



Düzce Üniversitesi Bilim ve Teknoloji Dergisi

Derleme Makalesi

Derin Öğrenme Yöntemleri ile 3B Nokta Bulutlarının Semantik Segmentasyonuna Genel bir Bakış

 Muhammed Ahmet DEMİRTAŞ

^a Yazılım Mühendisliği Bölümü, Mühendislik Fakültesi, Muş Alparslan Üniversitesi, Muş, TÜRKİYE

* Sorumlu yazarın e-posta adresi: ahmet.demirtas@kocaeli.edu.tr

DOI: 10.29130/dubited.1004211

Öz

Semantik segmentasyon, çevredeki nesnelere anlam vermek için etiketlenmiş her pikseli anlamlı bir sınıfa atayan bir veri işleme yöntemidir. Derin öğrenme (DÖ) tabanlı yöntemlerin geliştirilmesi, Nokta Bulutu (NB) ile segmentasyon yöntemlerine olan ilgiyi artırmıştır. 3 Boyutlu (3B) nokta bulutu semantik segmentasyonu, farklı tarama araçları ile elde edilen 3B veri setlerinde aynı bölgede aynı özelliklere sahip noktaları homojen bölgelere ayırmaktadır. 3B nokta bulutları ile 3B nesnelere anlamak için semantik segmentasyonun kullanılması önemli bir başlangıç olmuştur. Özellikle derin öğrenme yöntemlerinin kullanılması bu alanı odak noktası haline getirmiştir. 3B yapılandırılmamış büyük nokta bulutlarını işlerken, derin öğrenmeyi temel alarak geliştirilen yeni yöntemler, yaklaşımlar ve modeller üzerinde benzersiz sorunlarla karşılaşılması bu alanın gelişime açık olduğunu göstermektedir. Bu yeni yöntemlerin başarılarını anlamak için, kıyaslama veri kümeleri: ShapeNet, S3dis, ScanNet, SemanticKITTI üzerindeki performansları değerlendirilmiştir. 3B nokta bulutu ile segmentasyon alanına katkıda bulunan dikkate değer araştırmalar incelenmiş, avantajları, dezavantajları ve önerilen yöntemlerin katkıları sunulmuştur. Sunulan tüm yöntemlerin mimari yapısı, yaygın olarak kullanılan veri kümeleri üzerindeki başarıları tartışılmış ve gelecekteki araştırmalara öncülük edecek bilgiler önerilmiştir.

Anahtar Kelimeler: 3B Nokta Bulutu, Derin Öğrenme, Semantik Segmentasyon

Review on Semantic Segmentation of 3D Point Clouds with Deep Learning Methods

ABSTRACT

Semantic segmentation is a data processing method that assigns each labeled pixel to a meaningful class to give meaning to surrounding objects. The development of deep learning-based methods has increased the interest in point cloud segmentation processes. 3D point cloud semantic segmentation in 3D datasets obtained with different scanning tools divides, points with the same feature in the same region are divided into homogeneous regions. The use of semantic segmentation to understand 3D point clouds with 3D objects has been an important start. In particular, the use of deep learning methods has made this area a focal point. In particular, When processing large unstructured 3D point clouds, encountering unique problems on new methods and approaches developed on the based on deep learning has shown that this area needs to be improved. In order to understand the achievements of these new methods, their performance on 3D benchmark datasets ShapeNet, S3dis, ScanNet, SemanticKITTI was evaluated. Important research contributing to the segmentation field with 3D point cloud is analyzed, the advantages, disadvantages and contributions of the proposed methods are presented. The architectural structure of

all the presented methods and their achievements on widely used datasets are discussed and he offered information that would lead to future research.

Keywords: 3D Point Cloud, Deep Learning, Semantic Segmentation

I. GİRİŞ

Son yıllarda 3B teknolojilerinin hızla gelişmesiyle birlikte 3B tarayıcılar, LIDAR (Light Detection and Ranging) ve RGB-D kameralar (Kinect, RealSense ve Derinlik kameraları) 3B sensörler kullanılarak [1] elde edilen 3B veri kümelerinden anlamlı geometrik özellikler ve ölçülebilir bilgiler alınmıştır [2]. Gelişen teknoloji sayesinde elde edilen 3B veri setleri, hayatımızda sağlık alanında, kentsel dönüşüm projelerinde bilgisayarlı görü ile farklı alanlarda yeni model ve uygulamanın geliştirilmesine zemin hazırlamıştır [3]. 3B tarayıcıların yaygınlaşması ile 3B nokta bulutları farklı açılardan ve istenilen amaçlar doğrultusunda kolaylıkla elde edilmiştir. 3B veri kümelerinin sürekli artması sonucu, bu verileri hızlı ve güvenilir bir şekilde işleyecek kütüphanelere ihtiyaç doğurmuştur. Bu alanda en iyi bilinen nokta bulutu kütüphanesi (Point Cloud Library) 3B nesne algılama, sınıflandırma ve semantik segmentasyon algoritmalarını uygulamada kolaylıklar sağlamıştır [4].

Yukarıda bahsedilen gelişmeler sonucunda 3B modelleri anlamlı bölgelere ayırmak için, yoğun araştırmalar yapılmış yön, pürüzsüzlük, içbükey bilgilerinden bir küme oluşturacak noktalar kümelendirilmiştir. Nokta bulutlarını kümeleme aşamasında yaşanan bu problemlere Shamir dışbükey ayrıştırma, bölge genişlemesi, hiyerarşik ve spektral kümeleme yaklaşımlarını önermiştir [5]. Sunulan yaklaşımlar, bölge temelli nokta bulutlarını bölümlere ayırmada yaygın olarak kullanılmıştır [6], [7], [8], [9]. Bunlara ek olarak grafik tabanlı yöntemler Normalized Kesimler [10] ve Minimum Kesimler (uygun değer tam bölümlenme) [11], Grafik Kesimleri [12] en popüler yöntemlerdir.

Anguelov'e göre, 3B NB algoritmaları üç özelliğe sahip olmalıdır: Birincisi; farklı özellikleri öğrenerek farklı nesnelere ayırt etmeli ve özellik sayısı arttıkça algoritma bunu otomatik öğrenebilmelidir. İkincisi; komşu bilgilerine göre seyrek örneklenmiş bölgelerdeki noktaları etiketlemelidir. Üçüncüsü ise bu algoritmanın 3B tarayıcılara uygulamasıdır. Bunun amacı çeşitli lazer tarama araçlarının farklı NB üretebilmesidir. Uyarlanan bu algoritma aynı sahnede farklı özelliklere sahip olabilir [13]. Bu özellikler kullanılarak nesnelere yapıları hakkında bilgiler elde edilir ve semantik segmentasyon ile yorumlanabilir. Nesne yapılarından yola çıkarak semantik segmentasyon ile nesnelere bölümlenme, konumlandırma ve tanıma özellikleri analiz edilerek anlamlı bilgiler elde edilmektedir.

2 boyutlu (2B) görüntülerin semantik olarak bölümlenmesi akademik çalışmalarını bu yöne çevirmiştir fakat 2B görüntü üzerindeki tıkanıklıklar diğer özellikleri engellediği için performans kaybı yaşanmıştır. Bu tıkanıklıkların sonucunda 2B NB yöntemlerini 3B yeni modellere çevirmiştir. Bunun sonucunda 3B NB, 3B veri işleme araştırmalarında önemli bir yere sahip olmuştur. NB'nin literatüre katkıları, otonom araçlar, kentsel dönüşümler, navigasyon, kültürel mirasın korunması gibi birçok alanda verileri noktasal olarak etiketleyip yeni yöntemlerin geliştirilmesine olanak sağlamıştır [14]. Ayrıca 3B verilerin çeşitli alanlarda yaygın olarak kullanılmasının bir diğer nedeni 3B verilerinin farklı dosya formatlarında temsil edilmesidir.

