



Derleme Makalesi/Review Article

Derin öğrenme ile 3 boyutlu nokta bulutlarının sınıflandırılmasına genel bir bakış

*Review of the classification of 3 dimensional point clouds with deep learning***Muhammed Ahmet DEMİRTAŞ¹**¹Muş Alparslan University, Muş, Turkey, ma.demirtas@alparslan.edu.tr, ORCID: 0000-0003-4092-7284¹Kocaeli University, Kocaeli, Turkey, ahmet.demirtas@kocaeli.edu.tr

MAKALE BİLGİLERİ

*Makale Geçmişi:*Geliş 3 Şubat 2022
Revizyon 23 Mart 2022
Kabul 25 Mart 2022
Online 30 Mart 2022*Anahtar Kelimeler:**Nokta bulutu, derin öğrenme, 3B nesne sınıflandırma*

ÖZ

Nokta bulutu (NB) bir vektör uzayında nesneye ait bilgilerin x,y,z koordinat sisteminde matematiksel olarak temsil edilen noktalar kümesidir. Noktalar kaydedilen uzaysal koordinat sisteminde sınıflandırılarak nesne ya da bir alanı ifade eden anlamsal bilgileri tanımlarlar. 3 boyutlu (3B) nokta bulutu gelişen teknolojilerle beraber nesnelere sınıflandırma, algılama ve tanıma alanlarında son zamanlarda oldukça popüler bir hale gelmiştir. Lazer tarama sistemleri ile taranan nesnelere 3B nokta bulutuna dönüştürülmüş ve bu verilerin sanal ortama aktarılması ile farklı veri setleri elde edilmiştir. 3B noktasal verileri derin ağlar ile başarılı bir şekilde sınıflandırmak için geliştirilen yöntemler detaylı olarak incelenmiştir. Nokta bulutlarına 3B koordinat sistemi ile birlikte farklı bilgiler dâhil ederek (derinlik ya da RGB (red-green-blue) farklı boyutlarda veya farklı yoğunlukta nokta bulutları oluşturulmuştur. Ayrıca nokta bulutu veri kümesindeki her bir noktaya ait; harici veya dâhili bilgiler eklenmiş RGB değerleri ile nesnelere renklendirilmiştir. Bu çalışmada 3B nokta bulutunu derin ağlar ile sınıflandıran yöntemlerin başarı performansları, avantajları, dezavantajları analiz edilmiştir. Özellikle uygulanan algoritmalar, denenmiş yöntemler ve oluşturulan modeller karşılaştırılmış ve tartışılmıştır. Son olarak makalede gelecekteki çalışmalara hız ve yön vermesi için güncel yöntemler kapsamlı bir şekilde sunulmuştur.

ARTICLE INFO

*Article history:*Received 3 February 2022
Received in revised form 23 March 2022
Accepted 25 March 2022
Available online 30 March 2022*Keywords:**point cloud, deep learning, 3D object classification*

ABSTRACT

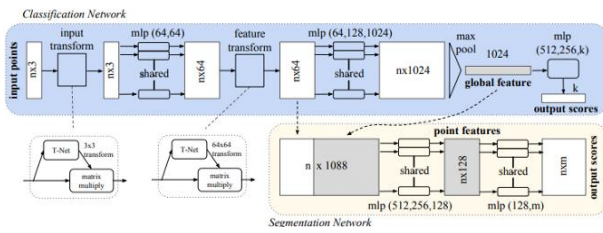
A point cloud is a set of points mathematically portrayed in a vectorial space or in the x, y, z coordinate plane of the object. Points are classified in the registered spatial coordinate plane and define semantic information that represents an object or an area. As technology develops, 3D point clouds have recently become eminently prominent in the range of classification, detection, and recognition of objects. Different data sets are attained by transforming the scanned objects into 3D point clouds and transferring them to the virtual environment with lidar scanning systems. Classification methods are developed to successfully classify 3D point data which is analyzed in detail. Dense point clouds of different sizes are created by encompassing distinctive information (depth or RGB) together with the 3D coordinate plane to the point clouds. In addition, for each point in the point cloud dataset; external or internal information has been supplemented and objects are colored with RGB values. In this paper, the success performances, advantages and disadvantages of the methods classifying the 3D point cloud with deep networks have been analyzed. In particular, the applied algorithms, the tried methods, and the created models are compared and discussed. As a result, the existing methods for giving speed and direction to future work are presented in a comprehensive manner.

