olduğundan bu alanda çalışma yapmak isteyenler tarafından mevcut açık veri setlerinin ve özelliklerinin bilinmesi faydalı olacaktır. Bu bölümde açık alanları özellikle meskun sahaları konu alan 3B semantik segmentasyon veri setlerinden örnekler sunulmuştur.

Oakland 3D Point Cloud Dataset; kentsel bir alanda (kampüs) mobil lazer tarayıcı sistem ile elde edilmiş yaklaşık 1.3 milyon eğitim, doksan bine yakın validasyon ve otuz bin civarında test noktasından oluşan bir veri setidir. Veri seti 5 ana kavramsal sınıf altında manuel olarak etiketlenmiştir. Bunlar; vejetasyon, kablo, yol, cephe ve direklerdir. Her nokta 3B koordinatlar, yansıma değerleri ve etiketler ile temsil edilmektedir (Munoz, ve diğ, 2009).



Şekil 3.16 : Oakland 3D Point Cloud Dataset veri örneği, Munoz ve diğ. (2009)'dan uyarlanmıştır.

Sydney Urban Objects Dataset, mobil lazer tarayıcı ile elde edilmiş bir veri setidir. Yol çevresi ile ilişkili yaya, taşıt, bisiklet, bisikletli, trafik ışığı, elektrik direği, ağaç gibi 26 farklı tematik sınıftan nesnelere ait noktalar bulundurmaktadır. Veri seti akıllı araçlar için gerçekçi algılama koşullarına örnek oluşturabilmek amacı ile oluşturulmuştur. Bu nedenle nesnelere ait farklı açılardan yapılmış alımlar ve örtüşmelerden dolayı kısmi eksiklikler içermektedir ve non-homojen nokta yoğunluğuna sahiptir. Veride noktalar 3B koordinatlarının yanı sıra algılama geometrisine ilişkin lazer id, azimut, orjine olan mesafe ve yansıma yoğunluğu ile de sunulmaktadır (De Deuge ve diğ, 2013).

ISPRS Vaihingen 3D Semantic Labeling Dataset, Vaihingen Kentinde HLT ile elde edilmiş yaklaşık 1.2 milyon noktadan oluşan bir veri setidir. Veri seti boyunca ortalama nokta yoğunluğu metrekare başına 6.7 metre olarak gerçekleşmiştir. Veri; 3B koordinat, yansımaya yoğunluğu, lazer dönüş numarası, lazer dönüşlerinin sayısı ve etiket halinde kayıt edilmiş noktalar halinde sunulmaktadır. Veri setinde; enerji iletim hattı, alçak boylu vejetasyon, su geçirmez yüzey(asfalt kaplı yol, otopark vs.), araç, çit, çatı, cephe, çalı ve ağaç olmak üzere 9 farklı kavramsal sınıf bulunmaktadır. (Niemeyer ve diğ, 2014).

IQmulus abd TerraMobilita, veriseti yoğun bir kentsel alanda MLT ile elde edilmiş 300 milyon noktadan meydana gelmektedir. Veri seti PLY formatında olmak üzere her nokta 3B koordinatları, yansımaya değeri, dönüş numarası ve sınıf bilgisi ile sunulmaktadır. Veri setinde tematik gruplar hiyerarşik biçimde oldukça detaylı biçimde tasarlanıp ayrıştırılmış sınıflar ile ifade edilmiştir. Ana sınıflar, diğer/arkaplan, yüzeyler ve objeler başlıkları ile kurgulanmıştır. Ana sınıflar kendi içlerinde detaylandırılmıştır (Vallet ve diğ, 2015).

Paris-rue-Madame Database, Paris'te 160 metre uzunluğunda bir sokak boyunca MLT ile alımı yapılmış bir nokta bulutudur. Ham verinin otomatik biçimde kabaca segmente edilmesi sonrasında sonuçların el ile iyileştirilerek etiketlenmiş bir veri setidir. Veri seti PLY formatında sunulmaktadır. Veri dosyaları içerisinde noktalar 3B koordinatları, yansımaya yoğunluk değerleri ve etiketleri ile saklanmaktadır. Veri 27 farklı tematik sınıfı içermektedir. Verinin içerisindeki tematik sınıflara örnek olarak; yol, yaya, cephe, araç, çöp kutusu, cepheler verilebilir (Serna ve diğ, 2014).

Semantic3D, 2017 yılında, yaklaşık dört milyar nokta ve sekiz tematik sınıf ile zamanının en büyük dış mekan nokta bulutu veri seti olarak ortaya çıkmıştır. Veri, İsviçrede bulunan; köy, sokak, meydan, kaleler gibi farklı karakter ve dokulardaki meskun sahalarda yersel lazer tarayıcı kullanılarak oluşturulan nokta bulutlarının nokta bazında etiketlenmesi ile elde edilmiştir. Veri seti Semantic-8 ve Reduced-8 adında iki versiyon halinde dağıtılmaktadır. Bu iki versiyon aynı eğitim bölütüne sahiptir fakat Reduced-8 daha küçük bir test bölütü içermektedir. YLT ile alım yapıldığından düşey yüzeylere dair nokta sayısı fazla, çatı ve binaların üst yüzeyleri gibi yer seviyesinden yüksek yüzeylere ilişkin veri yoğunluğu azdır. Veri içeriğindeki sekiz tematik sınıf; yapay yüzey, doğal yüzey, yüksek bitki, alçak bitki, binalar, araçlar, tarama artefaktları ve sert peyzaj objeleridir (Hackel ve diğ, 2017).



Şekil 3.17 : Semantic3D veri setinden bir örnek.

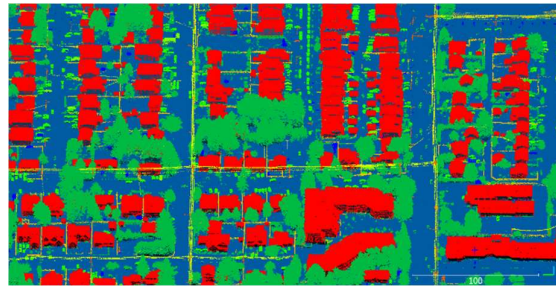
Paris Lille 3D. Bu veri seti bir MLT sistemi ile toplanmıştır. Toplamda yaklaşık 1940 metrelik 3 farklı alım kampanyasından elde edilen 143.1 milyon noktadan oluşmuştur. Veride tematik sınıflar tasarlanırken IQmulus&TerraMobilita veri setinin hiyerarşik sınıf ağacı temel alınmış, eldeki veriye uyarlamak için bazı sınıflarda değişiklik yapılmış ve yeni sınıflar eklenmiştir. Bu sınıf yapısı özellikle otonom araçlar için detaylı bir çevresel farkındalık yaratabilecek öğrenme süreçlerini yürütmeye uygunluk hedeflenerek tasarlanmıştır. Öyle ki objeler trafik ve çevre ile ilişkilerine göre hareketli, sabit ya da park etmiş olarak ayırt edilerek 50 farklı sınıf ile sunulmuştur. Veri seti PLY formatında her nokta için 10 özellik sütunu ile paylaşılmaktadır. Bunlar noktanın 3B koordinatları, lazer tarayıcının 3B koordinatları, alım anındaki GPS zamanı, yansıma yoğunluğu, ana sınıf etiketi ve detay sınıf etiketidir (Roynard ve diğ, 2018).

SemanticKITTI, Karlsruhe Institute of Technology (KIT) ve Toyota Technological Institute (TTI) ortaklığı ile 2012 yılında otonom araç alanındaki çalışmaların test edilmesi amacıyla başlatılmış ve Geiger ve diğ. (2012), tarafından yürütülen KITTI Vision Odometry Benchmark projesinde referans olması için toplanan mobil lazer tarama verileri nokta bazında etiketlenerek hazırlanmıştır. Veri seti; yer, yapı, taşıt, doğa, insan, nesne ve diğer olmak üzere 7 ana sınıfta bulunan 28 tematik grup tanımlamaktadır. Veri bir MLT tarafından her bir tarama noktasında yapılan 360 derecelik alımla elde edilen noktalardan oluşmaktadır. Doğal trafik ortamını yansıtan, yaklaşık 4.5 milyar nokta ile büyük bir veri kaynağıdır (Behley ve diğ, 2019).

Toronto3D, kentsel alanda bulunan yaklaşık 1km’lik bir rotada MLT ile toplanmış 78.3 milyon noktadan oluşan bir veri setidir. Veri setinde noktalar 3B koordinatları, RGB değerleri, lazer yansıma yoğunluğu, GPS zamanı ve tarama açısı numarası ile sunulmaktadır. Veri setinde gerçek trafik koşullarını yansıtıcı biçimde alım sırasında algılayıcı ve çevredeki objelerin hareketleri sonucu kısmi eksiklikler bulunmaktadır. Veri seti; yol, yol işareti, doğa, yapı, altyapı hattı, direk, araç, çit ve sınıflandırılmamış olarak 10 farklı tematik sınıf tanımlanmaktadır (Tan ve diğ, 2020).

SemanticPOSS, otonom sürüş sistemlerinin veri gereksinimleri göz önüne alınarak geliştirilmiş bir veri setidir. 2988 alım noktasında yaklaşık 216 milyon noktadan oluşan veri seti Pekin Üniversitesi kampüsünde toplanmıştır. Veri seti, doğal trafik koşullarında toplanmış ve 14 kavramsal sınıf içermektedir. Veri setinin sunum formatı SemanticKITTI nokta bulutu veri seti ile aynıdır. Veri setinde objeler sınıf etiketleri dışında örnek id’leri ile de tutulduğundan nesne takibi uygulamaları için de uygun bir eğitim veri setidir (Pan ve diğ, 2020).

DALES (Dayton Annotated Laser Earth Scan) , geniş çaplı bir sahayı kapsayan HLT kaynaklı bir veri setidir. Veri kentsel, ticari, kırsal gibi farklı karakter ve dokulara sahip sahalarda toplanmıştır. On kilometre kare büyüklüğünde bir alanda 1300 metre uçuş yüksekliği ile toplanan yaklaşık 500 milyona yakın noktadan oluşan veri setidir. 3B kentsel modelleme ve ormancılık gibi geniş alanların gözlemlenmesine gereksinim duyulan alanlarda geliştirilecek derin öğrenme çalışmalarına eğitim verisi sağlamak amacı ile oluşturulmuştur. Derin öğrenme çalışmalarına kolaylık sağlamak için veri 0.5 kilometre karelik 40 mozağe bölünmüştür. 30/70 oranında eğitim ve test seti olarak ayrılmış olan veri PLY formatında sunulmaktadır. Veri seti yer, vejetasyon, otomobil, kamyon, elektrik iletim hattı, direkler, çitler, yapılar olmak üzere 8 tematik sınıfa ayrılmıştır. Ek olarak sınıflandırılmayan fakat bütünlüğü korumak amacıyla veride tutulan diğer noktalar sınıfı, bulunmaktadır (Varney ve diğ, 2020).



Şekil 3.18 : DALES veri setinden bir örnek.

DublinCity, Dublin şehir merkezinde tarihi ve modern kentsel dokuların farklı örneklerini içeren HLT ile üretilmiş bir nokta bulutu verisidir. Yaklaşık 260 milyon noktadan oluşan veri seti metrekaare başına 250 ila 348 nokta arasında değişen yoğunluğa sahiptir. Farklı tarihi dönemlerden çok çeşitli formlarda yapıları içeren bu yüksek yoğunluklu nokta bulutu kentsel alanlar üzerine yapılacak derin öğrenme çalışmaları için değerli bir veri kaynağı olmaktadır. Veri seti hiyerarşik biçimde 3 seviyede, 13 farklı tematik sınıfa ayrılmış durumdadır. İlk seviye daha geniş tanımlı 4 sınıf içermektedir. Bunlar; yapı, yer, vejetasyon, tanımsız olarak adlandırılmıştır. 3. seviyeye giden yolda noktalar git gide daha detaylı şekilde sınıflandırılmıştır. Örneğin ilk seviyede yapı olarak sınıflandırılan bir nokta ileriki seviyelerde sırasıyla çatı ya da cephe ve kapı, duvar, pencere gibi daha yüksek ayrıntı seviyesinde etiketlenmiştir (Zolanvari ve diğ, 2019).

Hessigheim 3D (H3D) , Almanya'nın Hessigheim kasabasında 4 farklı epokta HLT ve dijital kamera ile elde edilen eğik resimlerden üretilmiş yüksek yoğunluklu bir veri setidir. Veri seti sayısal kamera ve lazer tarayıcı verilerinin füzyonu ile üretildiğinden çatı ve cephe yüzeyleri hakkında daha bütüncül bilgi içermektedir. Metrekare başına 800 nokta ile oldukça yoğun bir veri setidir. H3D, aynı şekilde ISPRS çalışma grupları tarafından üretilen Vaihingen 3D veri setine benzer bir tematik sınıf yapısına sahiptir. Bu verinin özel koşulları neticesinde kent mobilyası, toprak/çakıllık, düşey yüzey ve baca sınıfları şeklinde yeniden düzenlemeye gidilmiş ve 11 farklı sınıf tanımına ulaşmıştır. Veri seti yaklaşık 61 milyon noktadan oluşmaktadır (Kölle ve diğ, 2021).

LASDU veri seti, ağırlıklı kentsel ve endüstriyel dokuya sahip bir bölgede HLT ile elde edilmiştir. Yaklaşık 1200 metre irtifadan gerçekleştirilen alımda metrekaare başına ortalama 4 nokta yoğunluğa sahip nokta bulutu 1 kilometrekarelik bir alanı kapsamaktadır ve yaklaşık 3 milyon nokta içermektedir. Veri noktaları el ile etiketlenmiştir. Veri setinde semantik segmentasyon faaliyetleri için 5 farklı tematik sınıf ve bu sınıflara yerleştirilemeyen objeleri gruplamak için bir arkaplan sınıfı tanımlanmıştır (Ye ve diğ, 2020).

Campus3D, Li ve diğ (2020), Singapur Ulusal Üniversitesi kampüsünde insansız hava aracı ile elde edilmiş hava fotoğraflarından 1.58 kilometrekarelik bir alanı kapsayacak biçimde üretilmiş nokta bulutu veri setidir. Nokta bulutu, otomatik görüntü eşleme yöntemi ile Structure from Motion(SfM) tekniği kullanılarak üretilmiştir. Veri

setindeki tüm noktalar hem sınıf bazında hem de tekil nesne bazında etiketlenmiştir. Bu da semantik segmentasyonun da ötesinde tekil örnek/instance segmentasyon uygulamalarının geliştirilmesine imkan tanımaktadır. Veri seti yaklaşık 937 milyon noktadan oluşmaktadır. Sınıf yapısı olarak, veri seti 5 seviyeli bir hiyerarşi ile 24 gruba ayrılmıştır. İlk seviyede noktalar yapı, yer ve diğer olarak ayrılmış ilerleyen seviyelerde ilişkili oldukları sınıflara göre detaylandırılmışlardır.