II. MODELLERİN PERFORMANSINI KIYASLAYAN VERİ KÜMELERİNİN ÖZELLİKLERİ

3B NB örnekleme yoğunluğu, tarayıcının değişken olması veya 3B sensör sınırlamaları nedeniyle, ön plan genelde arka planla iç içe olmaktadır. Bu da segmentasyon algoritması tasarlarken bazı problemler ve zorlukları beraberinde getirmiştir. Ayrıca NB'lerin seyrek olması, rastgele yapısal özelliklerin bozuk olması, NB tabanlı semantik segmentasyonu zorlaştırmıştır. Tüm bu zorlukların üstesinden gelebilmek

için kıyaslama (benchmark) veri setleri üzerindeki semantik segmentasyon performans sonuçları değerlendirilmiştir. Semantik segmentasyon, sınıflandırma, nesne tespit ve takibi için yeni DÖ tabanlı yöntemler sunulmuştur. Bu yöntemleri kıyaslamak için yaygın kullanılan veri setleri: S3DIS [15] ShapeNet [16], KITTI Vision Benchmark Suite [17], [18], ScanNet [19], Semantic3D [20], PartNet [9], [21], ApolloCar3D [22] ve ScanObjectNN [23], ModelNet, [24] olmuştur. Ek olarak, 3B NB verilerini kaydetmek, altı serbestlik düzeyinde (6-DOF/6 degrees of freedom) poz tahmini ve 3B verileri yeniden yapılandırma [25], [25], [26] yöntemleri kullanılmıştır.

Tablo 1. En Yaygın Kullanılan Veri Setleri.

Veri Seti	Noktalar	Kategori	Year
Oakland	1.6M	5	2009
ISPRS	1.2M	9	2012
Paris-rue-Madame	20M	17	2014
IQmulus	300M	8	2015
ScanNet	-	20	2017
S3DIS	273M	13	2017
Semantic3D	4000M	8	2017
Paris-Lille-3D	143M	9	2018
SemanticKITTI	4549M	25	2019
Toronto-3D	78.3M	8	2020
DALES	505M	8	2020

Yukarıda verilen 3B NB'leri ile 3B nesne tanıma-tespit, sınıflandırma ve segmentasyon yapılırken meydana gelen çeşitli sorunlar ele alınmıştır.

Özet olarak nokta bulutlarını toplamak için geleneksel olarak kullanılan veri toplama araçları ile güncel olarak kullanılan temassız veri toplama araçlarının farkı; gerçek zamanlı olması, verilerin dinamik olması, yüksek yoğunluk ve verimliliğe sahip olmasıdır. Ayrıca nokta bulutları gibi 3B veriler, sadece iki boyutlu görüntülerde karşılaşılan aydınlatma ve poz (bakış açısı) sorunlarını gidererek karmaşık sahneler içinde zengin içeriğe sahip bilgiler vermiştir.

III. 3B NOKTA BULUTU İLE SEMANTİK SEGMENTASYON

3B NB segmentasyonu, aynı bölgede aynı özelliklere sahip nokta bulutlarını birden fazla homojen bölgeye ayırma işlemidir. 3B NB segmentasyonu, nesnelerin detaylı olarak geometrik bilgilerinin genel ve ayrıntılı bir şekilde özniteliklerinin anlaşılmasını sağlamaktadır.

Segmentasyon, eşit olmayan noktaların yoğunluğu ve NB'lerin açık yapısından yoksun olması nedeniyle zordur. Bu sorun robotikte akıllı araçlar, otonom haritalama ve navigasyon gibi uygulamalarda farklı yaklaşımlar ortaya çıkarmıştır [27]. Geleneksel NB segmentasyon algoritmaları üzerinde başlıca şunları içermektedir: öznitelik kümelemeye dayalı yöntemler [28], [29], model oluşturma yöntemleri [30], [31], bölge büyümesine dayalı yöntemler [31] grafik kesiminde [32], [33] ve kenarına dayalı yöntemlerdir [34].

Semantik segmentasyon bir görüntü içerisinde farklı özelliklerin tutulduğu anlamlı bölgelerdeki NB anlamlarına göre alt sınıflara ayırma işlemidir. 3B nesnelerin segmentasyonu dört genel paradigmadan oluşur: Projeksiyon, ayrık, hibrit ve nokta, tabanlı yöntemlerdir.

Projeksiyon ve ayrıklaştırma tabanlı yöntemlerinin ilk aşaması, nokta bulutlarının çoklu gösterimi [35], [36] küresel gösterim [37], [38], [39] hacimsel gösterim [40], [41], [42] Permutohedral ızgara gösterimi

[43], [44] ve hibrit gösterimi [45], [46] gibi düzenli bir ara temsile dönüştürmektedir. Sonra düzensiz NB tekrar gönderilerek yapısız NB'ye direkt uygulanmıştır.

3.1 PROJEKSİYON TABANLI YÖNTEMLER

Projeksiyon tabanlı yöntemler çoklu görünüm ve küresel görünüm yöntemlerini kullanarak 3B görüntülerin 2B olarak yansıtılmasıdır.

A.1. Çoklu görünüm gösterimi

Lewin vd., 2B düzlem üzerine ilk olarak birden fazla sanal kamera görüntüsünden oluşan 3B NB yansıtılar [36] ve oluşan sentetik görüntülerdeki pikselleri tahmin etmek için çoklu akış kullandılar. Her bir noktanın semantik etiketlenmesi, yeniden yansıtılan temsili değerlerin farklı görünümde birleştirilmesiyle elde edilmiştir. Benzer şekilde, Baulch vd., ilk olarak birden fazla kamera ile NB elde edilen anlık RGB (red green blue) ve derinlik görüntüsü oluşturmuşlar. Sonrasında 2B segmentasyon ağlarını kullanarak anlık görüntülerdeki pikselleri etiketlemiştir. RGB ve derinlik görüntülerinin tahmini sonuçları artık düzeltme yöntemi ile birleştirilmiştir [36], [47].

NB yerel yüzeylerinde Öklid ile örneklendiği varsayımına dayanarak Tatarchenko vd., yoğun NB segmentasyonu için teğet kıvrımlarını tanıtmışlardır [48]. Bu yöntemde noktalar yüzey geometrisinde doğrudan işlenmektedir. Aynı zamanda büyük ölçeklenebilir (milyonlarca noktayı) NB işleyebilirken geometrik ve yapısal bilgileri tam olarak kullanılamamaktadır.

A.2. Küresel Gösterim

3B NB ile hızlı ve doğru bir şekilde segmentasyon yapılması için Wu vd., SqueezeNet ve Koşullu Rastgele Alana (Conditional Random Field /CRF) dayalı uçtan uca bir ağ önermiştir [48]. Segmentasyon doğruluğunu iyileştirmek için, SqueezeSegV2 [38] denetimsiz bir alanda ardışık düzen uyarlamasını kullanarak alan kayması yönünden segmentasyon doğruluğunu ele almıştır.

Milioto vd., NB ile semantik segmentasyonu gerçek zamanlı işlemek için RangeNet ++ yöntemini önermişlerdir [39]. 2B etiketli verilerin anlamlı ara temsillerini 3B'ye dönüştürürken, KNN (k-near neighbours) temelli işlem hatalarını ve çıktılardaki bulanıklığı düzeltmek için kullanmıştır. Tek görüntülü projeksiyon yöntemi küresel projeksiyon yöntemine göre daha fazla bilgi tutar ve bu yöntem ile LIDAR ile alınan NB'lerin etiketlenmesi daha uygundur. Bu yöntem kullanılırken geçişler sırasında oluşan ara temsillerden kaynaklanan ayırma hataları veya tıkanma gibi sorunları meydana getirmiştir.