Doi: 10.24012/dumf.1067736

* Sorumlu Yazar

Giriş

Derin öğrenme (DÖ) geçmiş yıllarda çok çeşitli bilgisayarlı görü görevlerinde, özellikle görüntü alanında oldukça yaygın ve başarılı bir performans göstermiştir. Bu sayede hava ve karasal lazer tarayıcılar, otonom araçlar ile toplanan nokta bulutu (NB) verileriyle daha net algılanan nesnelere elde edilmiştir. NB nesnelere ait bilgileri x,y,z koordinat düzleminde sayısal olarak ifade eden ve binlerce noktadan oluşan 3B dijital bir temsildir. Şekil 1.2'de gerçek hayatta kullanılan nesnelere dijital temsilleri verilmiştir. Ek olarak Lidar (Light Detection and Ranging) veya RGB-D kameralardan gelen 3B nokta bulutlarından 3B geometrik veriler elde edilmiştir. Bu verilerin elde edilmesi ile nokta bulutlarının DÖ modelleri ile kullanılmasını popüler bir araştırma noktası haline getirmiştir. NB doğrudan mekânsal yapılar, araçlar, yayalar, trafik işaretleri ve şehir planlaması gibi birçok alanda kullanılmıştır. Kullanım alanının genişlemesi bu alanda kullanılan DÖ yöntemlerinin NB işleme aşamalarında karşılaştığı problemleri genelleme yoluna gidilebilir bunlar; seyrek noktalar ile doğru bilgilerin nasıl temsil edileceği, kaynak kullanımı, büyük hacme sahip verilerin işlenmesi ve verilerin doğru bir şekilde segmentasyon ya da sınıflandırma yapılmasıdır. NB işlemek için ilk DÖ ağı PointNet kullanılmıştır. PointNet 3B NB işlemek için yeni bir DÖ ağı tasarlamış ve yaygın kullanılan bir model haline gelmiştir. Ayrıca PointNet çok katmanlı algılayıcılar (MLP) ve T-Net'ten yararlanarak yapısız 3B NB işlemek için ayrı bir ağ tasarlamıştır. PointNet T-Net'i değişmezliği korumak için kullanmış ve tahmin edilen dönüşüm matrisi ile geometrik dönüşümleri çarparak bu değişmezliği korumuştur. Sonrasında NB MLP ve maksimum havuz katmanından geçirilerek sınıflandırma yapılmıştır.



Şekil 1. PointNet mimarisi [1].

NB verileri düzensiz bir formatta olduğu için 3B voksel ızgaralarına dönüştürülür. Bu dönüşümde gereksiz yer kaplayan verilerin oluşması bazı sorunlara neden olmuştur. Bu sorunların üstesinden gelmek için PointNet nesne sınıflandırma, parça bölütleme, anlamsal sahne çözümlemeye kadar çeşitli uygulamalar için birleşik bir mimari sunmuştur [1]. PointNet modelinin şekil 1,1'de gösterilen bu temel yapısı sınıflandırma ve segmentasyon tahminlerinde yaygın olarak kullanılmaktadır. Sınıflandırma yapılırken genel özellikler, çıktı sonuçları için bir MLP'den geçirilmiştir. Segmentasyonda her noktanın sınıflandırılması ile küresel özelliklerin ve her noktadan farklı düzeylerde ara özelliklerin birleştirilmesi sonucu MLP'den geçirilmiştir. Bu adımlar güncel CNN

modellerinde ise nokta bulutlarını bir dizi evrişimli katmalardan geçirip farklı düzeylerdeki özelliklerini almıştır. Bu özelliklerden yola çıkarak yeni bir model PointNet ++ önerilmiştir. PointNet basit olmasına rağmen verimliliği yüksek ve etkili bir yöntemdir. PointNet temel alınarak elde edilen PointNet ++ [2] sürümü geliştirilmiş ve bu sürümler gelecekte önerilecek birçok çalışmaya referans sağlamıştır.

3B NB işleme üç ana başlık altında incelenmiştir: Nesne tanıma, kümeleme ve sınıflandırma adımlarıdır. Bu yöntemler ve algoritmalar gerçek hayatta farklı alanlarda kullanılan nesnelere, iç ve dış mekânları; tanıma sınıflandırma ve kümeleme üzerinde önemli bir yere sahip olmuştur.