SensatUrban veri seti kentsel ölçekte, insansız hava aracı kullanılarak yüksek çözünürlüklü dijital hava fotoğraflarından elde edilmiş bir veri setidir. Mevutta en çok nokta sayısına sahip veri setidir. Veri seti, Birleşik Krallık'ta yer alan Birmingham, Cambridge ve York şehirlerinde yer alan büyük kentsel alanları içermektedir. Nokta bulutu yaklaşık 7.6 kilometre karelik alan kaplamakta, 2.8 milyar nokta içermektedir. Veri seti üretilirken, noktalara ait gri değer bilgisinin semantik segmentasyon başarısına etkisi ve farklı kentsel dokularda eğitilmiş ağların şehirler arası veri setlerinde kullanılabilirliği araştırma soruları göz önünde tutulmuştur. Veri seti hem nadir hem de eğik alım görüntülerinden oluşturulduğundan hem yatay hem düşey yüzeylerde bakımından bütüncül niteliktedir. Nokta bulutu otomatik olarak görüntü eşleme ve SfM tekniklerinden yararlanılarak üretilmiştir. Etiketleme işlemi ise el ile yaklaşık 600 iş saati sarf edilerek gerçekleştirilmiştir. Veri setinde 13 farklı tematik sınıf tanımlanmıştır. Tematik sınıflar tanımlanırken ilk olarak her sınıfın açık ve net bir tanımı olmasına ikinci olarak da her sınıfın bir sosyal ya da ekonomik faaliyet alanına hitap etmesine dikkat edilmiştir. Tematik sınıflar; yer, vejetasyon, yapı, duvar, köprü, otopark, karayolu, şehir mobilyası, araç, yayayolu, bisiklet ve su yüzeyleri olarak tanımlanmıştır. Veri seti eğitim süreçlerinde pratik olarak kullanım imkanı sağlamak için 400 metrekairelik mozaiklere bölünmüştür (Hu ve diğ, 2021).

Grand Theft Auto-V (2014) (GTA-V), popüler açık dünya video oyunu zengin, canlı ve detaylı şehir modeli ile araştırmacıların dikkatini çekmiştir. GTA-V video oyununu ya da başka modern açık dünya video oyunlarını simulasyon ortamı olarak kullanarak hızlı biçimde sentetik görüntü segmentasyonu veri seti elde edilebileceğini gösteren çalışmalar yapılmıştır (Richter ve diğ, 2016) ile (Johnson-Roberson ve diğ, 2017). Bu çalışmalara atıf ile aynı video oyununda bir mobil lazer tarayıcı bakış açısı, açısız çözünürlük gibi parametreler ile tanımlandıktan sonra alım kampanyası kurgulanmıştır. Nokta bulutu üretilirken lazer tarayıcı tarafından yayılan ışınların objelere temas ettiği noktaların 3B koordinatları, bu objelerin sınıfları ve her objenin

tekil nesne id'si otomatik olarak elde edilmiştir. Sentetik verinin gerçekliğini arttırmak amacıyla SemanticKITTI veri setindeki gürültü analiz edilmiş ve sentetik veriye aktarılmıştır (Yue ve diğ. 2018).

SynthCity, Griffiths ve Boehm (2019), büyük ölçekte nokta bazında etiketlenmiş nokta bulutu verisi üretimi işinin zahmetli ve maliyetli olması sebebi ile sentetik olarak üretilmiş bir veri setidir. Veri seti üretilirken açık kaynaklı Blender (2018) isimli 3B bilgisayar grafiği programı için Gschwandtner ve diğ. (2011) tarafından geliştirilmiş BlenSor eklentisinden faydalanılmıştır. Açık kaynaklı 3B realistik modeller, gerçekçi bir kent ortamı oluşturmak için Blender'da bir araya getirilmiştir. BlenSor eklentisi ile hareketli bir lazer tarayıcı tasarlanmış, kullanılan 3B modellere ait yüzeyler 9 farklı tematik sınıfla ilişkilendirilmiştir ve veri seti MLT ölçme uygulaması şeklinde kurgulanmıştır. Bu sayede nokta bulutunun etiketlenmesi işlemi tümüyle otomatik biçimde tamamlanmıştır. Sentetik verilerin bilgisayarlı görü ağlarının ön eğitiminde kullanılması daha önce denenmiştir (Rajpura ve diğ, 2017). Veri seti, 9 tematik sınıftan yaklaşık 368 milyon noktadan oluşmaktadır. Veriye gerçekçilik katmak için ölçme rotasına ve nokta koordinatlarına gauss dağılımlı gürültü eklenmiştir. Veri seti için oluşturulan kent modeli, Kuzey Amerika ve Avrupa kent dokusunu yansıtmaktadır.

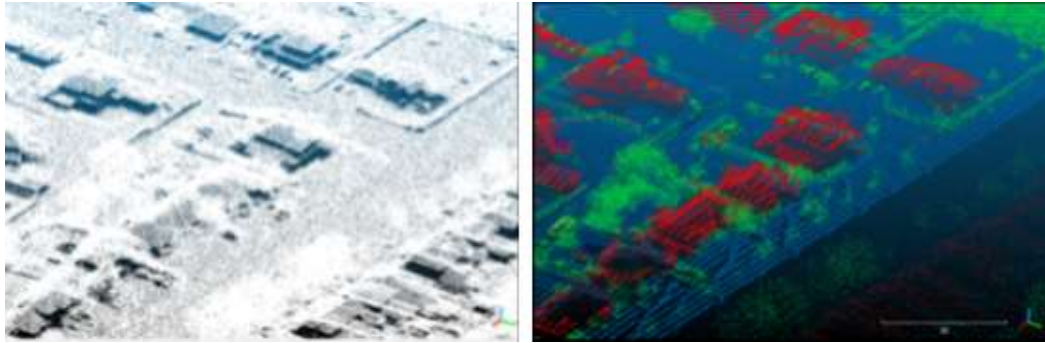
Bu bölümde bahsedilen veri setleri çeşitli gruplar tarafından halka açık biçimde kullanılması amacı ile üretilmiştir. Bu veri setleri farklı kullanım ve paylaşım şartları getiren açık kaynak lisans çeşitleri, örneğin Creative Commons lisansları ile yayınlanabilmektedir. Veri setlerini çalışmalarında kullanacak araştırmacıların verilerin herhangi bir teknik garanti ile paylaşılmadığına, gürültü ve hatalar barındırabileceğine, üretildikleri ortamların karakteristik özelliklerine, içerdikleri tematik sınıfların oransız ağırlıklarda temsil edilebileceğine dikkat etmelidir.

Farklı algılayıcı ya da üretim tekniklerinden elde edilen nokta bulutları, farklı geometrik ve radyometrik özelliklere sahip olurlar. Bu nedenle uygulama alanları ve pratiklerinde de farklılıklar görülebilmektedir. Örneğin, sayısal görüntülerden elde edilen nokta bulutları LIDAR kaynaklı nokta bulutlarından farklı karakterde olurlar (Baltsavias, 1999). RGB-D kameralar, piksel bazında derinlik bilgisini kayıt ederler fakat yapısal özellikleri gereği efektif menzilleri kısadır genellikle bina içi ya da algılayıcı-hedef mesafesinin kısa olduğu alanlarda tercih edilirler. LIDAR teknolojisi kullanılarak elde edilen nokta bulutları ise algılayıcı sistem ve algılama platformuna bağlı olarak hem spektral özellikler hem çözünürlük hem de uzamsal dağılım

yönünden diğer kaynaklardan farklılaşır. Bunların yanı sıra uydu görüntüleri (Han ve diğ, 2020) ve yapay açıklıklı radar (Zhu ve Bamler, 2010) gibi uzaktan algılama teknikleri de pek çok açıdan farklı özelliklerde nokta bulutları üretilmesini sağlarlar (Yu ve diğ, 2015). Bu sebeple tek bir model ile tüm uygulama alanlarındaki görevlerde kullanılabilecek çözümler geliştirilmesi mümkün değildir. Derin öğrenme ağlarının eğitilmesinde kullanılacak verinin seçiminde hedeflenen uygulamaların gereksinimleri gözetilmeli ve uygun veri setleri tespit edilmelidir.

3.5 3B Nokta Bulutu Semantik Segmentasyonunda Son Teknoloji Çözümler

Nokta bulutları nesnelere ait yüzeyleri yüksek çözünürlükle ifade edebilen 3B mekânsal bilgi kaynaklarıdır. Bu sebeple 3B modelleme çalışmalarında nesne geometrilerini elde etme ihtiyacı bakımından değerli bir veri olmaktadır. Modelleme çalışmalarında ihtiyaç duyulan geometrilerin oluşturulmasının otomasyonunda, noktaların yeryüzünde karşılık geldikleri varlıklar ile ilişkilendirilmeleri aşaması önemli bir adımdır. Semantik segmentasyon adı verilen bu işlem şekil 3.19’da görselleştirilmiştir.



Şekil 3.19 : Nokta bulutu semantik segmentasyonu öncesi ve sonrasını temsil eden görsel.

Derin öğrenme öncesi dönemde, nokta bulutlarının segmentasyonu için noktalar arasındaki mekansal ilişkileri geometrik ve/veya istatistiksel açıdan ifade etmeyi hedefleyen yöntemler tasarlanmıştır. Bu çalışmalara; kenar tabanlı, alan büyüme tabanlı ve nokta gruplarının oluşturduğu temel geometrik şekilleri tespit etmeye dayalı yöntemler örnek olarak verilebilir. Veri ve hesaplama imkanlarının artması ve makine öğrenmesi alanında yaşanan gelişmeler neticesinde derin öğrenme yaklaşımları 3B nokta bulutlarının işlenmesinde ve özel olarak semantik segmentasyonunda kullanılmaya başlamıştır (Xie ve diğ, 2020).

3B Nokta bulutlarının derin öğrenme ile semantik segmentasyonu görevi için geliştirilen çözümlerin kendi içlerinde genel olarak projektif yöntemler, volumetrik yöntemler ve nokta temelli yöntemler olarak ana gruplara ayrıldığı görülmüştür.

Bir önceki bölümde tanıtılan çalışmalara benzer şekilde projektif yöntemlerde nokta bulutlarına dair çoklu bakış noktalarından elde edilen görüntüler 2B evrimsel sinir ağları ile özellik çıkarımına tabi tutulmuştur. Volumetrik yöntemlerde nokta bulutları vokseller gibi düzenli ve ayrık veri tiplerine dönüştürülmüş ve 3B evrim işlemi ile işlenmiştir.

Nokta bulutu temelli metodlar, herhangi bir ara temsile dönüştürme zorunluluğunu gerektirmemeleri, depolama yönünden ve hesaplama bakımından ek yükümlülüklerle sebep olmamaları, ham algılayıcı verisine en yakın veri tipi olması ve mevcut hali ile nokta grupları tarafından ifade edilen mekansal ilişkilerden faydalanmaya imkan tanımaları sebepleri ile son dönemde popüler hale gelmektedir. Bu çalışmada bahsedilen anlayışa paralel olarak nokta bulutu temelli metodlardan faydalanan çalışmalar üzerinde durulmuştur.

Semantik segmentasyon için geliştirilen modeller arasında son teknoloji olarak nitelendirilebilecek ürünleri tespit etmek amacıyla büyük yerleşim alanlarını kapsayan nokta bulutu veri setlerine ilişkin raporlanmış performanslar incelenmiştir. Referans veri setleri seçilirken çalışmanın teması olan kentsel alanları konu alan içeriğe sahip nokta bulutları tercih edilmiştir. Yapılan inceleme sonucunda Hessigheim3D, DALES ve SensatUrban veri setleri içerik olarak araştırma temasına uygun olmaları sebebi ile seçilmiştir. Bu veri setlerine ilişkin genel bilgiler Çizelge 3.1’de özet olarak sunulmuştur. Seçilen veri setlerini işlemede kullanılmış modellere dair erişilen performans sonuçları Çizelge 3.2, Çizelge 3.3 ve Çizelge 3.4’te sunulmuştur.

Çizelge 3.1 : Seçilen veri setlerine ilişkin genel bilgiler.

Veri Seti	Veri Toplama Yöntemi	Nokta Sayısı	Veri Kapsam Alanı	Veri içeriği
Hessigheim3D	HLT	~73 Milyon	~135.000 m2	X,Y,Z, yansıma,R,G,B, dönüş sayısı, dönüş numarası
DALES	HLT	~505 Milyon	~10 km2	X,Y,Z, yansıma
SensatUrban	Hava Fotoğrafı	~2847 Milyon	~7.6 km2	X,Y,Z,R,G,B

Modellerin veri setleri üzerindeki sınıf bazında ve ortalama performansları karşılaştırıldığında Kernel Point Convolution (KPConv) (Thomas ve diğ, 2019), Random Local Attention Network (RandLANet) (Hu ve diğ, 2021) ve ConvPoint (Boulch, 2020) modellerinin kullanılan performans ölçütlerine göre ön plana çıktığı görülmüştür. En yüksek başarımlar değerlerine ulaştıkları görülen bu modellerin tamamının nokta temelli yöntemleri esas aldığı tespit edilmiştir.

Çizelge 3.2 : Hessigheim3D veri seti için raporlanan performans sonuçları Url-3'den alınmıştır.

Model veya Yöntem	F1 Skorları (%)										
	Kısa Boylu Bitki	Su geçirmez yüzey	Araç	Şehir Mobilyası	Çatı	Cephe	Çalı	Ağaç	Toprak Yüzey	Düşey Yüzey	Baca
Rand. Forest	90.36	88.55	66.89	51.55	96.06	78.47	67.25	95.91	47.91	59.73	80.65
SCN	92.31	88.14	63.51	57.17	96.86	83.19	68.59	96.98	44.81	78.2	73.61
KPConv	88.57	88.93	82.1	63.89	97.13	85.13	75.24	97.38	42.68	80.87	0
PointNet++	78.11	72.07	31.78	13.65	73.98	47.79	28.34	71.8	9.65	21.67	4.39

Hessigheim3D, SensatUrban ve DALES veri setlerine göre çok daha küçük çapta bir veri setidir. Bu veri seti için performans ölçütleri kesinlik (precision) ve duyarlılık(recall) değerlerinin harmonik ortalaması olan F1 skoru ile belirtilmiştir. Veri setinde noktalar 3B koordinatları, RGB değerleri, lazer ışını dönüş sayıları ve lazer ışını yansıma değerleri ile ifade edilmiştir. KPConv modeli; veri setinde bulunan kısa boylu bitki (çimen, çalı vb.), toprak yüzey ve baca sınıfları haricinde diğer sınıflarda en yüksek F1 skoruna ulaşan model olmuştur. Bu model asfalt yol/otopark ve beton zeminler gibi yüzeyleri içeren su geçirmez yüzey sınıfında, yapılara ilişkin yüzeyleri içeren çatı ve cephe sınıflarında ve ağaç sınıfında en yüksek F1 skorunu elde etmeyi başarmıştır. Araştırmacılar su geçirmez yüzey, toprak yüzey ve kısa boylu bitki sınıflarına dahil edilen bazı noktaların benzer geometrik ve radyometrik özellikler göstermesinden dolayı eğitilen modellerin nokta sayısı daha fazla olan sınıf yönünde yanlış segmentasyon sonuçları ürettiğini açıklamıştır. Çalı ve ağaç sınıflarına karşılık

gelen noktalar da benzer radyometrik ve geometrik özelliklere sahip olduklarından eğitilen modellerin daha yüksek ağırlıkla temsil edilen ağaç sınıfı lehine segmentasyon sonuçlar ürettiği raporlanmıştır. Şehir mobilyası sınıfına dahil nesnelerin de semantik segmentasyonu modeller için önemli bir zorluk olmuştur. Bu sınıfa dahil nesnelerin hem daha az sayıda nokta ile temsil edilmesi hem de şekilsel özellik bakımından farklı sınıflardan nesnelere (örneğin balkon) benzerlik göstermeleri segmentasyon zorluğunu arttıran bir etmen olarak ortaya çıkmıştır.

Çizelge 3.3 : DALES veri seti için raporlanan performans sonuçları (Varney ve diğ., 2020).