B. AYRIŞTIRMA TABANLI YÖNTEMLER

Bu yöntemler, genellikle nokta bulutlarını hacimsel veya permutahedral kafesler gibi yoğun veya seyrek gösterime dönüştürür. NB'yi işlerken yoğun ızgaralar vokselleştirilerek 3B evrişim ağları kullanılmıştır. Huang vd., nokta bulutunu doluluk vokseline böldü ve 3B voksellerin segmentasyonu için 3B CNN'den (convolutional neural network) geçirdi [48]. Bu yöntemde her bir voksel aynı etikete sahip olur; fakat performans bakımından NB'leri sınırlamıştır. Tchapmi vd., ince taneli ve global semantik segmentasyon elde etmek için SEGCloud'u önermiştir [49]. Bu yöntem, 3D-FCNN (Full Connected CNN) tarafından oluşturulan voksellerin tahminlerini NB'ye eşlemek için deterministik üç doğrusal enterpolasyon sunmuş ve etiketli her nokta için 3B FC-CNN kullanmıştır [50]. Meng vd, her bir vokselde yerel geometrik yapıları kodlamak için çekirdek tabanlı değişken bir yapıya sahip otomatik kodlayıcı bir mimari geliştirmiştir [41]. Meng her bir voksel noktasının dağılımını yakalamak ve sürekli bir temsil sağlamak için RBF (radial basis function) kullanılmıştır. VAE-Net (variational autoencoder/VV-Net) her bir voksel içindeki yoğun nokta dağılımını gizli bir alana eşledikten sonra sağlam özellikleri öğrenmek için simetri olan grupları ve eşdeğer CNN'leri kullanmıştır.

3B CNN'nin iyi bir şekilde ölçeklenmesi hacimsel tabanlı ağlarda, farklı uzaysal boyutlardaki nokta bulutlarının test edilmesini ve eğitilmesini sağlamaktadır. Tam Evrişimli Nokta Ağı (FCPN/ Fully-Convolutional Point Network) [40], farklı düzeydeki geometrik yapıları hiyerarşik olarak nokta bulutlarından ayırır. Bu yapıdaki bitişik özellikleri çıkarmak, uzun menzilli bağımlılıkları birleştirmek için 3B evrişimler ve ağırlık değerlerini ortalayan havuzlama sistemi kullanılır. Bu yöntem, büyük ölçekli nokta bulutu verilerini işler ve özellik çıkarım zamanında iyi ölçeklenebilir bir yapıya sahiptir. Dai vd., her bir vokseli anlamsal etiketleme ve 3B taramanın tamamlanması için ScanComplete'i önerdiler [51]. Bu yöntemde, tam evrişimli ağların ölçeklenebilirliği kullanılmış ve eğitim-test aşamasında farklı veri boyutlarındaki girdilere uyum sağlanmıştır. Tahmini sonuçlar çözünürlüğünü hiyerarşik olarak düzeltmek için kabadan inceye doğru bir (coarse-to-fine) stratejisi kullanılmıştır.

Genellikle hacimsel temsiller 3B nokta bulutlarına (NB'lerin) komşuluk yapısını korumaktadır. Düzenli bir veri yapısı genel 3B evrişim ağına doğrudan uygulanmıştır. Bu etkenler, düzenli bir performans artışı sağlamış fakat noktaları vokselleştirme aşamasında yapay ayırmalara ve bilgi kaybına neden olmuştur. Yüksek çözünürlükler büyük bellek alanına ve hesaplama maliyetine, düşük çözünürlükler ayrıntı kaybına sebep olmuştur. Yüksek veya düşük çözünürlük için doğru ve uygun bir ızgara seçiminin önemini göstermiştir.

Hacimsel gösterimde sıfır olmayan değerler yalnızca küçük bir yüzdeyi açıkladığı için seyreklerdir. Bu gibi yöntemlerde, düşük boyutlu veriler için yoğun bir sinir ağı kullanmak verimsizdir. Bu amaçla, Graham vd., indeksleme yapısına bağlı olarak alt manifoldlu seyrek evrişim ağlarını önermişler [51]. Bu yöntemde, konvolüsyon ağı dolu vokseller tarafından kısıtlandığı için bellek ve hesaplama maliyetini önemli derecede düşürmektedir. Bu alt manifoldlu seyrek konvolüsyon ağı, yüksek boyuta sahip öznelikler ve seyrek özelliklere sahip özneliklerin daha verimli bir biçimde işlenmesine yol açmıştır. Seyrek evrişim ağını etkin bir şekilde kullanmak için, Choy vd., 3B video algılamak amacıyla MinkowskiNet isminde 4B uzay-zamansal evrişimli bir sinir ağı önerdi [52]. Bu yöntemde yüksek boyutlu verileri verimli bir şekilde işlemek için genel seyrek bir evrişim ağı kullanılmış ve uyum sağlamak için üç yönlü-sabit koşullu rastgele bir alan üzerinde uygulanmıştır.

Su vd., İkili Evrişim Katmanlarına (Bilateral Convolution Layers/BCL) dayalı Seyrek Kafes Ağlarını (SPLATNet) önermiştir [52]. SPLATNet hem bir nokta bulutunu seyrek bir permutohedral kafese ekler ve BCL (bilateral convolution layers) kafesin dolu kısımlarına evrişim ile filtrelenen değerleri tekrar ham NB'ye ekler. Bu yöntemin başka bir avantajı, çoklu görüntüleri ve nokta bulutlarını esnek şekilde işlemesidir. Rosu vd., LatticeNet'i büyük nokta bulutunu verimli bir şekilde işlemek için önermiştir [52]. DeformsSlice isminde verilere bağlı bir enterpolasyon modülünü ve kafes özelliğini nokta bulutuna yeniden yansıtan bir yapı sunmuştur.

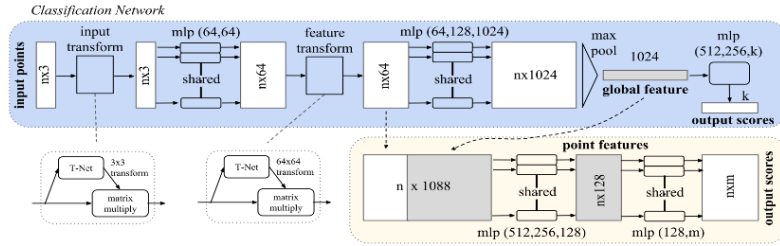
C. HİBRİT YÖNTEMLER

Bu yöntemler tüm verilerden maksimum derecede yararlanmak ve 3B tarama ile çok yönlü özellikleri öğrenmek için kullanılır. Dai ve Nießner, RGB özelliklerini ve geometrik özellikleri birleştirmek için ortak bir 3B çoklu görüntü ağı sunmuştur [45]. 2B akış özelliklerini çıkarmak için 3B CNN'i kullanmış ve öğrenilen 2B yerleştirmeleri ve 3B geometrik özellikleri birleştirmek için değişken bir geri yansıma katmanı olarak kullanılmıştır. Chiang vd., nokta bulutlarından 2B doku görünümü, 3B yapıları ve küresel bağlam özelliklerini öğrenmek için nokta tabanlı birleşik çerçeve önermiştir [45]. Bu yöntem, noktasal ağ yapısını ve geometrik özelliklerini bulmak için doğrudan uygulamıştır. Jaritz vd., kanonik NB uzayındaki 2B çoklu görüntülerden ve uzamsal geometrik özelliklerin görünüm özelliklerini gruplamak için Çoklu Görünüm PointNet'i (multi-view point network /MVPNet) sunmuşlardır [46].