Şekil 2. Nesnelere nokta bulutu ile temsil edilmesi [1].

Gerçek hayatta güvenlik analizleri ve trafik gürültüsü seviyeleri hakkında güncel bilgiler sağlama için, çok sayıda özel yazılımlarda artışa neden olmuştur [3].

Güncel modellerden ShapeNet ve ModelNet, düzenli 3B modellerden oluşan veri kümeleridir [4] [5]. Bununla birlikte, sağlam modeller geliştirmek amacıyla gürültülü ve eksik noktalar dikkate alınarak ScanNet ve KITTI gibi veri kümeleri gerçek ortam taramalarından elde edilmiştir [6] [7]. Bunun yanında NuScenes ve Lyft gibi otonom araç görevleri için tasarlanmış veri kümelerine farklı zaman ve ortamlar dâhil edilerek oluşturulmuştur [8].

Tüm bunları özetlemek gerekirse PointNet ve PointNet ++ uygulamaları gelecekte kullanılacak birçok çalışmaya temel olmuştur. Örneğin PointNet ++ 'dan PointWeb [9] SRN (Yapısal İlişki Ağı) [10] yöntemleri önerilmiştir. SRN, PointNet ++ donatmak için önerilmiş ve sonuç olarak PointNet ++'tan daha iyi performans elde etmiştir.

3 Boyutlu veri kümeleri oluşturma

Günümüzde 3B boyutlu Lidar ve Microsoft Kinect gibi tarayıcıların yaygın olarak kullanılması ile 3B NB verileri kolaylıkla oluşturulmaktadır. NB yapısız ve yoğun olması noktaları işleyecek bir model veya bir NB kütüphanesine ihtiyaç olduğunu göstermiştir. NB Kütüphanesi (Point Cloud Library (PCL)) [11] 3B nesnelere sınıflandırmak ve algılamak için en güncel algoritmaları içermektedir.

3B NB işleme aşamalarında toplanan nokta bulutlarının ön ve arka plana ayrılması gerekmektedir. Bu aşamalarda farklı NB'leri segmentlere ayırma, benzer özelliklere sahip noktaları ise anlamlı ve homojen olarak kümelemeyi hedeflemiştir. Segmentasyon, bir alan ya da mekândaki nesnelere bulma, tanıma, sınıflandırma ve özelliklerini çıkarma gibi çeşitli yönlerden analiz etmek

için yardımcı olmuştur [12]. 3B NB'yi anlamlı homojen kümelere ayırmak için girdi açısından grafik oluşturma, sınırları belirleme, pürüzsüzlük veya içbükey bilgileri ile bölümlenme yapmaktadır.

Shamir, bölümlenme tekniklerini sınır ağları üzerindeki segmentasyon problemini, bir optimizasyon problemi olarak formüleştirmiştir [13]. İki farklı ağ segmentasyonu tanımlamış: Bunlar parça ve yama segmentasyonudur. Segmentasyon, optimizasyon kriterlerine, farklı özellik ve tekniklere göre sınıflandırılmıştır. Segmentasyon, temel teknikler için genel algoritmalar, geometrik şekil işleme, anlamsal temelli şekil temsilini geliştirmek, 3B ağ ile şekil özelliklerini ve yapısını çıkarmıştır. Bu algoritmalar yaklaşım türlerine göre sınıflandırmış ve genel kümeleme algoritmasını kullanarak bağlantı sağlamıştır. Bu yaklaşımlar yaygın olarak, bölge tabanlı yöntemlerde özellikle NB işlerken noktaları bölümlere ayırmak ve geometrik ağ oluşturmak için kullanılmıştır [13, 14]. Bu yaklaşımların temel amacı 3B NB'yi bölümlere ayırmak olmuştur. Xu vd., göre 3B NB segmentasyon algoritmasının üç önemli özelliği vardır [15]. İlk olarak algoritma, nesnelere birbirinden ayırt etmeli ve niteliksel birçok farklı özelliklerden ayarlanabilmelidir. Özelliklerin sayısı arttıkça, algoritma bu özellikleri otomatik olarak nasıl değiştireceğini öğrenmelidir. İkincisi, bölümlenme algoritması, komşuluk bilgilerine göre seyrek bölgelerdeki noktaların etiketini çıkarmalıdır. Üçüncüsü, segmentasyon algoritması belirli 3B tarayıcıya uygulanmalıdır çünkü farklı lazer tarayıcılar farklı özelliklerde NB üretir ve aynı sahne bile farklı özelliklere sahip olabilirler. Sonuç olarak nesnelere ait 3B NB, birçok parçanın birleştirilmesi ve ön-arka planın birbirinden ayrılması ile nesnenin 3B verisi elde edilmiştir.