Model	ort. IoU %	Sınıf Bazında IoU değerleri (%)							
		Zemin	Yapı	Araba	Kamyon	Direkler	Enerji hattı	Çit	Bitki
KPConv	81.1	97.1	96.6	85.3	41.9	75	95.5	63.5	94.1
PointNet++	68.3	94.1	89.1	75.4	30.3	40	79.9	46.2	91.2
ConvPoint	67.4	96.9	96.3	75.5	21.7	40.3	86.7	29.6	91.9
SuperPoint	60.6	94.7	93.4	62.9	18.7	28.5	65.2	33.6	87.9
PointCNN	58.4	97.5	95.7	40.6	4.8	57.6	26.7	52.6	91.7
ShellNet	57.4	96	95.4	32.2	39.6	20	27.4	60	88.4

DALES veri setinde KPConv modelinin hem tüm sınıflar ortalaması olarak hem de her sınıf için tekil bazda en yüksek ortalama IoU değeri elde ettiği görülmüştür. Noktalara dair özellik vektörlerinin yalnızca 3B koordinatlardan ve lazer ışını yansıma değerlerinden oluştuğu bu veri setinde modeller; nokta sayısının çok olduğu zemin, yapı ve bitki sınıfları için genel olarak iyi başarı göstermiştir. Veri setinde daha az

nokta ile yer alan kamyon, direk, çit ve enerji nakil hattı sınıflarında tüm modeller daha düşük ortalama IoU değerleri elde etmiştir. KPConv modeli, genel başarı oranının daha düşük olduğu kavramsal sınıflarda da en yakın rakibinden daha büyük ortalama IoU değeri elde edebilecek performans göstermiştir. Özellikle diğer modellerin başarı oranının çok düşük olduğu enerji nakil hattı noktalarında %95.5 ortalama IoU değeri elde etmeyi başarmıştır.

SensatUrban veri seti için bildirilen performans sonuçları incelendiğinde KPConv modelinin 13 semantik sınıftan 8 tanesinde en yüksek mIoU değerini elde ettiği görülmüştür. Modellere ait performans ölçütleri hem yalnızca 3B koordinatlar kullanılarak hem de noktalara ait RGB değerlerinin de dahil edilerek yürütülen eğitimler sonucu elde edilen ortalama IoU değerleri ile sunulmuştur. Kullanılan modellerin genel olarak zemin, yapı ve bitki semantik sınıflarına ait noktaları, diğer sınıflara karşılık gelen noktalara nazaran daha yüksek doğrulukla tespit edebildiği görülmüştür. Makine öğrenmesi, özellikle derin öğrenme çalışmalarında veri setinde bulunan sınıflara ait örneklerin sayısının dengesizliği modellerin başarısını etkilemektedir. Bisiklet sınıfının en az sayıda nokta ile temsil edilen sınıf olmasının modellerin bu semantik sınıf için sıfıra yakın mIoU değerleri elde etmesine sebep olduğu düşünülmüştür. Demiryolu, otopark, köprü, şehir mobilyası ve yayayolu düşük ortalama IoU değerleri elde edilebilen kavramsal sınıflar olmuştur. RGB değerlerinin dahil edildiği ve edilmediği eğitim sonuçları karşılaştırıldığında, radyometrik özelliklerin modellere dahil edilmesinin başarıyı arttırdığı görülmüştür. Otopark, karayolu, köprü, otoyol gibi benzer geometrik ve uzamsal özellikler gösteren sınıflar için radyometrik büyüklüklerin dahil edilmesiyle daha yüksek ortalama IoU değerleri elde edilebildiği tespit edilmiş yalnızca uzamsal boyutları ile ayırt edilmesi zor kavramsal sınıfları içeren uygulamalarda radyometrik verinin kullanılmasının faydalı olacağı gösterilmiştir.

Veri setlerinin, modellerin ve performans testi sonuçlarının incelenmesinin ardından KPConv, RandLANet, SCN (Graham ve diğ, 2018), ConvPoint, PointNet++ modellerinin diğer modellere nazaran daha başarılı sonuçlar elde ettiği görülmüştür. Çalışmamızın teması olan kentsel alanların modellenmesi alanına uygun veri setlerinin tamamında en başarılı performans değerlerine ulaşılmasını sağlayan KPConv, (Kernel Point Convolution) modelinin daha detaylı incelenmesine ve uygulama imkanlarının araştırılmasına karar verilmiştir.

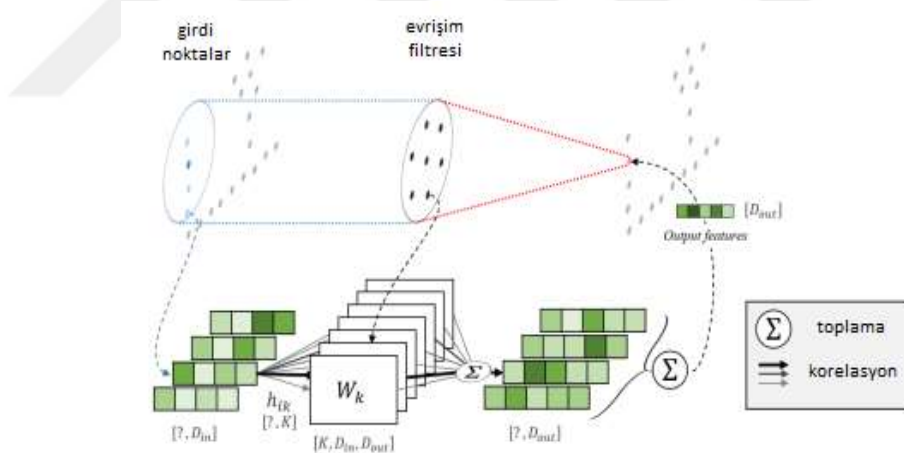
Çizelge 3.4 : SensatUrban veri seti için eğitim sonuçları (Hu ve diğ, 2021).

Model	OA(%)	mAcc(%)	mIoU(%)	Sınıf Bazında Ortalama IoU (%)												
				Zemin	Bitki	Bina	Duvar	Köprü	Otopark	Demiryolu	Karayolu	Şehir Mobilyası	Araç	Yayayolu	Bisiklet	Su
PointNet (- RGB)	83.5	33.52	28.85	67.35	92.66	84.72	16.02	0	13.65	2.68	17.09	0.33	54.54	0	0	26.04
PointNet(+ RGB)	90.57	56.3	49.69	83.55	97.67	90.66	22.56	43.54	40.35	9.29	50.74	29.58	68.24	29.27	0	80.55
PointNet++ (- RGB)	90.85	56.94	50.71	79.05	98.37	94.22	66.76	39.74	37.51	0	51.53	38.82	81.71	5.8	0	65.68
PointNet++ (+ RGB)	93.1	64.96	58.13	86.38	98.76	94.72	65.91	50.41	50.53	0	58.4	46.95	82.31	38.4	0	82.88
SPGraph (- RGB)	84.81	42.12	35.29	69.6	94.18	88.15	34.55	20.53	15.83	16.34	31.44	10.54	55.01	0.98	0	21.57
SPGraph (+ RGB)	85.27	44.39	37.29	69.93	94.55	88.87	32.83	12.58	15.77	15.48	30.63	22.96	56.42	0.54	0	44.24
KPConv (- RGB)	91.47	57.43	51.79	80.43	98.82	94.93	74.17	44.53	32.11	0	54.32	37.83	84.88	14.48	0	56.79
KPConv (+ RGB)	93.92	71.44	64.5	87.04	99.01	96.31	77.73	58.87	49.88	37.84	62.74	56.6	86.55	44.86	0	81.01
RandLA-Net (- RGB)	88.9	67.96	51.53	77.3	97.92	91.24	51.94	47.46	45.04	9.71	49.79	34.21	79.97	21.13	0	64.18
RandLA-Net (+ RGB)	91.24	74.68	58.14	82.23	98.39	92.69	56.62	49	54.19	25.10	60.98	38.69	83.42	38.74	0	75.8

3.6 Kernel Point Convolution (KPConv) Modelinin İncelenmesi

Kernel Point Convolution (KPConv), nokta bulutlarının sınıflandırılması ve semantik segmentasyonu için geliştirilmiş bir 3B nokta evrişim katmanıdır. KPConv çalışmasında araştırmacılar sınıflandırma için KP-CNN (Kernel Point Convolutional Neural Network) ve semantik segmentasyon için KP-FCNN (Kernel Point Fully Convolutional Neural Network) adını verdikleri iki mimari geliştirmiştir.

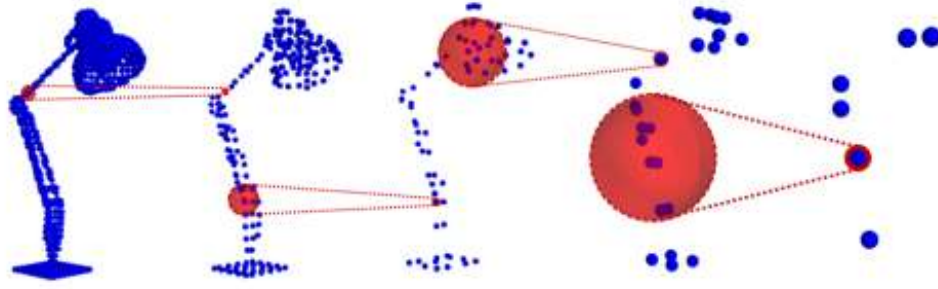
Araştırmacılar bu çalışmada nokta bulutlarının değişken yoğunluklu ve düzensiz yapısına uygun bir çözüm geliştirmeyi hedeflemişlerdir. Nokta bulutları, verideki yerel mekansal ilişkileri tespit etmede başarılı klasik 2B evrişim işlemine uygun yapıda değildirler. Bu sebeple araştırmacılar veri setindeki nokta gruplarının oluşturduğu şekilsel özelliklerden faydalanmak için herhangi bir ara temsile dönüştürme işlemi gerektirmeyen Kernel Point Convolution (Çekirdek Nokta Evrişimi) adını verdikleri özel bir 3B nokta evrişim işlemi tasarlamışlardır.



Şekil 3.20 : KPConv katmanının işlem prensibinin 2B üzerinde görselleştirilmesi
Thomas ve diğ. (2019)’dan uyarlanmıştır.

KPConv katmanlarında evrişim filtresi şekil 3.20’de de görülebileceği gibi; yarıçapı bir hiperparametre ile katman derinliğine göre kontrol edilen küreler içerisine homojen biçimde yerleştirilmiş noktalar ile temsil edilen filtre elemanlarından oluşacak şekilde tasarlanmıştır. Filtre nokta sayısı da farklı bir hiperparametre ile düzenlenmiştir. Evrişim işlemi için ilk olarak filtre, komşuluk alanının merkez noktasında konumlandırılmıştır. Komşuluk alanındaki noktalar ve filtre elemanları arasındaki 3B öklid mesafeleri hesaplanmıştır. Bu mesafeler nokta özellik vektörlerinin ağırlıklandırılmasında kullanılmıştır. Ağırlıklandırılmış özellik vektörleri filtre

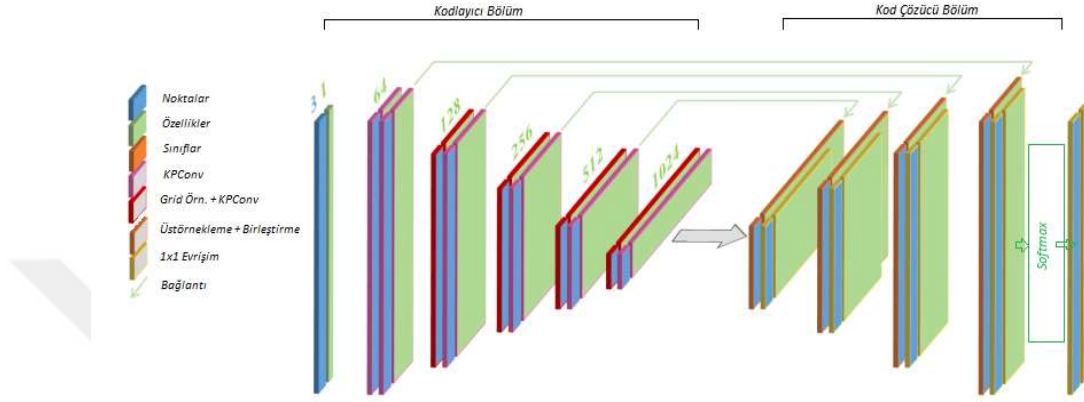
elemanı noktalar ile ilişkilendirilmiş eğitilebilir parametreler ile işleme sokulup her nokta için çıktı özellik vektörleri hesaplanmıştır. Bu özellik vektörleri toplanarak komşuluk alanını temsilen, merkez nokta ile ilişkilendirilen bir özellik vektörü elde edilmiştir. Bu sayede nokta bulutlarının permütasyonlarına duyarsız bir evrişim işlemi tasarlanmıştır.



Şekil 3.21 : Ağ boyunca altörnekleme ile nokta sayısının azaltılması ve yerel komşuluk alanlarının genişletilmesi Thomas ve diğ. (2019)’dan alınmıştır.

Evrişimsel mimaride; ağ boyunca artan derinlikle, orjinal girdiyi mekansal ilişkilerden faydalanarak daha az sayıda veri noktası ve daha soyut özellik vektörleri ile ifade edebilecek temsiller oluşturulması hedeflenir. KPConv çalışmasında oluşturulan modelde de bu yaklaşımı uygulamak için girdi nokta bulutu her katmanda bir alt örnekleme işlemine tabi tutulmuş ve nokta sayısı azaltılmıştır. Bununla birlikte daha geniş komşuluk alanları tanımlanarak filtrelerin algılama alanları (receptive field) şekil 3.21’deki gibi büyütülmüştür. Ağ boyunca çıktı özellik vektörlerinin boyutları artırılarak basitten karmaşığa bir temsiller hiyerarşisi elde edilmiştir. KPConv çalışmasında alt örnekleme yöntemi olarak grid altörnekleme algoritması tercih edilmiştir. Bu algoritmada nokta bulutu uygun boyut ve çözünürlükte bir 3B grid ile çakıştırılır. Boş olmayan grid hücrelerinin ağırlık merkezleri kayıt edilir ve aynı nesneyi daha az noktayla temsil eden bir nokta bulutu elde edilir. KPConv çalışmasında grid örneklemeden her katmanda, grid hücreleri genişletilip grid çözünürlüğü düşürülerek nokta sayısının azaltılması ile faydalanılmıştır. Evrişim filtresinin algılama alanını genişletmek için komşuluk alanı ve filtre yarıçapları her katmanda bir hiperparametre ile kontrollü biçimde büyütülmüştür. Grid altörnekleme diğer çalışmalarda yaygın biçimde kullanılan ve iteratif bir yöntem olan en uzak nokta örnekleme algoritmasına göre daha az işlem yüküne sahip ve daha basit bir algoritma olduğu için tercih edilmiştir.

KPConv katmanlarında elde edilen komşuluk alanlarına dair özellik vektörlerini bir araya getirmek için havuzlama(pooling) işlemi yerine hemen grid altörnekleme işlemi sonucu elde edilen noktalar tekrar bir KPConv katmanı ile işlenmesi yöntemi tercih edilmiştir.



Şekil 3.22 : Semantik segmentasyon için tasarlanan KP-FCNN modeli Thomas ve diğ. (2019)'dan uyarlanmıştır.