D. NOKTAYA DAYALI YÖNTEMLER

Nokta tabanlı ağlar doğrudan düzensiz ve yapılandırılmamış olması nedeni ile doğrudan standart CNN'lere uygulanamaz. Bu amaçla, ilk çalışma da PointNet'in [53] simetrik havuzlama ile paylaşılan MLP'leri, (multi-layer perceptron) her bir noktanın özelliklerini öğrenmesi için önerilmiştir. PointNet'e dayalı, nokta tabanlı ağlar, MLP yöntemleri, noktasal evrişim ağları, RNN (recurrent neural network) tabanlı ve grafik tabanlı birçok yöntem geliştirilmiştir. Tablo 2'de noktasal tabanlı yöntemler temellerine göre listelenmiştir.

Nokta tabanlı yöntemler, yüksek performans elde etmek için ağ yapısına MLP uygularlar. MLP ile çıkarılan noktasal özellikler nokta bulutlarındaki yerel geometrik yapıyı ve noktaların karşılıklı etkileşimlerini yakalayamaz [53].

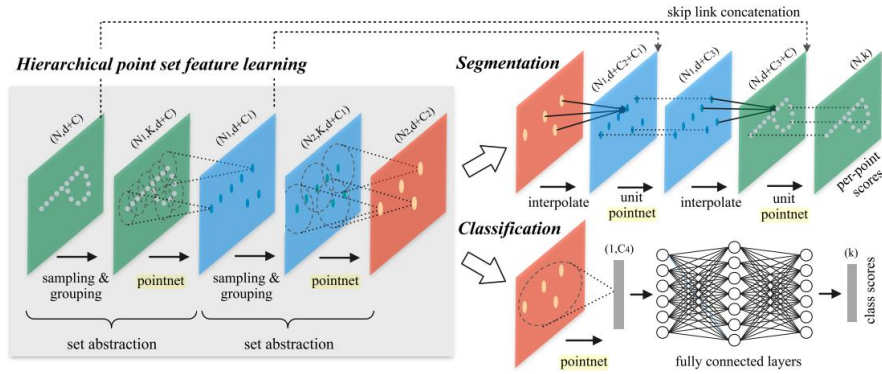


Şekil 1. PointNet yapısı.

Şekil 1'de PointNet yapısı incelendiğinde her noktada daha geniş bir alan yakalamak ve daha zengin yerel yapıları öğrenmek için komşu özellik havuzlaması, odak tabanlı toplama ve yerel-küresel özellikleri birleştirmeye bağlı birkaç özel ağ tanıtılmıştır

D.1. Komşu Özellik Havuzu

Bu yöntemler yerel geometrik desenleri yakalamak için yerel komşu nokta bilgilerini toplayarak her noktadan bir özellik öğrenir. Özellikle, PointNet++ [54], Şekil 2'te gösterildiği gibi hiyerarşik ve gittikçe daha büyük yerel bölgelerden öğrenen noktaları gruplandırır.



Şekil 2. Hiyerarşik özellik öğrenme mimarisinin 2B Öklid uzayındaki segmentasyonu ve sınıflandırmasıdır[54].

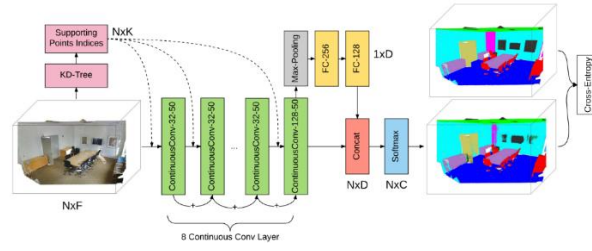
Noktaların değişken olması ve nokta buluntularındaki yoğunluk değişiminin neden olduğu sorunları çözmek için çok ölçekli ve çok çözünürlüklü gruplama yöntemleri önerilmiştir. Jiang vd., yönlendirme kodlaması ve farklı boyutlar için PointSIFT modülünü önermişlerdir[55]. Bu modül, sıralı bir evrişim yoluyla üç aşamalı sekiz uzamsal yönelimden gelen bilgileri kodlar ve kümeler. Yüksek boyutlu özellikleri farklı boyutlarla uyum sağlaması için birleştirmiştir. PointNet++ gruplama yöntemlerinden farklı olarak Engelmann [56], alanı ve özellik uzayındaki iki komşuyu ayrı ayrı belirlemek için K-ortalama (K-means) kümeleme ve KNN kullanmıştır. Özellik uzayında aynı sınıfa ait noktaların daha yakın olmasının beklendiği varsayımına göre bu özellikleri düzenli bir şekilde öğrenmek için ikili bir mesafe kaybı ve bir ağırlık merkezi kaybı meydana gelmiştir. Farklı noktalar arasındaki karşılıklı

etkileşimleri modellemek için Zhao [57] PointWeb'e yerel tam bağlı ağ oluşturarak yerel bölgedeki tüm nokta çiftleri arasındaki ilişkileri tanımlamıştır. Uyarlanabilir Özellik Ayarlama (AFA) modülünü, ağı bilgi alışverişini sağlamak ve özellik iyileştirme için önermiştir. Bu işlem, ağı ayırt edici özellik gösterimini öğrenmesine yardımcı olmaktadır. Zhang, eş merkezli küresel kabuklardan/yüzeylerden elde edilen istatistiklere dayanarak Shellconv adında değişmez bir permutasyona sahip evrişim ağı önermiştir [58]. Bu yöntem ilk olarak eş merkezli çok ölçekli alanı sorgular ardından maksimum havuzlama ile farklı kabuklar/yüzeyler içindeki istatistikleri özetler. Sonunda MLP'ler ve 1B (1 boyutlu) evrişimi ile çıktıyı elde eder. Hu, büyük ölçekli NB segmentasyonu için RandLA-Net'i önermiştir [59]. RandLA-Net' bellek alanı ve hesaplama maliyeti bakımından yüksek performans göstermek için gelişigüzel noktalar kullanmıştır. Geometrik özellikleri yakalamak ve korumak için yerel özellik toplama modülü önerilmiştir.

Yerel-küresel birleştirme: Zhao, nokta bulutlarındaki yerel yapıları ve küresel bağlamı birleştirmek için değişmez bir permutasyon olan PS2-Net önermiştir [60]. Edgeconv [61] ve NetVLAD [62], görüntüde küresel özellikleri ve yerel bilgileri elde etmek için tekrarlı bir şekilde kümelemiştir.

D.2. Nokta Eğilimli Evrişim Yöntemleri

Nokta evrişim yöntemleri, NB'leri verimli bir şekilde işlemek için evrişim ağları önermektedir. Hua vd., çekirdek hücrelerine komşu noktaları bağlayıp çekirdek ağırlıklarıyla birleştirilen noktasal bir evrişim önermiştir [63]. Şekil 3'de gösterildiği gibi Wang, PCNN isiminde parametrik sürekli evrişim katmanlarına sahip bir ağ önermiştir [64].



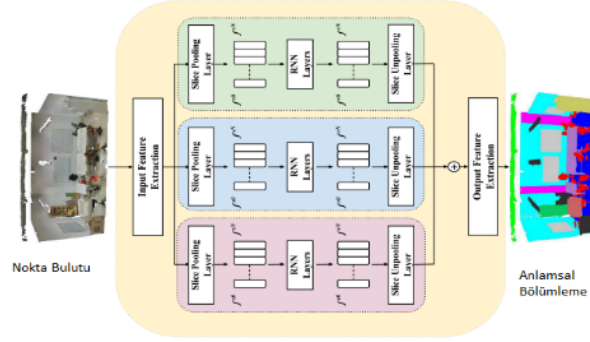
Şekil 3. Anlamsal Etiketleme için Derin Parametrik Sürekli CNN yapısı.