KPConv çalışmasında diğer 3B nokta evrişimsel mimari çalışmalarında olduğu gibi, bahsedilen geometrik ilişkilerden faydalanabilmek için yerel komşuluk alanlarını işlem birimi olarak kabul eden bir yaklaşım kabul edilmiştir. Nokta komşuluk alanları oluşturulurken merkez olarak grid altörnekleme algoritması sonucu tespit edilen noktalar kullanılmıştır. Merkez ile ilişkilendirilecek komşu noktalar belirlenirken, uzunluğu katman derinliğine göre bir hiperparametre ile kontrol edilen küreler kullanılmıştır. Küre içine rastlayan noktalar merkez nokta ile birlikte komşuluk alanını tanımlamıştır. Nokta bulutları algılama geometrisine bağlı olarak değişken yoğunlukta olabildiklerinden 'k' en yakın komşuluk algoritması (kNN) yerine belirli boyuttaki küresel hacimlerden faydalanılmıştır. Değişken nokta yoğunluğu sebebi ile kNN algoritmasının uzak noktaları bir arada gruplayan kümeler oluşturabilmesi bu tercihe gerekçe olarak sunulmuştur.

3B nokta bulutlarının semantik segmentasyonu için tasarlanan KP-FCNN modeli şekil 3.22'de görsel olarak sunulmuştur. Model tasarlanırken semantik segmentasyon görevleri için popüler olarak kullanılan U-Net (Ronneberger ve diğ, 2015) mimarisi referans alınmıştır. Bu mimari kodlayıcı(encoder) ve kod çözücü(decoder) olmak üzere iki bölümden oluşmaktadır. Kodlayıcı bölümde nokta komşuluk alanlarını düşük seviyeden yüksek seviye soyutluğa sahip özellik vektörleri ile ifade edebilen temsillerin oluşturulması hedeflenmiştir. Kod çözücü bölümde alt örnekleme süresince

seyrekleştirilen nokta bulutunun orjinal mekansal boyutlarına ve özelliklerine geri dönüştürülmesi için en yakın komşuluk üst örnekleme ve tek katmanlı algılayıcılar ile uygulanan 1x1 evrişim işleminden faydalanılmıştır. Segmentasyon niteliğini arttırmak için U-Net mimarisinde olduğu gibi kodlayıcı bölüm katmanlarında elde edilen özellik vektörleri ara bağlantılarla eşlenik mekansal boyutlara sahip kod çözücü bölüm özellik vektörlerine birleştirilmiştir(concatenation). Son kod çözücü katman sonrasında nokta bazında softmax fonksiyonu uygulanmış ve her nokta için sınıf skorları hesaplanmış. En yüksek skora sahip sınıf ile ilişkilendirme yapılmıştır.

KP-FCNN mimarisinin 5 kodlayıcı katman ile uyarlanmış örneği 14.9 milyon eğitilebilir parametre içermektedir. Araştırmacılar bu modeli eğitmek için hesaplama donanımı olarak RTX2080Ti (11GB VRAM) ve Titan Xp (12GB VRAM) ekran kartlarını kullandıklarını raporlamışlardır. Modelin eğitilebilir bir örneğini yeniden oluşturmak için gerekli kaynak kodları araştırmacılar tarafından açık kaynaklı olarak Url-4 adresinden erişime açık biçimde sunulmuştur.

KPConv çalışmasında derin öğrenme görevleri için nokta bulutlarının doğrudan girdi olarak işlenmesi ile son-teknoloji olarak tanımlanabilecek nitelikte sonuçlar elde edilebileceği gösterilmiştir. İteratif, hesap yoğun ve sofistike nokta yeniden örnekleme algoritmaları (örneğin Furthest Point Sampling) yerine hesaplama yönünden daha hafif ve tekdüze olan grid altörnekleme yöntemi kullanıldığında da modellerin iyi performans gösterebildiğine yönelik bulgu elde edilmiştir. Bu durum daha yeni ve benzer performans değerlerine ulaşabilen RandLA-Net çalışmasında rastgele altörnekleme ile iyi başarı elde edilmesine paralel bir durum ortaya koymaktadır. Komşuluk alanı merkezlerinin oluşturulmasında hesap yükü hafif ve daha tekdüze algoritmaların kullanılmasının faydalı olacağı düşünülmektedir. Nokta bulutlarının, üretim yöntemleri ve algılama geometrisine bağlı olarak farklı doğrultular boyunca değişken yoğunluk gösterebilmeleri sebebi ile komşuluk alanları belirlenirken 'k' en yakın komşuluk algoritması yerine sabit yarıçaplı küreler tanımlanmasının faydalı olacağı öne sürülmüştür. Komşuluk alanlarındaki mekansal ilişkileri ifade eden özellik vektörlerinin elde edilebileceği bir evrişim işlemi geliştirilmiştir. Model boyunca ilerleyen katmanlarda gitgide daha az noktayla daha büyük komşuluk alanları kurulmuş ve bir temsiller hiyerarşisi elde edilmiştir. KPConv çalışması nokta bulutlarının derin öğrenme ile modellenmesi alanının gelişiminde önemli bir kilometre taşı olarak ortaya çıkmıştır.

4. 3B NOKTA BULUTLARININ SEMANTİK SEGMENTASYONU UYGULAMASI

Bu tez çalışmasında uygulama olarak 3B nokta bulutlarının semantik segmentasyonu için geliştirilen modellerin pratik olarak kullanılabilirliğini deneyimlemek amacı ile son teknoloji olarak tanımlanabilecek nitelikte bir modelin eğitilmesi ve test edilmesi hedeflenmiştir. Uygulama için seçilen modelin eğitim ve test süreçleri için yazılım, donanım ve veri gereksinimlerine karşılık çözümler geliştirilmiştir.

Eğitim ve test uygulaması için çalışmamızın teması olan kentsel alanların modellenmesine yönelik veri setlerinde başarıları kanıtlanmış olan ve KPConv çalışmasında semantik segmentasyon amacıyla geliştirilen KP-FCNN mimarisi tercih edilmiştir.

Eğitim veri seti olarak DALES veri seti; nokta sayısı, kapsadığı alanın büyüklüğü, tanımladığı semantik sınıflar ile kentsel modelleme ihtiyaçlarına uygun olması ve hava lidarı ile üretildiğinden gerçek ölçme koşullarını yansıtmaları sebepleri ile tercih edilmiştir.

KP-FCNN modelini eğitmek ve test etmek amacıyla, bu modelin aslına uygun bir uyarlamasını hazır olarak içeren açık kaynaklı kod depoları taranmıştır. Bulunan aday depolardan kullanıma hazır hale getirilmesi en kolay görülen ve Url-5 adresinde yer alan depo, kod kaynağı olarak seçilmiştir. Bulunan diğer aday kaynak kod imkanları; Torch-Points-3D (Chaton ve diğ. 2020) ve Open3D-ML (Zhou ve diğ. 2018) olarak gerçekleşmiştir. URL-5 adresinde yer alan depo katkı sağlayıcıları arasında orjinal KPConv çalışmasının yazarları bulunması, kullanılan kodların KPConv için resmi kaynak kodu deposundan uyarlanmış olması, geniş kullanım imkanları tanıyan MIT lisansı altında sunulması ve DALES veri setini KP-FCNN mimarisi ile işlemek için gerekli ara programların büyük oranda hazır halde sunulması sebepleri ile seçilmiştir.

Veri ve yazılım ihtiyaçlarına temel düzeyde çözümler getirilmesinin ardından donanım ihtiyaçları üzerine eğilinmiştir. KPConv çalışmasında eğitim ve test süreçlerinin 11GB VRAM bellekli RTX2080Ti ve 12GB VRAM bellekli Titan Xp ekran kartlarını içeren

bilgisayarlarda yürütüldüğü not edilmiştir. VRAM ekran kartı donanımının tek önemli özelliği olmasa da eğitim ve test esnasında modelin ve girdilerin sığabileceği büyüklükte belleğe sahip ekran kartına ihtiyaç duyulduğundan donanım gereksinimleri için belirleyici bir ölçüt olmuştur. Bahsedilen nitelikte bir ekran kartı donanımını doğrudan içeren bir iş istasyonu bilgisayar kurmaya çalışmak yerine kullandığın kadar öde sistemi ile çalışan bulut çözümler sabit maliyetlerden kaçınmak ve bulut çözümleri deneyimleyebilmek amacı ile araştırılmıştır.

En popüler bulut bilişim hizmet sağlayıcıları olan Google Cloud Platform (GCP), Google Colab, Amazon Web Services (AWS) ve Microsoft Azure arasından AWS; esnekliği, hizmetlerinin genel kapsamı, ekosistemini öğrenmenin kolay olması, geniş ve detaylı dokümantasyonu sebepleri ile tercih edilmiştir.

4.1 KP-FCNN Modelinin Bulut Ortamında Eğitilmesi

Semantik segmentasyon görevinde kullanılmak üzere KP-FCNN ağıнын eğitilmesi uygulaması için karar verilen bulut bilişim çözümünü gerçekleştirmek üzere kullanılacak iş akışı, bu alanda ileriki çalışmalara örnek sağlamak ve bulut bilişim çözümlerinin mevcut imkanlarını kullanıcıya sunmak amacıyla detaylı biçimde tasarlanmıştır.

4.1.1 Model eğitimi için bulut geliştirme ortamının düzenlemesi

Bulut hizmet sağlayıcısı olarak tercih edilen AWS tarafından makine öğrenmesi iş süreçleri için sunulan ürünler araştırılmıştır. Yapılan inceleme sonucunda iki potansiyel ürün tespit edilmiştir. İlk bulunan çözüm adayı Amazon tarafından modern makine öğrenmesine uygun donanım ve açık kaynaklı kod kütüphaneleri ile desteklenmiş biçimde sunulan SageMaker ürünü olmuştur. Amazon SageMaker adı altında makine öğrenmesi iş süreçleri için birden fazla hizmet sunmaktadır. Bunlar arasında veri etiketleme ortamı hizmeti, erişimi kolaylaştırmak için bulut depolama hizmetlerine arayüzler, Amazon'un kendi makine öğrenmesi modellerinden yararlanma imkanı örnek olarak verilebilir.

Bulunan diğer ürün ise bir hizmet olarak altyapı (IaaS) çözümü olan Elastic Compute Cloud (EC2) olmuştur. EC2 farklı amaçlara uygun biçimde tasarlanmış merkezi işlem birimleri, bellek çeşitleri ve ek olarak ekran kartları içerebilen makine örneklerinden oluşan bir kullandığın-kadar öde seçenek havuzu sunmaktadır. Makine örnekleri ya da

diğer adıyla sanal bilgisayarlar donanımlarına göre belirlenen saatlik ücretler ile kullanılmaktadır. Kullanıcılar seçtikleri makine örneklerinde yararlanmak istedikleri işletim sistemi ve yazılım kütüphaneleri kendilerine bilgisayar açıldığında hazır biçimde ulaştırılması için Amazon Machine Image adındaki makine görüntüsü kataloglarından seçebilmekte ya da kendi makine görüntülerini oluşturabilmektedirler. EC2 ürünleri yüksek derecede konfigürasyon esnekliği sunmaktadırlar fakat bu durum aynı zamanda kullanıcılara veri işlem süreçlerini gerçekleştirmek için daha çok sorumluluk yüklemektedir.

Temel olarak Amazon bir EC2 makine örneği için kullanıcı tarafından seçilen makine görüntüsünü yüklemekte, gerekli donanıma sahip bir sanal bilgisayarı internet üzerinden açık bir alan adı ile sunmakta ve kullanım süresince makine örneğinin saatlik ücreti karşılığında hem donanım hem de ağ ihtiyaçlarını karşılamaktadır. Bu sayede kullanıcılar ağ üzerinden bağlanabilecekleri yüksek düzeyde özelleştirilebilir bir sanal bilgisayara erişim imkanı bulmaktadır.

SageMaker ve EC2 makine örnekleri saatlik fiyatları karşılaştırıldığında EC2 makine örnekleri lehine bir maliyet durumu ortaya çıktığı görülmüştür. Çizelge 4.1'de bu karşılaştırmaya referans olan fiyatlar gösterilmiştir. Bulut hizmet sağlayıcıları, sunucularını farklı coğrafi merkezlere dağıtık halde konumlandırırlar. Fiyatların, AWS servis merkezlerinden eu-central-1/Frankfurt için çalışmanın yapıldığı tarihlerde geçerli olduğuna ve zaman içinde değişiklik gösterebileceğine dikkat edilmelidir.

Kullanılacak bulut bilişim ürünün seçilmesi için, makine öğrenmesi iş sürecini oluşturacağımız kaynak kodunun bağımlılıkları ve donanım gereksinimleri incelenmiştir.

Tercih edilen KPCConv uyarlamasını içeren açık kaynaklı kod deposunda yazılımsal bağımlılıklar; işletim sistemi için Ubuntu 16.04, makine öğrenmesi kütüphanesi olarak Python programlama dilinde yazılmış Tensorflow 1.12.0 veya 1.14.0 olarak belirtilmiştir.

Makine öğrenmesi kütüphanesi için ekran kartı donanımının hesaplama gücünden yararlanmak amacıyla ekran kartı programlama kütüphanelerinden CUDA 9.0 ve cuDNN 7.3.1 belirtilen Tensorflow versiyonları ile uyumlu olduklarından tercih edilmiştir. Tensorflow kütüphanesi için gerekli minimum Python dili ve yorumlayıcı versiyonu 3.7.0 olarak tercih edilmiştir.

Çizelge 4.1 : SageMaker ve EC2 tarafından derin öğrenme için sunulan makine fiyatları, Url-6'dan alınmıştır.

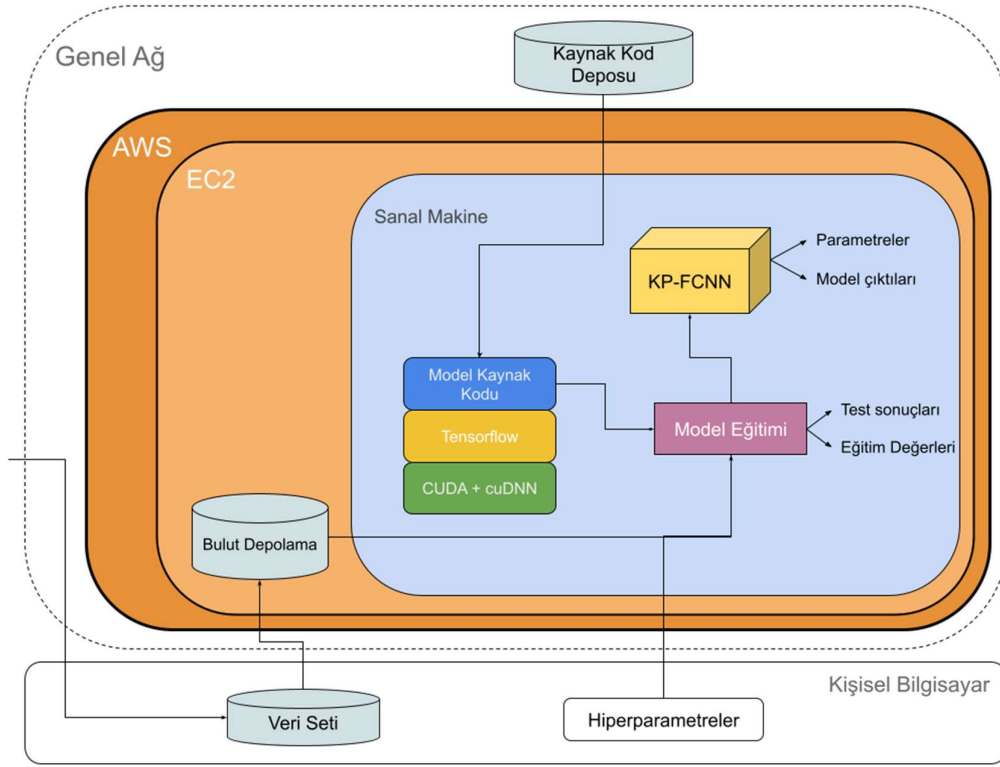
SageMaker		Elastic Compute Cloud	
Makine Örneği	Saatlik Fiyat (USD)	Makine Örneği	Saatlik Fiyat (USD)
ml.p3.2xlarge	4,779	p3.2xlarge	3,823
ml.p3.8xlarge	18,350	p3.8xlarge	15,292
ml.p3.16xlarge	35,172	p3.16xlarge	30,584
ml.p2.xlarge	1,658	p2.xlarge	1,326
ml.p2.8xlarge	12,730	p2.8xlarge	10,608
ml.p2.16xlarge	24,398	p2.16xlarge	21,216
ml.g4dn.xlarge	0,921	g4dn.xlarge	0,658
ml.g4dn.2xlarge	1,316	g4dn.2xlarge	0,940
ml.g4dn.4xlarge	2,107	g4dn.4xlarge	1,505
ml.g4dn.8xlarge	3,808	g4dn.8xlarge	2,720
ml.g4dn.12xlarge	6,846	g4dn.12xlarge	4,890
ml.g4dn.16xlarge	7,616	g4dn.16xlarge	5,440