Bu katmandaki çekirdek fonksiyonu MLP kullanılarak parametreleri ayırır ve sürekli vektör uzayını kapsar. Thomas, çekirdek nokta konvolüsyon (Kernel Point Convolution /KPCConv) temelli tam evrişimli çekirdek nokta ağını (Kernel Point Full Convolutional Network /KP-FCNN) önermiştir [65]. KPCConv'un çekirdek noktalarındaki evrişim ağırlıkları Öklid ile belirlenir ve bu ağı çekirdek nokta sayısı sabit değildir. Çekirdek noktalarının konumları bir küre uzayında iyi bir şekilde kapsama alanı oluşturmak için bu konumlar en uygun şekle sokma problemi olarak formülleştirilmiştir. Yarıçap komşuluğunda daha tutarlı bir alıcı alan oluşturmak ve her katmanda farklı yoğunluktaki nokta bulutundan yüksek verimlilik elde etmek için alt ızgara örnekleme kullanılmıştır. Engelmann, alıcı alanın toplama tabanlı yöntem performansının etkisini göstermek için çözünürlüğü yüksek görselleştirmeler yapmıştır [47]. Ayrıca, KNN yerine genişlemiş komşu özelliklerini toplamak için Genişletilmiş Nokta Evrişimi (Dilated Point Convolution/DPC) ağını önermiştir. Bu işlemin alıcı alanı büyütmede etkili olduğu ve diğer kümelenebilirliğe dayalı ağlarla kolayca bütünleştiği gösterilmiştir.

D.3. Tekrarlanan Sinir Ağı Tabanlı Yöntemler

RNN (recurrent neural network) hem nokta bulutlarının bağlam ve içerik özelliklerini yakalamak hem de semantik segmentasyon için kullanılmıştır. PointNet [53] temel olarak Engelmann, girdi katmanındaki nokta bilgilerine ulaşmak için nokta bulutunu çok ölçekli ve ızgara bloklarına dönüştürmüştür [66]. PointNet ile elde edilen özelliklere ve çıkış katmanındaki bilgilere ulaşmak için Tekrarlayan Konsolidasyon Birimleri (Recurrent Consolidation Units /RCU) ile beslenir. Deneysel sonuçlar, bölümlenme performansının iyileştirilmesi için uzamsal bağlam bilgilerini birleştirmedeki önemi göstermiştir. Huang vd., yerel bağımlılıkları modellemek için bir modülü önermiş ve düzensiz

nokta özellik kümelerini sıralı özellik vektör dizisine dönüştürmek için bir dilim havuzlama (Recurrent Slice Network /RSNet) katmanı kullanmıştır [67].



Şekil 4. RSNet X,Y,Z eksenini boyunca dilimleme yönü gösterimi

Şekil 3'te gösterildiği gibi, Ye vd., ilk olarak kabadan-inceye yerel yapıyı yakalamak için Noktasal Piramit Havuzlama (3P) modülü önerdiler ve uzun menzilli uzamsal bağımlılık aralığını elde etmek için iki yönlü hiyerarşik RNN kullanılmıştır [67]. Sonrasında ağıncı uçtan uca öğrenmesi için RNN kullandılar. Bununla birlikte, yerel komşuluk özelliklerini küresel yapı özellikleriyle birleştirirken kullanılan bu yöntemler NB verilerinden zengin geometrik özelliklerini ve yoğunluk dağılımını kaybettiler [68]. Sabit yapı ve statik havuzlama işlemlerinin neden olduğu sorunları hafifletmek için, Zhao vd., küresel sahne karmaşıklığını ve yerel geometrik özellikleri dikkate alarak Dinamik Toplama Ağını (DAR-Net) önermişlerdir [68]. Ortamlar arası özellikler ortama kendi kendine uyum sağlayan alıcılar ve düğüm aralıkları dinamik olarak toplanmıştır. Liu vd., yüksek boyutlu NB'yi verimli ve daha başarılı bir semantik ayrışım yapılabilmesi için 3D-CNN-DQN-RNN'yi önerdi [69]. Bu ağ 3B bir CNN ağı ile noktaların uzaydaki yayılımını ve renk özelliklerini öğrenmektedir. Derin Q ağını (DQN'i (Deep Q-Network) belirtilen sınıfa ait nesnelere sınırlandırmak için kullanmıştır. Sonuç olarak birleştirilen özellik vektörü, segmentasyon sonuçlarına ulaşmak için bir RNN ile beslenmiştir.

Tablo 2. Nokta Tabanlı Derin Öğrenme Yöntemlerinin özeti.

Point Net Yöntemler	Tabanlı Özellik	PointNet++ Yöntemler	Tabanlı Özellik	Temel Yapı
PointNet	MLP	PointNet++	PN	MLP
G+RCU	MLP	PointWeb	PN	
ESC	MLP	PointSIFT	PN	
HRNN	MLP			
RSNet	1x1 Conv	MCC	MCCConv.	Noktasal Evrişim
DPC	PointConv	PointConv	PointConv	
PWCNN	PWConv	A-CNN	AConv	
KPCConv	KPCConv	RandLA-Net	LocSE	
PCNN	PCCConv	PolarNet	PN	

FlexConv	flexConv			
PointCNN	χ -Conv			
DGCN	EdgeConv	LS-GCN	Spec.Conv.	Grafiksel Evrişim
SPG	PN	PAN	LAE-Conv	
DeepGCNs	DGConv	TGNet	TGConv	
SPH3D-GCN	SPH3D-GConv	HDGCN	DGConv	
PGCRNet	Conv1D	3DCon.Net	PN	
AGCN	MLP	ψ -CNN	ψ -Conv	

Tablo 2’ de gösterildiği gibi PointNet ve PointNet++ tabanlı birçok derin öğrenme yöntemi geliştirilmiş ve kıyaslama veri setleri üzerinde semantik segmentasyon yapılmıştır. PointNet ve PointNet++ yöntemler gibi temel teşkil edecek yöntemlerin kısa sürede başarılı bir şekilde geliştirilmesi için yöntemler detaylı olarak anlatılmıştır.

D.4. Grafik Tabanlı Yöntemler

3B NB’deki şekilleri ve geometrik yapıları yakalamak için grafik tabanlı yöntemlere başvurulmuştur. Landrieu vd., NB’leri birbirine bağlı basit şekiller süper noktalar kümesi ile temsil edilen bağlam bilgisini ve yapı bilgilerini yakalamak için bağlı grafik (yani süper nokta grafiği) kullanmıştır [70]. Sonrasında büyük ölçekli NB bölümlene problemini üç alt probleme bölerek: homojen geometrik bölümlene, süper nokta gömme ve bağlamsal bölümlene gibi çözümler kullanılmıştır. Bölme adımında başarıyı artırmak için Landrieu ve Boussaha NB’yi süper nokta haline getirmek amacıyla denetimli bir yapı önermişler [71].

Kang vd., boyut sayısı büyük bir uzayda yerel geometrik ilişkileri doğru bir şekilde yakalamak için grafik gömme modülü (GEM/ Graph Embedding Module) ve piramit dikkat ağı (PAN/ Pyramid Attention Network) temelli PyramNet’i önerdiler [71]. GEM modülü, bir NB yönlendirilmiş bir döngüsel olmayan grafik olarak formüle eder ve bitişik benzerlik matrisi için Öklid mesafesi yerine kovaryans matrisi kullanılır. PAN modülünde ise farklı anlam yoğunluklarına göre özelliklerin çıkarılması için dört farklı boyuta sahip evrişim çekirdekleri kullanılır. Wang vd., komşu yerel bir kümeden ilgili özellikleri seçici olarak öğrenmek için Grafik Dikkat Evrişimi (GAC) önermiştir. Bu işlem, uzaydaki konumlarına, özniteliklerine, komşu noktalarına ve boyutlarına dikkat ağırlıkları atanarak yapılmıştır [72]. GAC, bölümlene için ayırıcı öznitelikleri almak için eğitilebilir ve genel CRF modeline benzeyen özniteliklere sahiptir. Ma vd., grafik temsili ile yakın özellikleri belirlemek için PointGCR modülü önermiştir [73]. PointGCR, baştan sona eğitilebilir bir yapıdır. Performans iyileştirmesi için herhangi bir segmentasyon ağıyla kolayca birleştirilebilir.