3B nokta bulutlarından derin öğrenme nispeten yeni bir çalışma alanı olduğu ve nokta bulutu verisinin karakteristik özellikleri nedeniyle standart makine öğrenmesi kütüphaneleri mevcut halleri ile yetersiz kalmaktadırlar. Özellikle yerel komşuluk alanlarının oluşturulması, 3B nokta evrişim işlemi ve doğruluk ölçütlerinin hesaplanması için gerekli kodlar araştırmacılar tarafından bu kütüphanelere uyumlu olacak biçimde sıfırdan yazılmakta ve ayrı biçimde dağıtılmaktadır. Kullanılmaya karar verilen açık kaynaklı kod deposu da komşuluk alanları ile ilgili işlemler ve evrişim filtresi için C++ dilinde Tensorflow uyumlu olarak CUDA ve cuDNN kütüphaneleri bileşenleri kullanılarak yazılmış program ve program parçaları içermektedir. Nvidia tarafından geliştirilmiş ve açık kaynaklı olarak sunulan CUDA genel kullanım amaçlı ekran kartı programlama kütüphanesidir. cuDNN ise derin yapay sinir ağlarında kullanılan matematiksel işlemlerin ekran kartı üzerinde yürütülebilmesini sağlamak için geliştirilmiş bir kütüphanedir. Tensorflow ise Google tarafından CUDA ve cuDNN imkanları kullanılarak Python dilinde açık kaynaklı olarak geliştirilmiş makine öğrenmesi kütüphanesidir.

Gerekli bulut donanımını seçmek için KPConv çalışmasında raporlanmış donanım ve seçilen kaynak kod deposunda ima edilen yazılım bağımlılıkları incelenmiştir. Uygulanması planlanan KPConv uyarlamasının daha interaktif veri işlem süreçleri için tercih edilen “Jupyter Notebook” ortamında kullanıma uygun olmadığı düşünüldüğünden SageMaker yerine EC2 çözümleri saatlik maliyetler ve gerekli geliştirme ortamının kurulması işlemi için sarf edilecek emek göz önüne alınarak tercih edilmiştir.

CUDA ve cuDNN kütüphanelerinden faydalanabilmek için Nvidia firması tarafından bu tür işler için özel olarak üretilen ekran kartlarına gereksinim duyulmuştur. EC2 makine örnekleri arasından orjinal çalışmada kullanılan ekran kartlarına yakın performanslı adaylar içerenler tespit edilmiştir. Genel olarak “g4dn” sınıfından makine örnekleri 16GB VRAM’e sahip T4 model ekran kartları içerdikleri için tercih edilmiştir. Veri işlem ön adımları için 8 sanal çekirdekli ve 32GB rastgele erişimli hafıza (RAM) içeren “g4dn.2xlarge” makine örneği (instance) seçilmiştir. Seçilen makine örneğinde kullanılmak üzere makine öğrenmesi iş süreçlerine uygun biçimde ön ayarlı olarak sunulan makine görüntüleri incelenmiştir.

CUDA 9.0 ve cuDNN 7.3.1 bağımlılıklarını bir arada karşılayabilmek için Ubuntu 16.04 işletim sistemini içeren makine görüntüsü tercih edilmiştir. Yazılım bağımlılıklarını karşılamak için seçilen makine görüntüsü ile çalıştırılan “g4dn.2xlarge” makine örneğine SSH (Secure Shell) protokolü ile bağlanılmış ve ilk olarak CUDA 9.0 ve cuDNN 7.3.1 kütüphaneleri resmi kaynaklarından erişilerek kurulmuştur. Tensorflow kütüphanesini kullanmak için Python programlama dilinin 3.7.0 versiyonu kurulmuştur. Sonrasında Tensorflow kütüphanesinin 1.14.0 versiyonunun ekran kartı destekli sürümü Python paket yönetim sistemi ile indirilmiştir. Seçilen açık kaynaklı kod deposu “Git” adındaki versiyon kontrol sistemi ile “Github” web sitesinden URL’si ile (Uniform Resource Locator) kullanılarak indirilmiştir. KPConv modelinin çalıştırılması için gereken özel işlemler kaynak kodlarından derlenmiştir ve gerekli geliştirme ortamı hazır hale getirilmiştir. Eğitim için kullanılacak DALES veri setine ancak kişisel e-mail yoluyla ulaşılabildiğinden PLY dosya formatında saklanan nokta bulutlarından oluşan sıkıştırılmış veri arşivi bulut ortamına SFTP (Secure File Transfer Protocol) yöntemiyle aktarılmıştır. Şekil 4.1’de oluşturulan geliştirme ortamı bileşenleri birbirleri ile olan ilişkileri ile görselleştirilmiştir.



Şekil 4.1 : KP-FCNN modelinin bulut ortamında eğitilmesi için kurulan yazılım geliştirme ortamı bileşenlerinin ilişkisini gösteren şema.

4.1.2 Model eğitimi

Derin öğrenme modellerinin eğitiminde model mimarisi kadar, kullanılan bileşenlere dair hiperparametrelerin uygun biçimde seçilmesinin önemli etkiye sahip olduğu görülmüştür. Çalışma kapsamında bulut bilişim kaynaklarını kullandığından maliyet kısıtları gereği gereksinim duyulan hiperparemetreler; başarılı sonuç ürettiği raporlanmış değerler araştırılarak seçilmiştir.

KP-FCNN modelinin eğitimine esas olacak hiperparametrelerin seçilmesi için orjinal KP-Conv çalışması, resmi kaynak kodu deposu ve kullanmaya karar verdiğimiz kaynak kod deposu incelenmiştir. Yapılan inceleme sonucunda 5 kodlayıcı ve kod çözücü katmandan oluşan KP-FCNN mimarisinin kullanılmasına karar verilmiştir. Optimizasyon algoritması olarak momentum meyl azalımı yöntemi 0.98 sürtünme katsayısı (gradyenler küçülmeye başladığında yani olası bir maksimum noktasına yaklaşıldığında yavaşlatıcı etki yapar) ile seçilmiştir. Momentum temelli optimizasyon algoritmaları ağ ağırlıkları için güncellemeleri meyl azalımlı öğrenme sürecinde ortaya çıkan gradyenlerin büyüklüklerini dikkate alarak

hesaplayabilmektedirler. Başlangıç öğrenme katsayısı olarak 0.01 seçilmiştir. Ağ boyunca özellik vektörlerini normalize halde tutmak için KPCnv katmanları sonrasında batch-normalizasyon katmanları kullanılmıştır. Batch-norm katmanlarında hareketli ortalamaların hesaplanması için gerekli momentum katsayısı 0.98 olarak belirlenmiştir.

Son katmanda nokta bazında sınıflandırma sonuçlarını elde etmek için 8 sınıflı softmax katmanı kullanılmıştır. Hata fonksiyonu olarak softmax sınıflandırıcı ile uyumlu çapraz-entropi fonksiyonu kullanılmıştır.

KPCnv katmanları için filtre elemanı olarak kullanılacak nokta sayısı 15 olarak belirlenmiştir. İlk katmanda uygulanacak grid yeniden örnekleme işleminde birim hücre boyutu 25x25x25cm olarak ayarlanmıştır. 'dl0' olarak ifade edilen bu değer sonraki katmanlarda oluşturulacak komşuluk alanlarının boyutlarına etki ettiğinden önemli bir hiperparametredir. Devam eden katmanlarda grid hücre boyutları, bir önceki katmana göre her ekseninde 2 katına çıkartılarak elde edilmiştir.

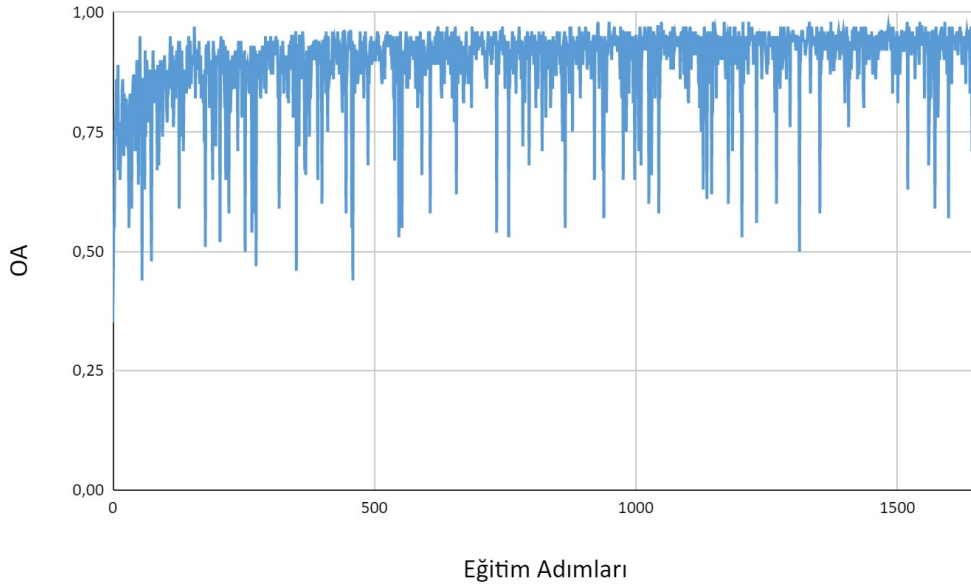
İlk katman komşuluk alanlarının yarıçapları 20 metre olarak belirlenmiştir. Komşuluk alanı yarıçapı da her katmanda grid hücresi boyutlarındaki değişim oranında artırılmıştır.

Eğitim için kullanılan DALES veri seti nokta koordinatları, nokta bulutu dosyası bazında en sol alt nokta (0,0,0) pozisyonunda konumlandırılacak şekilde kaydırılmıştır.

Veri çoğaltma yöntemi olarak girdi nokta bulutlarının X-ekseni yönünde simetrisi alınmış, düşey eksen etrafında rastlantısal dönüklüğe tabi tutulmuş, nokta koordinatlarına ortalaması sıfır ve varyansı 0.01 olan gauss filtresi ile gürültü eklenmiş, koordinatlara 3 koordinat ekseninde anizotropik biçimde 0.9 ve 1.1 değerleri arasında rastlantısal ölçekleme uygulanmıştır.

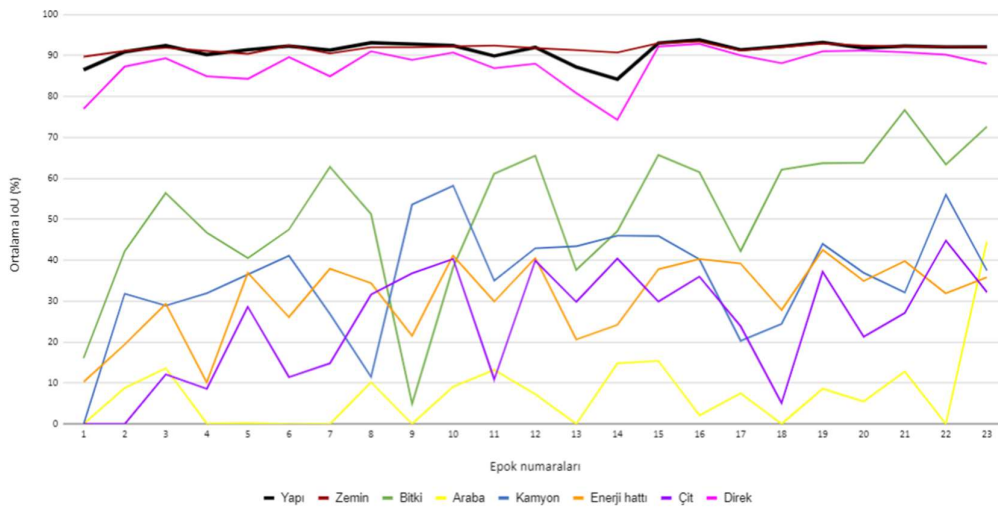
Eğitim esnasında her beş epokta bir kere model kayıt edilmiştir. Her epok, 4 nokta komşuluk alanı içeren 551 adımdan meydana gelmiştir. Eğitim boyunca her epok sonunda 50 komşuluk alanından oluşan validasyon seti modele beslenmiş sınıf bazında ve sınıflar ortalaması olmak üzere 2 IoU değerleri raporlanmıştır.

Eğitim boyunca bir epoğun yaklaşık 2 saatte tamamlandığı görülmüştür. Maliyet kısıtları sebebi ile eğitim 23. epokta sona erdirilmiştir. Eğitilen modele dair eğitim boyunca raporlanan sonuçlar Şekil 4.2, 4.3 ve 4.4'te görselleştirilmiştir.

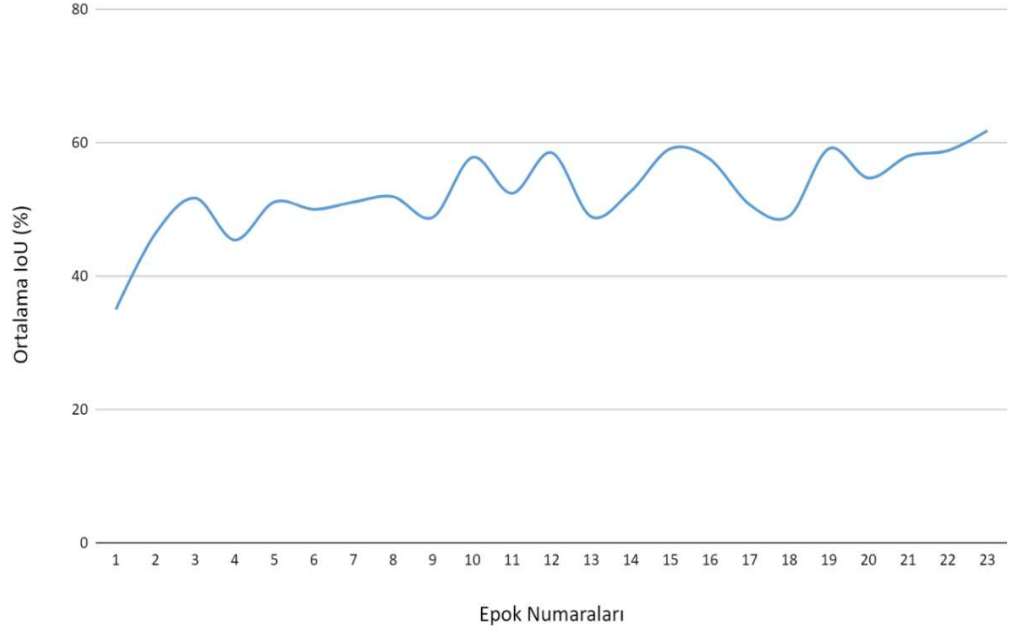


Şekil 4.2 : 3 Epok boyunca eğitim seti üzerinde hesaplanan OA değerleri.

Sunulan sonuçlar incelendiğinde doğru semantik sınıfa atanan nokta sayısının tüm noktalara oranı olarak hesaplanan OA değerinin eğitimin en başında doyum noktasına ulaşp eğitim boyunca bu değerlerde seyir etmesine rağmen, epok sonlarında hesaplanan IoU değerlerinin yaşadığı değişkenlik modeller değerlendirilirken göreve uygun metriklerin seçilmesi gerekliliğini göstermiştir. OA değerinin doyum noktasına hızlıca ulaşmasına sebep olarak veri setindeki nokta sınıfları arasındaki dengesizlik olduğu düşünülmüştür.



Şekil 4.3 : Eğitim boyunca validasyon seti üzerinde hesaplanan ortalama IoU değerlerinin değişimi.



Şekil 4.4 : Eğitim boyunca validasyon seti üzerinde hesaplanan sınıflar ortalaması olarak hesaplanan IoU değeri.

4.1.3 Eğitim sonuçları ve modelin test edilmesi

23 epok boyunca eğitilen model parametreleri ile birlikte kayıt edilerek test verisi üzerinde denenmiştir. Elde edilen test sonuçları, bir diğer çalışmada raporlanan sonuçlar ile karşılaştırılmış ve yorumlanmıştır.

Çizelge 4.2 : Test seti üzerinde elde edilen sonuçlar.

Eğitim	Sınıflar Ortalaması IoU (%)	Sınıf Bazında Ortalama IoU (%)							
		Zemin	Bitki	Araba	Kamyon	Enerji Hattı	Çit	Direk	Yapı
Çalışma kapsamında gerçekleştirilen eğitim.	61.6	95.7	91.0	6.6	1.6	65.6	50.4	29.3	92.9
Varney ve diğ. (2020)'da raporlanan sonuçlar.	81.1	97.1	94.1	85.3	41.9	95.5	63.5	75	96,6

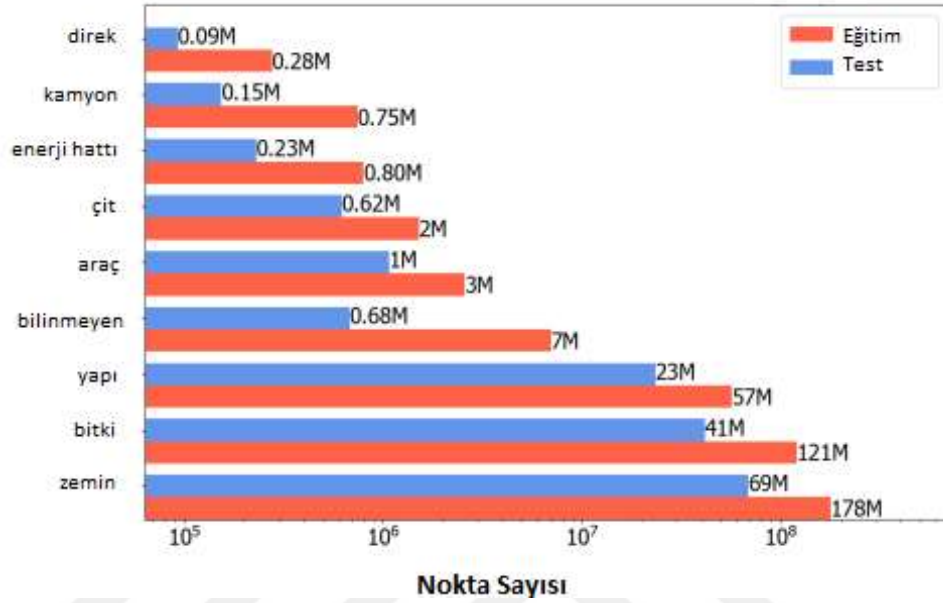
Eğitim sürecine ve modelin test verisi üzerinde denenmesine ilişkin sonuçlar incelendiğinde eğitimin ağ henüz optimum değerlere yakınsamadan erken biçimde 23. epokta sonlandırılmasına rağmen veri setinde en büyük ağırlıkla temsil edilen zemin, yapı ve bitki sınıflarında iyi sonuçlar elde edilebildiği görülmüştür. Bu duruma karşılık daha az sayıda örnek noktayla temsil edilen araba, enerji hattı, çit ve direk sınıfları için daha uzun süreli bir eğitim sürecinin yürütülmesi, bu mümkün değilse de hata fonksiyonun sınıf ağırlıklarını göz önüne alacak biçimde düzenlenmesi gerektiği sonucuna varılmıştır.

Bütün veri setinde en az nokta ile temsil edilen kamyon sınıfına ait nesnelerin tespitinde modelin başarız olduğu görülmüştür. Çizelge 4.2’de sunulduğu üzere çalışma kapsamında elde edilen sonuçlar model tümüyle yakınsayana kadar yürütülmüş bir eğitim sonrasında elde edilen sonuçlar karşılaştırıldığında zemin, yapı ve bitki sınıflarında benzer performans elde edildiği görülmüştür. Diğer sınıflarda ise %45’e varan ortalama IoU farkları, modelin bu sınıfları tespit etmeyi daha iyi öğrenmesi için daha uzun süre eğitilmesi gerektiğini teyit etmiş ve bu sınıfların, segmentasyona esas temsillerini yalnızca 3B koordinatlar ve yansıma değerlerinden elde etmenin daha zor olduğunu düşündürmüştür.

Eğitilen model tarafından test verisi üzerinde elde edilen semantik segmentasyon sonuçları Şekil 4.7’de sol tarafta gerçek durum, sağ tarafta model tarafından üretilen sonuçlar alt alta yerleştirilerek sunulmuştur. Özellikle yapı-çit, yapı-kamyon, kamyon-araç, araç-bitki, bitki-çit sınıf çiftleri arasında modelin bulanık sonuçlar ürettiği görülmüştür. DALES veri setinde çit sınıfına bahçe ve istinat duvarları gibi detay noktalarının da dahil edilmiş olması yapı cephelerine benzer düşey yüzeyler içeren bu sınıf için yapılarla olan bulanık sonuçları açıklayabileceği düşünülmüştür.

Araç ve kamyonlar ile kamyonlar ve yapılar arasındaki geometrik benzerlik düşünüldüğünde modelin bu sınıfları birbirine karıştırması normal kabul edilebilir. Ayrıca şekil 4.5’te görülebileceği üzere araba ve özellikle de yapı sınıfının veri setinde kamyon sınıfına göre daha büyük yoğunlukla temsil edilmesinin modelin bu sınıflar lehine sonuç üretme yatkınlığını açıklayabileceği düşünülmüştür. Model eğitimi

enasasında nokta bazında sınıf puanlarını hesaplayan softmax katmanına ‘bilinmeyen’ sınıfı dahil edilmemiştir.



Şekil 4.5 : DALES veri setindeki sınıflara ait nokta sayılarının dağılımı.

Çizelge 4.3 : Test seti üzerinden elde edilen doğruluk ölçütleri.

	Zemin	Bitki	Araç	Kamyon	Enerji Hattı	Çit	Direk	Yapı
Kesinlik	0,96	0,96	0,93	0,54	0,91	0,85	0,88	0,97
Duyarlılık	0,99	0,95	0,71	0,02	0,95	0,58	0,35	0,96
F1 Skoru	0,98	0,96	0,81	0,03	0,93	0,69	0,50	0,96

Eğitilen modelin test verisi üzerindeki kesinlik ve duyarlılık değerlerinin sınıf bazında incelenmesi için Çizelge 4.3 düzenlenmiştir. Zemin, bitki, enerji hattı ve yapı kavramsal sınıflarına ait örneklerde modelin başarılı kabul edilebilecek ve dengeli kesinlik ile duyarlılık değerleri üretebildiği görülmüştür. Modelin; veri setinde daha az nokta ile temsil edilen enerji nakil hattı kablolarına ait ölçme noktalarından oluşan enerji hattı sınıfındaki örnekleri tespit edebilme ve doğru sınıflandırma kapasitesi, benzer ağırlıkta temsil edilen sınıflar ile karşılaştırıldığında daha başarılı bulmuştur. Veri setinde daha büyük ağırlıkla temsil edilen sınıflarda duyarlılık açısından modelin daha başarılı sonuçlar ürettiği tespit edilmiştir. Tüm sonuçlar incelendiğinde kamyon

kavramsal sınıfı haricinde diğer sınıflar için modelin kesinlik değerlerinin düşük duyarlılığa sahip sınıflar da dahi daha yüksek gerçekleştiği görülmüştür.

Modelin test edilmesi sonucunda oluşturulan hata matrisi (confusion matrix) Şekil 4.6'da sunulmuştur. Hata matrisi modelin ürettiği sınıflandırma sonuçlarını kullanıcıya asimetrik biçimde inceleme imkanı sunduğu için faydalı bir gösterge olmaktadır. Hata matrisinin diagonal elemanları modelin doğru sınıflandırdığı örneklerin sayısını veya oranını belirtmektedir. Diagonal elemanların alt ve üstündeki değerler model tarafından yapılan yanlış sınıflandırmanın (FP), sağ ve solundaki değerler modelin ilgili örnekleri doğru sınıf ile ilişkilendirememesinin (FN) birer ölçütüdür.

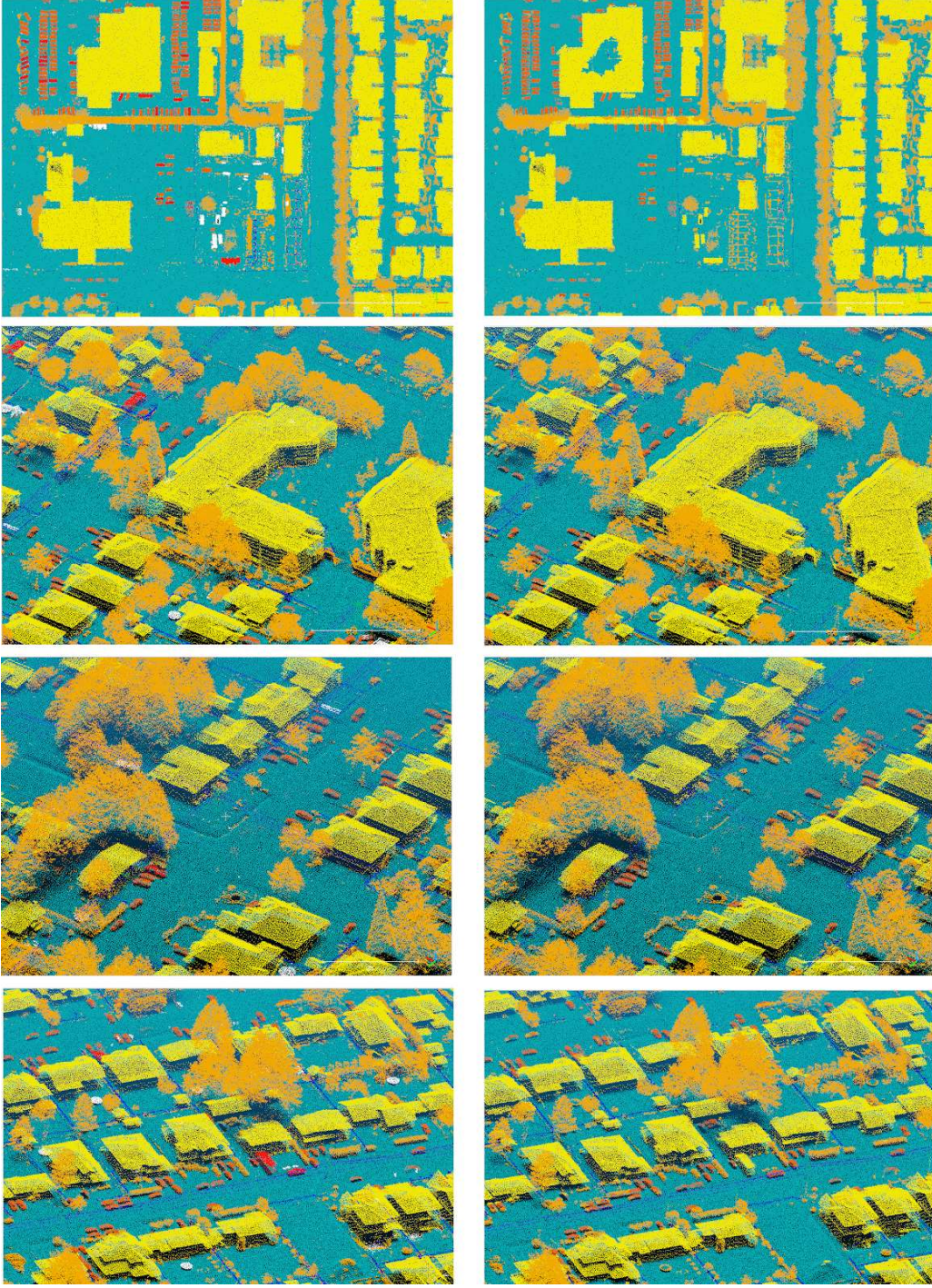
Hata matrisi girdileri kolay okunabilirlik sağlamak amacıyla en sağ sütunda sınıfa ait nokta sayıları olarak belirtilen değerlere bölünerek normalize edilmiştir. Model tarafından bir sınıfa atanan örnekler arasında FP sonuçların oranını ifade eden 'Error of Comission' (EoC) ve FN sonuçların gerçekte sınıfa ait örneklerin sayısına oranını gösteren 'Error of Omission' (EoO) değerleri hesaplanmıştır. Bu bağlamda EoC yanlış sınıf ile ilişkilendirmenin, EoO doğru sınıf ile ilişkilendirememenin birer ölçütü olmaktadır.

Hata matrisi incelendiğinde gerçekte 'bilinmeyen' sınıfına ait noktaların çoğunlukla bitki ve zemin sınıflarına atandığı görülmüştür. Softmax katmanında 'bilinmeyen' sınıfına yer verilmediğinden bu sınıfa dair EoC ve EoO değerleri hesaplanmamıştır. Kamyon sınıfına ait örneklerin büyük oranda yapı sınıfı ile karıştırıldığı buna karşın yapı sınıfına ait örneklerin kamyon sınıfı ile çok daha az oranda karıştırıldığı tespit edilmiştir.

Çit ve direk sınıfından örneklerin bitki sınıfı ile karıştırıldığı görülmüştür. Daha önce Çizelge 4.3 ve Çizelge 4.2'de sunulan sonuçlara paralel şekilde zemin, bitki, enerji hattı ve yapı sınıfına ait örnekler için modelin performansı diğer sınıflara oranla çok daha yüksek bulunmuştur. EoO değerleri incelendiğinde model tarafından tespit edilebilirliği en düşük sınıflar sırası ile kamyon, direk ve çit olarak gerçekleşmiştir. En düşük EoC değerine sahip sınıfların sırası ile yapı, zemin ve bitki olduğu görülmüştür. Bu veriler ışığında modelin yapısal hacimlerdeki değişimleri takip etmede kullanılabilecek bir 3B şehir modelinin üretilmesi işinde faydalı olabileceği düşünülmüştür.