D.5. Modellerin Değerlendirilmesi

Projeksiyon tabanlı ve ayrıklaştırma tabanlı yöntemler düzenli temsilleri sayesinde 2B görüntü işlemede kullanılan olgun ağ mimarilerinden yararlanabilir. Projeksiyon tabanlı yöntemlerde 3B’tan 2B’ye yansıtılırken bilgi kaybında neden olmaktadır. Ayrıklaştırma tabanlı yöntemlerde ise çözünürlüğün artması sonucu kübik olarak artan hesaplama ve bellek maliyetine neden olmuştur. Bu amaçla, indeklemeye yapıları üzerine inşa edilen seyrek evrişim, uygulanması etkili bir çözüm noktası olabilir.

Nokta tabanlı ağlar en sık araştırılan yöntemlerdir. Noktasal temsiller doğal olarak açık komşu bilgisine sahip değildir ve güncel olarak kullanılan nokta tabanlı yöntemlerin çoğu maliyeti yüksek komşu arama algoritması kullanmaktadır. (örneğin, KNN [74] veya top sorgusu [54]). Bu yöntemler doğal olarak bu yöntemlerin etkinliğini sınırlamaktadır. Nokta bulutunu yapısız verilerden öğrenmek hala zor bir problemdir. Bazı yaklaşımlar [65], [70], [58] dikkate değer genel bir performansa göstermiş olmasına rağmen bazılarının performansı hala sınırlıdır. 3B sahne ve nesne segmentasyonun performans karşılaştırması Tablo 3'te verilmiştir.

Tablo 3: SemanticKITTI, S3DIS, ScanNet ve Semantic3D semantik segmentasyon başarı tablosu. Genel Doğruluk (Overall Accuracy/OA), Birlik Üzerinden Ortalama Kesişim (Mean Intersection over Union /mIoU) değerlendirme sonuçları % olarak verilmiş ve “-” veri setinde kullanılmadığı anlamına gelmektedir.

Tablo 3. Modellerin farklı veri Setleri üzerindeki performans başarıları.

Yöntem	Ortalama kesişim	Genel Doğruluk	Yöntem	Ortalama kesişim	Genel Doğruluk	
SGPNet	85.8	-	DGCNN	85.1	-	ShapeNet
KD-Net	82.3	-	RGCNN	84.3	-	
SEGCloud	79.4	-	RSNet	84.9	-	
PointNet	83.7	-	SO-Net	84.6	-	
Yi	81.4	-	Spider-CNN	85.3	-	
O-CNN+CRF	85.9	-	LDGCNN	85.1	-	
SSCNN	84.7	-	GAPNet	84.7	-	
PN++(MSG+DP)	85.1	-	-	-	-	
SEGCloud	48.92	-	3DMAXNet	47.5	79.5	S3dis
PointNet	47.71	78.62	SP-GRAPH	62.1	85.5	
MS+CU	47.8	79.2	3R-RNN	56.3	86.9	
G+RCU	49.7	81.1	PointCNN	62.74	88.1	
DGCNN	56.1	56.1	PointSHIFT	70.23	88.72	
RSNet	53.83	-	ASIS	59.3	86.2	
SGPN	50.37	80.78	A-CNN	-	87.3	
SEGCloud	-	73.0	SEGCloud	61.3	88.1	ScanNet
PointNet	14.69	73.9	TMLC-MSR	54.2	86.2	
PN++(SSG)	-	83.3	DeepPr3SS	58.85	88.9	

PN++(MSG+DP)	34.26	-	SnapNet	59.1	88.6	
RSNet	39.35	84.5	SPGraph	73.2	94.0	
PointCNN	-	83.4	-	-	-	
PointSHIFT	-	86.2	-	-	-	
PointNet	34.4	79.7	-	-	-	Kıttı
3P-RNN	41.6	87.8	-	-	-	
G+RCU	36.2	80.6	-	-	-	

Tablo 1 ve Tablo 2’de veri setleri analiz edildiğinde ve son zamanlarda yapılan çalışmalar değerlendirildiğinde Tablo 3 ‘e göre noktasal tabanlı derin öğrenme yöntemleri hızlı bir şekilde artış göstermiştir.

IV. SONUC

Bu makalede, 3B sahnelerin veya nesnelerin semantik segmentasyon yöntemlerini anlamak için güncel yöntemler derinlemesine araştırılmıştır. Araştırılan yöntemlerin avantajları, dezavantajları ve bu yöntemlerin performans sonuçları değerlendirilmiştir. 3B NB’leri başarılı bir şekilde bölümlere ayırmak için farklı derin öğrenme yöntemleri geliştirilmiş ve NB işleme aşamasındaki zorluklara rağmen gelişime açık olduğu gösterilmiştir. Yeni algoritmaların gelişimine katkıda bulunmak için 3B NB segmentasyon yöntemlerinin en popüler 3B veri setleri üzerinde başarı performansları karşılaştırılmıştır.

3B NB’lerin başarılı bir şekilde eğiterek semantik segmentasyon yapılabilir. Tüm anlatılan modellerin hedeflerini bölge büyütme, başarılı model geliştirme, matematiksel modeller üretme ve geometrik teknikler olmak üzere dört başlık altında geliştirilmiştir.

Derin öğrenme yöntemleri ile 3B görüntü işlemede karşılaşılan problemleri çözmek için NB öznelikleri modeller üzerinde denenmiştir. Modellerin başarılı sonuçlar vermesiyle tüm dikkatleri bu yöntemler üzerine çevirmiş ve NB’yi odak noktası haline getirmiştir. Bu nedenle kıyaslama veri setleri üzerinde eğitilen derin öğrenme modellerinin başarıları karşılaştırılmıştır. Bu modeller detaylı olarak araştırılmış, her yöntemin avantajları, dezavantajları ve uygulama farklılıkları tartışılmıştır. Derin öğrenme yöntemlerinde 3B NB kullanılarak semantik segmentasyona genel bir bakış yapılmış ve değerli bilgiler verilmiştir. Özellikle NB ile semantik segmentasyon alanında yapılan çalışmaların önemli yönleri ve yöntemleri sunulmuştur. Bu makalede araştırmalar sonucunda geliştirilen yöntemlere öneri olarak gelecekte karşılaşılabilecek bazı problemler tespit edilmiştir. Bunlar; 3B’den 2B geçerken önemli derecede bilgi kaybının olması, nokta tabanlı ağların maliyetli komşu arama yöntemleri ve büyük ölçekli düzensiz nokta bulutlarını işleme problemleri daha fazla araştırılabilir.

Yazarların Katkısı

3B nokta bulutu alanında literatürü zenginleştirmek ve yeni araştırmacılar için kapsamlı bir alan taraması yapılmıştır. Özellikle bu alanda kaynak sağlaması ve gelecekte 3B nokta bulutu alanında yapılacak çalışmalara yol gösterici nitelikte bir makale sunulmuştur. Zengin bir kaynakça içeriğine sahip makalemizde bu önemli konu hakkında değerli bilgiler sağlayacağına ve yeni araştırmalar için yol gösterici bir araştırma olacağına inanıyoruz.