		Model Tarafından Üretilen Sınıf Etiketleri									Sınıf Ait Nokta Sayısı
		Diğer /Arkaplan /Bilinmeyen	Zemin	Bitki	Araç	Kamyon	Enerji Hatları	Çit	Direk	Yapı	
Gerçek Sınıf Etiketleri	Diğer /Arkaplan /Bilinmeyen	0,00000	0,32772	0,49157	0,00238	0,00042	0,00523	0,01071	0,00181	0,16016	681571
	Zemin	0,00000	0,99224	0,00676	0,00000	0,00000	0,00000	0,00000	0,00000	0,00100	68871897
	Bitki	0,00000	0,03703	0,95011	0,00049	0,00002	0,00002	0,00114	0,00002	0,01118	41464228
	Araç	0,00000	0,01811	0,23248	0,70743	0,00120	0,00004	0,00020	0,00001	0,04052	1070554
	Kamyon	0,00000	0,04944	0,29934	0,19106	0,01761	0,00000	0,00080	0,00000	0,44175	154142
	Enerji Hatları	0,00000	0,00002	0,02957	0,00004	0,00000	0,94560	0,00077	0,00848	0,01551	230412
	Çit	0,00000	0,01189	0,32453	0,00126	0,00000	0,00260	0,57637	0,00066	0,08269	624069
	Direk	0,00000	0,00071	0,39247	0,00072	0,00001	0,15919	0,00358	0,34606	0,09726	92724
	Yapı	0,00000	0,03145	0,01178	0,00005	0,00000	0,00000	0,00032	0,00000	0,95639	23454294
Sınıfa Atanan Nokta Sayısı		0	70868493	41013281	810730	5041	238802	422845	36352	23248347	136643891
Error of Comission		-	0,036	0,039	0,066	0,461	0,088	0,149	0,117	0,035	
Error of Omission		-	0,008	0,050	0,293	0,982	0,054	0,424	0,654	0,044	

Şekil 4.6 : Modelin test edilmesi sonucunda üretilen hata matrisi.



Şekil 4.7 : Model tarafından üretilen sonuçlar. Sol taraf test verisinden alınan etiketlere, sağ taraf model tarafından üretilen sonuçlara göre renklendirilmiştir.

5. SONUÇLAR VE BULGULAR

Bu çalışmada son yıllarda ölçme teknolojilerinde yaşanan ilerlemelere bağlı olarak gittikçe yaygınlaşan 3B nokta bulutlarının derin öğrenme yöntemleri ile işlenmesi alanında yaşanan gelişmeler incelenmiş, modern ve gelişmiş veri işleme imkanlarını daha erişilebilir hale getiren bulut bilişim çözümlerinin bu çalışma alanında kullanılmasına örnek olabilecek bir eğitim süreci tasarlanıp uygulanmıştır.

3B nokta bulutlarının derin öğrenme yöntemleriyle modellenmesi alanında yaşanan gelişmeler, nokta bulutlarının bir veri tipi olarak özellikleri ve bu özelliklerinden doğan zorluklar ve fırsatlar göz önüne alınarak incelenmiştir. Nokta bulutları; ilgi objelere dair geometrileri yüksek çözünürlükle ve aslına uygun biçimde modellemeye imkan tanıyan bir veri kaynağı olmaları sebebiyle; araştırmacılar tarafından bu güçlü özellikten faydalanmayı sağlayacak derin öğrenme çözümleri geliştirilmesi hedeflenmiştir. Günümüzde otonom mobilite, robotik ve 3B modelleme gibi çalışma alanlarında 3B nokta bulutları ile temsil edilen nesneleri sınıflandırma, tespit etme ve segmente etme görevlerinde başarılı sonuçlar elde edebilen derin öğrenme modelleri geliştirilmiştir. Son teknoloji olarak değerlendirilebilecek modeller incelendiğinde genel olarak nokta bulutları üzerinde oluşturulan yerel komşuluk alanlarının ifade ettiği mekansal özelliklerden faydalanabilmek için doğrudan nokta bulutlarını girdi kabul eden yaklaşımların popülerleştiği görülmüştür.

Literatür taramasının ardından pratik uygulama imkanları araştırılmıştır. Genel olarak derin öğrenme çalışmalarına imkan tanıyan ve çalışmaları hızlandıran yazılım ve donanım ürünleri gözden geçirilmiştir. Alanda kullanılan yazılım ürünleri çoğunlukla açık kaynaklı olsalar da en yaygın kullanıcı kitlesine ve ürün yelpazesine sahip olanların yalnızca belirli üreticiler tarafından sağlanan ekran kartı donanımlarına yüksek oranda bağımlılık durumu altında olduğu görülmüştür. Özel olarak 3B nokta bulutları için geliştirilen derin öğrenme çözümlerinin henüz standart bir hale getirilip daha ana akım açık kaynaklı derin öğrenme yazılım kütüphaneleri tarafından sunulmaya uzak olduğu tespit edilmiştir. 3B nokta bulutlarının işlenmesi için geliştirilen derin öğrenme modellerinin özellikle örnekleme, komşuluk alanı oluşturma

ve evriřim iřlemleri iin standart zmler bulunmadıėından bu iřlevleri gerekleřtirmek iin zel yazılımlar retilmesi gerekliliėi fark edilmiřtir. Standarttan uzaklařtırıcı etki gsteren bu durumun karmařık yazılım ve donanım baėımlılıklarına sebep olarak pratik uygulamalar iin zorluk yarattıėı tespit edilmiřtir. Modelleri oluřturup eėitmek iin gerekli kaynak kodların genel olarak zgr lisanslar ile aık kaynaklı biimde sunulması hem yeni teknolojiye eriřimi demokratikleřtirmeye katkı saėlamakta hem de alıřmalarda raporlanan sonuların yeniden retilbilirliėi ve test edilebilirliėini arttırarak arařtırmalar iin hızlandırıcı etki gstermektedir. 3B nokta bulutlarının derin ėrenmesi alanında da bu tutumun yaygın olması olumlu bulunmuřtur.

Derin ėrenme srelerinin saėlıklı olarak yrtlebilmesi iin genel olarak yksek performanslı hesaplama donanımı gerektirmesi sonucunda hesaplama ve depolama nitelikleri bakımından geliřmiř ve esnek biimde konfigre edilebilir bilgisayarlara eriřim nemli bir ihtiya olarak ortaya ıkmıřtır. Derin ėrenme modellerini arařtırma, inceleme, prototip oluřturma, eėitme, test etme ve kullanıma sunmak iin ihtiya duyulan eřitli donanım ve yazılım konfigrasyonlarını elde etmede sabit maliyetleri ortadan kaldırıp kullandıėın kadar de mantıėı ile iřletilen bulut biliřim zmleri n plana ıkmıřtır. 3B nokta bulutlarından derin ėrenmede son teknoloji modellerin eėitimi iin bulut biliřim imkanları hizmet saėlayıcılar ve rn yelpazeleri bazında incelenmiřtir. gereksinimlere ve kaynaklara uygun maliyette ve nitelikte bulut hizmetleri seilmesine rnek bir pratik uygulama yapılmıřtır. 3B nokta bulutu iřlemede kullanılan derin ėrenme aėının optimum deėerlere yakınsama sresinin gnler ile ifade edilebilecek zaman kesitlerine ulařabildiėi grlmřtr.

Nokta bulutlarından derin ėrenmede kullanılacak veri setlerinin ierikleri, uyguladıkları sınıflandırma sistemi ve amaları bakımından uyumlu olması gerekliliėi n plana ıkmıřtır. řekilsel ya da geometrik zellikleri ile ayırt edilmesi zor sınıflar tanımlayan veri setleri, bařarılı bir makine ėrenmesi sreci iin noktalara dair radyometrik byklkler ya da bařka zellikler de iermelidir. rneėin SensatUrban veri setinde RGB deėerlerinin eėitime dahil edilmesi karayolu, kpr, otopark alanı gibi sınıfların segmentasyon bařarısını arttırmıřtır. Bu nedenle eėitim veri setleri tasarlanıp oluřturulurken ve tercih edilirken bahsedilen uyumu saėlayacak nlemler alınmasının nemli olduėu tespit edilmiřtir.

3B nokta bulutlarının semantik segmentasyonu görevi için geliştirilen derin öğrenme ağlarının genel olarak yapı, zemin ve ağaçlık alan sınıflarını ayırt etmede başarılı olduğu, son-teknoloji standartlarını belirleyen modellerin enerji nakil hattı gibi detay sınıflarına ait geometrileri dahi otomatik olarak elde etmeyi sağlayacak yetkinliğe ulaşmaya başladığı görülmüştür. Bu modellerin yapısal hacimlerin çevre ile ilişkisini konu alan mekansal analizlere altlık olarak kullanılacak şehir modellerinin geometrik bilgi ihtiyaçlarını karşılamada katkı sağlayabileceği görülmüştür.

Modern ölçme teknikleri sayesinde geniş alanlar için hızlıca toplanabilen 3B nokta bulutlarının veri işlem süreçlerinin otomasyonu ile emek, zaman ve maliyet tasarrufu sağlanacağı dahası 3B modelleme, robotik, otonom mobilite gibi daha yüksek katma değerli uygulamalara yönelik çalışmaları hızlandıracağı düşünülmektedir. Son yıllarda derin öğrenme alanında 3B nokta bulutlarının işlenmesi için başarılı çözümlerin geliştirildiği görülmüştür. Bu teknolojiden pratik anlamda yararlanmak için popüler makine öğrenmesi kütüphanelerinin mevcut kabiliyetleri henüz yetersiz gelmekte ek yazılım geliştirme süreçlerine gereksinim duyulmaktadır. Derin öğrenme modellerinin eğitiminde kullanılmak üzere sentetik ve gerçek verilerden oluşan eğitim setleri git gide çoğalmaktadır. Özellikle semantik segmentasyon görevi için kent ölçeğinde nokta bazında etiketlenmiş veri setleri günümüzde mevcut hale gelmiştir. Bu veri setleri genellikle kendi içlerinde benzer yerleşim dokusunda sahneleri içerdiklerinden eğitilen modellerin genelleştirilebilirliğine katkı sağlamak için daha çeşitli veri setlerine ihtiyaç devam etmektedir. Bulut bilişim alanında hizmet sağlayıcıları hem kendilerine has hem de benzer ürünler içeren yelpazelerde farklı çözümler sunmaktadırlar. Bunlar arasından ihtiyaçlara cevap verebilecek ve kaynaklara uygun ürünlerin seçilmesi önemli bir karar konusu olarak ortaya çıkmaktadır.



KAYNAKLAR

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., . . . Zhang, X.** (2016). TensorFlow: A system for large-scale machine learning. *ArXiv, abs/1605.08695*.
- Ahmed, E., Saint, A., Shabayek, A. E., Cherenkova, K., Das, R., Gusev, G., . . . Ottersten, B. E.** (2018). A survey on Deep Learning Advances on Different 3D Data Representations. *arXiv: Computer Vision and Pattern Recognition*.
- Al-Athel, S. A., Fadika, L. M., Láng, I., Shijun, M., de Botero, M. M., Singh, N., . . . Sokolov, V.** (1987). Report of the World Commission on Environment and Development: "Our Common Future".
- Baltsavias, E.** (1999). A comparison between photogrammetry and laser scanning. *Isprs Journal of Photogrammetry and Remote Sensing*, 54, 83-94.
- Behley, J., Garbade, M., Milioto, A., Quenzel, J., Behnke, S., Stachniss, C., ve Gall, J.** (2019). SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 9296-9306.
- Biljecki, F., Stoter, J. E., Ledoux, H., Zlatanova, S., ve Çöltekin, A.** (2015). Applications of 3D City Models: State of the Art Review. *ISPRS Int. J. Geo Inf.*, 4, 2842-2889.
- Billen, R., Cutting-Decelle, A.-F., Marina, O., de Almeida, J.-P., Cagliioni, M., Falquet, G., . . . Zlatanova, S.** (2014). 3D City Models and urban information: current issues and perspectives.
- Blender - a 3D modelling and rendering package.** (2018). Blender Foundation.
- Bottou, L. E.** (1998). Online Learning and Stochastic Approximations.
- Boulch, A.** (2019). Generalizing discrete convolutions for unstructured point clouds. *3DOR@Eurographics*.
- Boulch, A.** (2020). ConvPoint: Continuous convolutions for point cloud processing. *Comput. Graph.*, 88, 24-34.
- Buyuksalih, G., Baskaraca, P., Bayburt, S., Buyuksalih, I., ve Rahman, A. A.** (2019). 3D City Modeling of Istanbul Based On LIDAR Data and Panoramic Images - Issues and Challanges. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*.
- Chaton, T., Nicolas, C., Horache, S., ve Landrieu, L.** (2020). Torch-Points3D: A Modular Multi-Task Framework for Reproducible Deep Learning on 3D Point Clouds. *2020 International Conference on 3D Vision (3DV)*.
\url{https://github.com/nicolas-chaulet/torch-points3d} adresinden alındı

- Chen, X., Ma, H., Wan, J., Li, B., ve Xia, T.** (2017). Multi-view 3D Object Detection Network for Autonomous Driving. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 6526-6534.
- Chollet, F.** (2017). Deep Learning with Python.
- Chollet, F.** (2018). Keras: The Python Deep Learning library.
- Ciresan, D. C., Meier, U., Gambardella, L. M., ve Schmidhuber, J.** (2011). Convolutional Neural Network Committees for Handwritten Character Classification. *2011 International Conference on Document Analysis and Recognition*, 1135-1139.
- De Deuge, M., Quadros, A., Hung, C., ve Douillard, B.** (2013). Unsupervised feature learning for classification of outdoor 3d scans. *Australasian Conference on Robotics and Automation*, 2, s. 1.
- Duchi, J. C., Hazan, E., ve Singer, Y.** (2011). Adaptive Subgradient Methods for Online Learning and Stochastic Optimization. *J. Mach. Learn. Res.*
- Engelmann, F., Kontogianni, T., Hermans, A., ve Leibe, B.** (2017). Exploring Spatial Context for 3D Semantic Segmentation of Point Clouds. *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*, 716-724.
- Esteves, C., Allen-Blanchette, C., Makadia, A., ve Daniilidis, K.** (2018). Learning SO(3) Equivariant Representations with Spherical CNNs. *ECCV*.
- Etzioni, A., ve Etzioni, O.** (2017). Incorporating Ethics into Artificial Intelligence. *The Journal of Ethics*, 21, 403-418.
- Fukushima, K.** (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36, 193-202.
- Geiger, A., Lenz, P., ve Urtasun, R.** (2012). Are we ready for autonomous driving? The KITTI vision benchmark suite. *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 3354-3361.
- Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., Meena, S. R., Tiede, D., ve Aryal, J.** (2019). Evaluation of Different Machine Learning Methods and Deep-Learning Convolutional Neural Networks for Landslide Detection. *Remote. Sens.*, 11, 196.
- Glorot, X., Bordes, A., ve Bengio, Y.** (2011). Deep Sparse Rectifier Neural Networks. *AISTATS*.
- Goodfellow, I. J., Bengio, Y., ve Courville, A. C.** (2015). Deep Learning. *Nature*, 521, 436-444.
- Graham, B., Engelcke, M., ve van der Maaten, L.** (2018). 3D Semantic Segmentation with Submanifold Sparse Convolutional Networks. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 9224-9232.
- Grand Theft Auto - V.** (2014). *GTA-V*. New York, NY: Rockstar Games.
- Griffiths, D. J., ve Boehm, J.** (2019). SynthCity: A large scale synthetic point cloud. *ArXiv, abs/1907.04758*.

- Gschwandtner, M., Kwitt, R., Uhl, A., ve Pree, W.** (2011). BlenSor: Blender Sensor Simulation Toolbox. *ISVC*.
- Guo, Y., Wang, H., Hu, Q., Liu, H., Liu, L., ve Bennamoun, M.** (2021). Deep Learning for 3D Point Clouds: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43, 4338-4364.
- Hackel, T., Savinov, N., Ladicky, L., Wegner, J. D., Schindler, K., ve Pollefeys, M.** (2017). Semantic3D.net: A new Large-scale Point Cloud Classification Benchmark. *ArXiv, abs/1704.03847*.
- Han, Y., Qin, R., ve Huang, X.** (2020). Assessment of dense image matchers for digital surface model generation using airborne and spaceborne images@ an update. *Photogrammetric Record*, 35, 58-80.
- Hochreiter, S., ve Schmidhuber, J.** (1997). Long Short-Term Memory. *Neural Computation*, 9, 1735-1780.
- Hu, Q., Yang, B., Khalid, S., Xiao, W., Trigoni, A., ve Markham, A.** (2021). Towards Semantic Segmentation of Urban-Scale 3D Point Clouds: A Dataset, Benchmarks and Challenges. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 4975-4985.
- Hu, Q., Yang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., . . . Markham, A.** (2020). RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 11105-11114.
- Hua, B.-S., Tran, M.-K., ve Yeung, S.-K.** (2018). Pointwise Convolutional Neural Networks. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 984-993.
- Ioffe, S., ve Szegedy, C.** (2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *ArXiv, abs/1502.03167*.
- Ivan, I., Singleton, A., Horák, J., ve Inspektor, T.** (2017). The Rise of Big Spatial Data. *The Rise of Big Spatial Data*.
- Jobin, A., Ienca, M., ve Vayena, E.** (2019). Artificial Intelligence: the global landscape of ethics guidelines. *ArXiv, abs/1906.11668*.
- Johnson-Roberson, M., Barto, C., Mehta, R., Sridhar, S. N., Rosaen, K., ve Vasudevan, R.** (2017). Driving in the Matrix: Can virtual worlds replace human-generated annotations for real world tasks? *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 746-753.
- Joseph-Rivlin, M., Zvirin, A., ve Kimmel, R.** (2019). Momen^t: Flavor the Moments in Learning to Classify Shapes. *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, 4085-4094.
- Kingma, D. P., ve Ba, J.** (2015). Adam: A Method for Stochastic Optimization. *CoRR, abs/1412.6980*.
- Klokov, R., ve Lempitsky, V. S.** (2017). Escape from Cells: Deep Kd-Networks for the Recognition of 3D Point Cloud Models. *2017 IEEE International Conference on Computer Vision (ICCV)*, 863-872.
- Kothari, S. C., ve Oh, H.** (1993). Neural Networks for Pattern Recognition. *Adv. Comput.*, 37, 119-166.

- Kölle, M., Laupheimer, D., Schmohl, S., Haala, N., Rottensteiner, F., Wegner, J. D., ve Ledoux, H.** (2021). The Hessigheim 3D (H3D) Benchmark on Semantic Segmentation of High-Resolution 3D Point Clouds and Textured Meshes from UAV LiDAR and Multi-View-Stereo. *ArXiv, abs/2102.05346*.
- Krizhevsky, A., Sutskever, I., ve Hinton, G. E.** (2012). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60, 84-90.
- Kukacka, J., Golkov, V., ve Cremers, D.** (2017). Regularization for Deep Learning: A Taxonomy. *ArXiv, abs/1710.10686*.
- Kumar, S. K.** (2017). On weight initialization in deep neural networks. *ArXiv, abs/1704.08863*.
- Lan, S., Yu, R., Yu, G., ve Davis, L. S.** (2019). Modeling Local Geometric Structure of 3D Point Clouds Using Geo-CNN. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 998-1008.
- LeCun, Y., Boser, B. E., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W. E., ve Jackel, L. D.** (1989). Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation*, 1, 541-551.
- Lei, H., Akhtar, N., ve Mian, A. S.** (2019). Octree Guided CNN With Spherical Kernels for 3D Point Clouds. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 9623-9632.
- Li, X., Li, C., Tong, Z., Lim, A., Yuan, J., Wu, Y., . . . Huang, R.** (2020). Campus3D: A Photogrammetry Point Cloud Benchmark for Hierarchical Understanding of Outdoor Scene. *Proceedings of the 28th ACM International Conference on Multimedia*.
- Li, Y., Bu, R., Sun, M., Wu, W., Di, X., ve Chen, B.** (2018). PointCNN: Convolution On X-Transformed Points. *NeurIPS*.
- Li, Y., Ma, L., Zhong, Z., Liu, F., Cao, D., Li, J., ve Chapman, M. A.** (2021). Deep Learning for LiDAR Point Clouds in Autonomous Driving: A Review. *IEEE Transactions on Neural Networks and Learning Systems*, 32, 3412-3432.
- Linsen, L.** (2001). Point cloud representation.
- Liu, Y., Fan, B., Meng, G., Lu, J., Xiang, S., ve Pan, C.** (2019). DensePoint: Learning Densely Contextual Representation for Efficient Point Cloud Processing. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 5238-5247.
- Liu, Y., Fan, B., Xiang, S., ve Pan, C.** (2019). Relation-Shape Convolutional Neural Network for Point Cloud Analysis. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 8887-8896.
- Maturana, D., ve Scherer, S. A.** (2015). VoxNet: A 3D Convolutional Neural Network for real-time object recognition. *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 922-928.
- Munoz, D., Bagnell, J. A., Vandapel, N., ve Hebert, M.** (2009). Contextual classification with functional Max-Margin Markov Networks. *CVPR*.
- Niemeyer, J., Rottensteiner, F., ve Soergel, U.** (2014). Contextual classification of lidar data and building object detection in urban areas. *Isprs Journal of Photogrammetry and Remote Sensing*, 87, 152-165.

- Pan, Y., Gao, B., Mei, J., Geng, S., Li, C., ve Zhao, H.** (2020). SemanticPOSS: A Point Cloud Dataset with Large Quantity of Dynamic Instances. *2020 IEEE Intelligent Vehicles Symposium (IV)*, 687-693.
- Park, Y., ve Guldman, J.-M.** (2019). Creating 3D city models with building footprints and LIDAR point cloud classification: A machine learning approach. *Comput. Environ. Urban Syst.*, 75, 76-89.
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., . . . Chintala, S.** (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. *NeurIPS*.
- Pirotti, F., Guarnieri, A., ve Vettore, A.** (2013). State of the Art of Ground and Aerial Laser Scanning Technologies for High-Resolution Topography of the Earth Surface. *European Journal of Remote Sensing*, 46, 66-78.
- Qi, C., Su, H., Mo, K., ve Guibas, L. J.** (2017). PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 77-85.
- Qi, C., Yi, L., Su, H., ve Guibas, L. J.** (2017). PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. *NIPS*.
- Qian, N.** (1999). On the momentum term in gradient descent learning algorithms. *Neural networks : the official journal of the International Neural Network Society*, 12 1, 145-151.
- Rajpura, P. S., Hegde, R. S., ve Bojinov, H.** (2017). Object Detection Using Deep CNNs Trained on Synthetic Images. *ArXiv, abs/1706.06782*.
- Richter, S. R., Vineet, V., Roth, S., ve Koltun, V.** (2016). Playing for Data: Ground Truth from Computer Games. *ArXiv, abs/1608.02192*.
- Riegler, G., Ulusoy, A. O., ve Geiger, A.** (2017). OctNet: Learning Deep 3D Representations at High Resolutions. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 6620-6629.
- Ronneberger, O., Fischer, P., ve Brox, T.** (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *MICCAI*.
- Rosenblatt, F.** (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65 6, 386-408.
- Roynard, X., Deschaud, J.-E., ve Goulette, F.** (2018). Paris-Lille-3D: A large and high-quality ground-truth urban point cloud dataset for automatic segmentation and classification. *The International Journal of Robotics Research*, 37, 545-557.
- Ruder, S.** (2016). An overview of gradient descent optimization algorithms. *ArXiv, abs/1609.04747*.
- Rumelhart, D. E., Hinton, G. E., ve Williams, R. J.** (1986). Learning representations by back-propagating errors. *Nature*, 323, 533-536.
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., . . . Fei-Fei, L.** (2015). ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, 115, 211-252.

- Serna, A., Marcotegui, B., Goulette, F., ve Deschaud, J.-E.** (2014). Paris-rue-Madame Database - A 3D Mobile Laser Scanner Dataset for Benchmarking Urban Detection, Segmentation and Classification Methods. *ICPRAM*.
- Shekhar, S., Gunturi, V., Evans, M. R., ve Yang, K.** (2012). Spatial big-data challenges intersecting mobility and cloud computing. *MobiDE '12*.
- Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I., ve Salakhutdinov, R.** (2014). Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.*, 15, 1929-1958.
- Su, H., Maji, S., Kalogerakis, E., ve Learned-Miller, E. G.** (2015). Multi-view Convolutional Neural Networks for 3D Shape Recognition. *2015 IEEE International Conference on Computer Vision (ICCV)*, 945-953.
- Tan, W., Qin, N., Ma, L., Li, Y., Du, J., Cai, G., . . . Li, J.** (2020). Toronto-3D: A Large-scale Mobile LiDAR Dataset for Semantic Segmentation of Urban Roadways. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 797-806.
- Thomas, H., Qi, C., Deschaud, J.-E., Marcotegui, B., Goulette, F., ve Guibas, L. J.** (2019). KPConv: Flexible and Deformable Convolution for Point Clouds. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 6410-6419.
- Turing, A. M., ve Haugeland, J.** (1950). *Computing Machinery and Intelligence*. Cambridge: MIT Press.
- United Nations.** (2019). World Population Prospects 2019: Highlights. *Statistical Papers - United Nations (Ser. A), Population and Vital Statistics Report*.
- Url-1 <<https://assets.bwbx.io/>> erişim tarihi: 1 Ocak 2022.
- Url-2 <<https://paperswithcode.com/sota/image-classification-on-imagenet>> erişim tarihi: 9 Ocak 2022.
- Url-3 <<https://ifpwww.ifp.uni-stuttgart.de/benchmark/hessigheim/results.aspx>> erişim tarihi: 27 Aralık 2021.
- Url-4 <<https://github.com/HuguesTHOMAS/KPConv>> erişim tarihi: 20 Aralık 2021.
- Url-5 <https://github.com/Arjun-NA/KPConv_for_DALES> erişim tarihi: 22 Aralık 2021.
- Url-6 <<https://aws.amazon.com>> erişim tarihi: 26 Aralık 2021.
- Vallet, B., Brédif, M., Serna, A., Marcotegui, B., ve Paparoditis, N.** (2015). TerraMobilita/iQmulus urban point cloud analysis benchmark. *Comput. Graph.*, 49, 126-133.
- van Maarseveen, M. F., Martínez, J., ve Flacke, J.** (2018). GIS in Sustainable Urban Planning and Management.
- Varney, N. M., Asari, V. K., ve Graehling, Q.** (2020). DALES: A Large-scale Aerial LiDAR Data Set for Semantic Segmentation. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 717-726.
- Vosselman, G.** (2013). Point cloud segmentation for urban scene classification. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 257-262.

- Wang, P., Chen, P., Yuan, Y., Liu, D., Huang, Z., Hou, X., ve Cottrell, G.** (2018). Understanding Convolution for Semantic Segmentation. *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 1451-1460.
- Wang, Y., Wei, G.-Y., ve Brooks, D. M.** (2019). Benchmarking TPU, GPU, and CPU Platforms for Deep Learning. *ArXiv, abs/1907.10701*.
- Wei, S.** (2008). Building boundary extraction based on lidar point clouds data.
- Xie, Y., Tian, J., ve Zhu, X.** (2020). Linking Points With Labels in 3D: A Review of Point Cloud Semantic Segmentation. *IEEE Geoscience and Remote Sensing Magazine*, 8, 38-59.
- Xu, D., Anguelov, D., ve Jain, A.** (2018). PointFusion: Deep Sensor Fusion for 3D Bounding Box Estimation. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 244-253.
- Yan, X., Zheng, C., Li, Z., Wang, S., ve Cui, S.** (2020). PointASNL: Robust Point Clouds Processing Using Nonlocal Neural Networks With Adaptive Sampling. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 5588-5597.
- Yang, J., Zhang, Q., Ni, B., Li, L., Liu, J., Zhou, M., ve Tian, Q.** (2019). Modeling Point Clouds With Self-Attention and Gumbel Subset Sampling. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 3318-3327.
- Yang, Z., ve Wang, L.** (2019). Learning Relationships for Multi-View 3D Object Recognition. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 7504-7513.
- Ye, Z., Xu, Y., Huang, R., Tong, X., Li, X., Liu, X., . . . Stilla, U.** (2020). LASDU: A Large-Scale Aerial LiDAR Dataset for Semantic Labeling in Dense Urban Areas. *ISPRS Int. J. Geo Inf.*, 9, 450.
- Ying, Y., Koeva, M., Kuffer, M., ve Zevenbergen, J.** (2020). Urban 3d modelling methods: a state-of-the-art review. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 699-706.
- Yu, X., Hyypä, J., Karjalainen, M., Nurminen, K., Karila, K., Vastaranta, M., . . . Katoh, M.** (2015). Comparison of Laser and Stereo Optical, SAR and InSAR Point Clouds from Air- and Space-Borne Sources in the Retrieval of Forest Inventory Attributes. *Remote. Sens.*, 7, 15933-15954.
- Yue, X., Wu, B., Seshia, S. A., Keutzer, K., ve Sangiovanni-Vincentelli, A. L.** (2018). A LiDAR Point Cloud Generator: from a Virtual World to Autonomous Driving. *Proceedings of the 2018 ACM on International Conference on Multimedia Retrieval*.
- Zaheer, M., Kottur, S., Ravanbakhsh, S., Póczos, B., Salakhutdinov, R., ve Smola, A.** (2017). Deep Sets. *NIPS*.
- Zhao, H., Jiang, L., Fu, C.-W., ve Jia, J.** (2019). PointWeb: Enhancing Local Neighborhood Features for Point Cloud Processing. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 5560-5568.

- Zhou, G., ve Di, W.** (2020). Extraction of houses from point cloud lidar: problems and challenge. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 831-837.
- Zhou, Q.-Y., Park, J., ve Koltun, V.** (2018). Open3D: A Modern Library for 3D Data Processing. *arXiv:1801.09847*.
- Zhu, X., ve Bamler, R.** (2010). Very High Resolution Spaceborne SAR Tomography in Urban Environment. *IEEE Transactions on Geoscience and Remote Sensing*, 48, 4296-4308.
- Zolanvari, S. M., Ruano, S., Rana, A., Cummins, A., da Silva, R. E., Rahbar, M., ve Smolic, A.** (2019). DublinCity: Annotated LiDAR Point Cloud and its Applications. *BMVC*.







ÖZGEÇMİŞ

Ad-Soyad : Mert BEŞİKTEPE

ÖĞRENİM DURUMU:

- **Lisans** : 2018, Yıldız Teknik Üniversitesi, Harita Mühendisliği Bölümü

MESLEKİ DENEYİM VE ÖDÜLLER:

- 2018 yılında Yıldız Teknik Üniversitesi Harita Mühendisliği Bölümünden 3.'lük derecesi ile mezun oldu.
- 2019 yılında İstanbul Teknik Üniversitesi Geomatik Mühendisliği Programı'nda lisanüstü eğitime başladı.
- 2021 yılında İller Bankası Genel Müdürlüğü Kaynak Geliştirme Dairesi İmar ve Planlama Müdürlüğünde mühendis olarak göreve başladı.

YÜKSEK LİSANS TEZİNDEN TÜRETİLEN YAYINLAR, SUNUMLAR VE PATENTLER:

- **Beşiktepe M., Güney C.** 2021. Nokta Bulutlarının Modellenmesinde Derin Öğrenme Yaklaşımı. *International Aegean Scientific Research Symposium*, Aralık 25-26, 2021 İstanbul, Türkiye.