V. KAYNAKLAR

- [1] Z. Liang vd., “Stereo matching using multi-level cost volume and multi-scale feature constancy”, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no 1, pp. 300–315, 2021.
- [2] Y. Guo, F. Sohel, M. Bennamoun, M. Lu, ve J. Wan, “Rotational projection statistics for 3D local surface description and object recognition”, *Int. J. Comput. Vis.*, vol. 105, no 1, pp. 63–86, 2013.
- [3] C. Xiaozhi, M. Huimin, W. Ji, L. Bo, ve X. Tian, “Multi-View 3D Object Detection Network for Autonomous Driving | Spotlight 4-2B - YouTube”, *Comput. Videos*, pp. 1907–1915, 2017.
- [4] R. B. Rusu ve S. Cousins, “3D is here: Point Cloud Library (PCL)”, *Proc. - IEEE Int. Conf. Robot. Autom.*, pp. 1–4, 2011.
- [5] A. Shamir, “Segmentation and shape extraction of 3D boundary meshes”, *Eurographics*, no September, pp. 137–149, 2006.
- [6] R. B. Rusu, Z. C. Marton, N. Blodow, M. Dolha, ve M. Beetz, “Towards 3D Point cloud based object maps for household environments”, *Rob. Auton. Syst.*, vol. 56, no 11, pp. 927–941, 2008.
- [7] P. J. Besl ve R. C. Jain, “Segmentation Through Variable-Order Surface Fitting”, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no 2, pp. 167–192, 1988.
- [8] S. Pu ve G. Vosselman, “Automatic extraction of building features from terrestrial laser scanning”, *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.*, no. 36, 2006.
- [9] R. B. Rusu, A. Holzbach, N. Blodow, ve M. Beetz, “Fast geometric point labeling using conditional random fields”, *2009 IEEE/RSJ Int. Conf. Intell. Robot. Syst. IROS 2009*, pp. 7–12, 2009.
- [10] J. Shi ve J. Malik, “Normalized cuts and image segmentation”, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no 8, pp. 888–905, 2000.
- [11] R. Leahy, “An Optimal Graph Theoretic Approach to Data Clustering: Theory and Its Application to Image Segmentation”, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol 15, no11, pp. 1101–1113, 1993.
- [12] Y. Boykov ve G. Funka-Lea, “Graph cuts and efficient N-D image segmentation”, *Int. J. Comput. Vis.*, vol. 70, no 2, pp. 109–131, 2006.
- [13] D. Anguelov vd., “Discriminative learning of Markov random fields for segmentation of 3D scan data”, *Proc. - 2005 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognition, CVPR 2005*, vol. II, pp. 169–176, 2005.
- [14] Y. Zhu vd., “Target-driven visual navigation in indoor scenes using deep reinforcement learning”, *Proc. - IEEE Int. Conf. Robot. Autom.*, vol 1, pp. 3357–3364, 2017.
- [15] I. Armeni vd., “3D Semantic Parsing of Large-Scale Indoor Spaces Supplementary Material”, *2016 IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 1534–1543, 2016.
- [16] A. X. Chang vd., “ShapeNet: An Information-Rich 3D Model Repository”, 2015.
- [17] J. Behley vd., “SemanticKITTI”, *Iccv*, vol iii, 2019.

- [18] A. Geiger, P. Lenz, ve R. Urtasun, “Are we ready for autonomous driving? the KITTI vision benchmark suite”, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., ss. 3354–3361, 2012.
- [19] A. Dai, A. X. Chang, M. Savva, M. Halber, T. Funkhouser, ve M. Nießner, “ScanNet: Richly-annotated 3D reconstructions of indoor scenes”, arXiv, 2017.
- [20] T. Hackel, N. Savinov, L. Ladicky, J. D. Wegner, K. Schindler, ve M. Pollefeys, “Semantic3D.Net: a New Large-Scale Point Cloud Classification Benchmark”, ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci., c. 4, sayı 1W1, ss. 91–98, 2017.
- [21] K. Mo vd., “PartNet: A Large-scale Benchmark for Fine-grained and Hierarchical Part-level 3D Object Understanding”, arXiv, pp. 909–918, 2018.
- [22] X. Song vd., “ApolloCar3D: A Large 3D Car Instance Understanding Benchmark for Autonomous Driving”, arXiv, ss. 5452–5462, 2018.
- [23] M. A. Uy, Q. H. Pham, B. S. Hua, T. Nguyen, ve S. K. Yeung, “Revisiting point cloud classification: A new benchmark dataset and classification model on real-world data”, Proc. IEEE Int. Conf. Comput. Vis., c. 2019-October, pp. 1588–1597, 2019.
- [24] Z. Wu vd., “3D ShapeNets: A deep representation for volumetric shapes”, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 07-12-June, pp. 1912–1920, 2015.
- [25] X. F. Han, H. Laga, ve M. Bennamoun, “Image-based 3D Object Reconstruction: State-of-the-Art and Trends in the Deep Learning Era”, arXiv, vol. 8828, 2019.
- [26] G. Elbaz, T. Avraham, ve A. Fischer, “3D point cloud registration for localization using a deep neural network auto-encoder”, Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017, vol. 2017-Janua, sayı July, pp. 2472–2481, 2017.
- [27] A. Nguyen ve B. Le, “3D point cloud segmentation: A survey”, IEEE Conf. Robot. Autom. Mechatronics, RAM - Proc., pp. 225–230, 2013.
- [28] Y. Ishikawa, R. Hachiuma, N. Ienaga, W. Kuno, Y. Sugiura, ve H. Saito, “Semantic Segmentation of 3D Point Cloud to Virtually Manipulate Real Living Space”, Proc. 2019 12th Asia Pacific Work. Mix. Augment. Reality, APMAR 2019, 2019.
- [29] B. Liu, S. He, D. He, Y. Zhang, ve M. Guizani, “A Spark-Based Parallel Fuzzy c -Means Segmentation Algorithm for Agricultural Image Big Data”, IEEE Access, vol. 7, pp. 42169–42180, 2019.
- [30] R. Schnabel, R. Wahl, ve R. Klein, “Efficient RANSAC for point-cloud shape detection”, Comput. Graph. Forum, vol. 26, no 2, pp. 214–226, 2007.
- [31] B. Yang ve Z. Dong, “A shape-based segmentation method for mobile laser scanning point clouds”, ISPRS J. Photogramm. Remote Sens., vol. 81, pp. 19–30, 2013.
- [32] J. Yan, J. Shan, ve W. Jiang, “A global optimization approach to roof segmentation from airborne lidar point clouds”, ISPRS J. Photogramm. Remote Sens., vol. 94, pp. 183–193, 2014.

- [33] A. Golovinskiy ve T. Funkhouser, “Min-cut based segmentation of point clouds”, 2009 IEEE 12th Int. Conf. Comput. Vis. Work. ICCV Work. 2009, pp. 39–46, 2009.
- [34] T. Rabbani, F. van den Wildenberg, ve G. Vosselman, “Segmentation of point clouds using smoothness constraint”, *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. 36, no 5, pp. 248–253, 2006.
- [35] A. Boulch, B. Le Saux, ve N. Audebert, “Unstructured point cloud semantic labeling using deep segmentation networks”, *Eurographics Work. 3D Object Retrieval, EG 3DOR*, c. 2017-April, pp. 17–24, 2017.
- [36] F. J. Lawin, M. Danelljan, P. Tosteberg, G. Bhat, F. S. Khan, ve M. Felsberg, “Deep projective 3D semantic segmentation”, *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10424 LNCS, pp. 95–107, 2017.
- [37] B. Wu, A. Wan, X. Yue, ve K. Keutzer, “SqueezeSeg: Convolutional neural nets with recurrent CRF for real-time road-object segmentation from 3D LiDAR point cloud”, *arXiv*, pp. 1887–1893, 2017.
- [38] B. Wu, X. Zhou, S. Zhao, X. Yue, ve K. Keutzer, “SqueezeSegV2: Improved model structure and unsupervised domain adaptation for road-object segmentation from a LiDAR point cloud”, *arXiv*, pp. 4376–4382, 2018.
- [39] A. Milioto, I. Vizzo, J. Behley, ve C. Stachniss, “RangeNet ++: Fast and Accurate LiDAR Semantic Segmentation”, *IEEE Int. Conf. Intell. Robot. Syst.*, vol i, pp. 4213–4220, 2019.
- [40] D. Rethage, J. Wald, J. Sturm, N. Navab, ve F. Tombari, “Fully-Convolutional Point Networks for Large-Scale Point Clouds”, *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11208 LNCS, pp. 625–640, 2018.
- [41] H. Y. Meng, L. Gao, Y. K. Lai, ve D. Manocha, “VV-NET: Voxel VAE net with group convolutions for point cloud segmentation”, *arXiv*, pp. 8500–8508, 2018.
- [42] B. Graham, M. Engelcke, ve L. Van Der Maaten, “3D semantic segmentation with submanifold sparse convolutional networks”, *arXiv*, pp. 9224–9232, 2017.
- [43] J. Kautz, “Supplementary Material for SPLATNet: Sparse Lattice Networks for Point Cloud Processing”, *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 2–4, 2018.
- [44] R. Alexandru Rosu, P. Schütt, J. Quenzel, ve S. Behnke, “LatticeNet: Fast Point Cloud Segmentation Using Permutohedral Lattices”, 2020.
- [45] Angela Dai ve M. Nießner, “3DMV: Joint 3D-multi-view prediction for 3D semantic scene segmentation”, *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11214 LNCS, pp. 458–474, 2018.
- [46] M. Jaritz, J. Gu, ve H. Su, “Multi-view pointnet for 3D scene understanding”, *arXiv*, 2019.
- [47] F. Engelmann, T. Kontogianni, B. Leibe, ve R. Field, “2020 IEEE International Conference on Robotics and Automation (ICRA) Dilated Point Convolutions : On the Receptive Field Size of Point Convolutions on 3D Point Clouds tions (DPC). In a thorough ablation study , we show that the ment of 3D scene unders”, vol 4, pp. 9463–9469, 2020.

- [48] M. Tatarchenko, J. Park, V. Koltun, ve Q. Y. Zhou, “Tangent Convolutions for Dense Prediction in 3D”, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 3887–3896, 2018.
- [49] L. Tchapmi, C. Choy, I. Armeni, J. Gwak, ve S. Savarese, “SEGCloud: Semantic segmentation of 3D point clouds”, Proc. - 2017 Int. Conf. 3D Vision, 3DV 2017, pp. 537–547, 2018.
- [50] J. Long, E. Shelhamer, T. Darrell, ve U. Berkeley, “Fully Convolutional Networks for Semantic Segmentation”, Proc. - 2019 Int. Conf. Comput. Vis. Work. ICCVW 2019, pp. 847–856, 2019.
- [51] A. Dai, D. Ritchie, M. Bokeloh, S. Reed, J. Sturm, ve M. Niebner, “ScanComplete: Large-Scale Scene Completion and Semantic Segmentation for 3D Scans”, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 4578–4587, 2018.
- [52] C. Choy, J. Y. Gwak, ve S. Savarese, “4D spatio-temporal ConvNets: Minkowski convolutional neural networks”, arXiv, 2019.
- [53] L. J. (2017). Qi, C. R., Su, H., Mo, K., & Guibas, “PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation”, Proc. - 2016 4th Int. Conf. 3D Vision, 3DV 2016, pp. 601–610, 2016.
- [54] C. R. Qi, L. Yi, H. Su, ve L. J. Guibas, “PointNet++: Deep hierarchical feature learning on point sets in a metric space”, arXiv, vol Nips, 2017.
- [55] M. Jiang, Y. Wu, T. Zhao, Z. Zhao, ve C. Lu, “PointSIFT: A SIFT-like Network Module for 3D point cloud semantic segmentation”, arXiv, 2018.
- [56] F. Engelmann, T. Kontogianni, J. Schult, ve B. Leibe, “Know what your neighbors do: 3D semantic segmentation of point clouds”, arXiv, 2018.
- [57] H. Zhao, L. Jiang, ve C. F. Jiaya, “PointWeb: Enhancing Local Neighborhood Features for Point Cloud Processing”, vol. 1, pp. 5565–5573.
- [58] Z. Zhang, B. S. Hua, ve S. K. Yeung, “ShellNet: Efficient point cloud convolutional neural networks using concentric shells statistics”, arXiv, pp. 1607–1616, 2019.
- [59] Q. Hu vd., “RandLA-Net: Efficient semantic segmentation of large-scale point clouds”, arXiv, pp. 11108–11117, 2019.
- [60] Y. Zhao, H. Deng, F. Tombari, T. Universit, ve S. Ag, “3D Point Capsule Networks Supplementary Material”, Cvpr, pp. 2–6, 2019.
- [61] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, ve J. M. Solomon, “Dynamic Graph CNN for Learning on Point Clouds”, ACM Trans. Graph., vol. 38, no 5, s. Article 146, 2019.
- [62] R. Arandjelovic, P. Gronat, A. Torii, T. Pajdla, ve J. Sivic, “NetVLAD: CNN Architecture for Weakly Supervised Place Recognition”, IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no 6, pp. 1437–1451, 2018.
- [63] B. S. Hua, M. K. Tran, ve S. K. Yeung, “Pointwise convolutional neural networks”, arXiv, pp. 984–993, 2017.

- [64] S. Wang, S. Suo, W. C. Ma, A. Pokrovsky, ve R. Urtasun, “Deep Parametric Continuous Convolutional Neural Networks”, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 2589–2597, 2018.
- [65] H. Thomas, C. R. Qi, J. E. Deschaud, B. Marcotegui, F. Goulette, ve L. J. Guibas, “KPConv: Flexible and deformable convolution for point clouds”, arXiv, pp. 6411–6420, 2019.
- [66] F. Engelmann, T. Kontogianni, A. Hermans, ve B. Leibe, “Exploring spatial context for 3D semantic segmentation of point clouds”, arXiv, 2018.
- [67] W. Wang, “Supplementary : Recurrent Slice Networks for 3D Segmentation on Point Clouds Qianguai Huang”, vol 1, pp. 2015–2016, 2017.
- [68] Z. Zhao, M. Liu, ve K. Ramani, “DAR-Net: Dynamic aggregation network for semantic scene segmentation”, arXiv, 2019.
- [69] F. Liu, S. Li, L. Zhang, ve C. Zhou, “3DCNN-DQN-RNN : A Deep Reinforcement Learning Framework for Semantic”, IEE Int. Conf. Comput. Vision (ICCV) 2017, vol July, pp. 5678–5687, 2017.
- [70] Loic Landrieu¹ ve M. Simonovsky, “Large-scale Point Cloud Semantic Segmentation with Superpoint Graphs”, J. Exp. Theor. Phys., vol. 89, no 4, pp. 734–739, 2018.
- [71] L. Landrieu ve M. Boussaha, “Point cloud oversegmentation with graph-structured deep metric learning”, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2019-June, pp. 7432–7441, 2019.
- [72] L. Wang, Y. Huang, Y. Hou, S. Zhang, ve J. Shan, “Graph attention convolution for point cloud semantic segmentation”, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2019-June, pp. 10288–10297, 2019.
- [73] Y. Ma, Y. Guo, H. Liu, Y. Lei, ve G. Wen, “Global context reasoning for semantic segmentation of 3D point clouds”, Proc. - 2020 IEEE Winter Conf. Appl. Comput. Vision, WACV 2020, pp. 2920–2929, 2020.
- [74] Y. Li, R. Bu, ve X. Di, “PointCNN : Convolution On X -Transformed Points”, sayı NeurIPS, 2018.