Yaygın kullanılan veri kümeleri

Bu bölümde, 3B nesnelere sınıflandırma, bölümlenme ve nesne algılama-tanıma için en sık kullanılan 3B NB veri kümeleri sunulmuştur. Açık kaynak veri kümelerinin kullanılma nedeni, farklı yaklaşımların performanslarını değerlendirmek, avantajlarını, dezavantajlarını belirtmek, uygulanan modelin yapısı hakkında bilgi almak ve önerilecek yeni yaklaşımlara kaynak olmasıdır.

ShapeNet 3B ham model verileri, halka açık çevrimiçi havuzlardan ve mevcut araştırma veri kümelerinden toplamıştır. Popüler ve herkese açık Trimble 3B ve Yobi3 iki depodaki 3B çokgen modellerden toplanmıştır. Trimble 3B Depo 2,4 milyon kullanıcı tarafından tasarlanmış, 3B modeller ve sahnelerden oluşmuştur. Yobi3D, çok çeşitli diğer çevrimiçi depolardan toplanan 350 bin ek modelden oluşmaktadır. Bu veri kümesi çok geniş bir nesne sınıfı, sahne kategorisi ve çok çeşitli şekillerden meydana gelmiştir. ShapeNet, 55 kategoride 51300 3B model içeren geliştirilebilir halka açık bir veri kümesidir [16]. Ek olarak ShapeNet farklı alt kümelere sahiptir. ShapeNetSem, 270 kategoriye ve 12000 model sahip ShapeNet alt kümesinden oluşturulmuştur. Bu düzenli ve küçük veri kümesi temel modelin kapasitesini değerlendirmek için kullanılmıştır. ShapeNetCore, tek düzenli 3B modeller ve manuel olarak doğrulanan

kategori ve hizalama notları içeren ShapeNet veri kümesinin bir alt kümesidir. Bu veri seti tahmini 51.300 3B model ve 55 ortak obje grubundan oluşmaktadır. ShapeNetCore PASCAL 3D + [17] veri setine ait 12 nesne grubunu da içermektedir.

ModelNet projesi üç bölüm sunmaktadır: ModelNet10, ModelNet40 ve Aligned40. ModelNet40 karşılaştırması en yaygın kullanılan modellerdir [16]. En yaygın nesne kategorilerini bulmak için Sun [18] veri tabanından elde edilen istatistiklerden yararlanmıştır.

Stanford Büyük Ölçekli 3B Kapalı Alanlar (S3DIS) veri kümesi, mimari tarz ve görünüm açısından çeşitliliği korumak için üç bina ve beş büyük ölçekli iç mekân sahneden oluşur. NB otomatik oluşturulmuş ve yapısal elemanlar (zemin, duvar) ortak mobilyalar dâhil olmak üzere 12 anlamsal nesne tespit edilmiş. Anlamsal bölümlenme için S3DIS [19], 6.000 m²'nin üzerinde, 215 milyon noktaya 271 odaya sahip birkaç binadan oluşan yeni bir veri kümesi oluşturulmuştur.

Semantic3D [20] dış mekân sahne bölümlendirmesi için lidar ile 110000 m² alandan ve 4 milyar noktadan oluşmaktadır. Bu veri kümesi şimdiye kadar kullanılan en büyük 3B NB veri kümesidir.

ScanNet, 3B kamera ile 1000'den fazla tarama yapılmış 1513 sahneden 2,5 M(milyon) kareden oluşan bir görünüme sahiptir. ScanNet kamera pozları, yüzey yapısı ve örnek düzeyinde semantik segmentasyon ile açıklanmış bir RGB-D video veri kümesidir [6]. Bu veri kümesi, sınıflandırma, semantik voksel etiketleme ve CAD (computer aided design) modeli gibi çok sayıda 3B sahne anlama görevleri içinde kullanılmaktadır.

Kitti veri seti çoklu görüntü verileri, yüksek çözünürlüğe sahip renkli ve gri kamera, Velodyne lazer tarayıcı ve bir GPS yerleştirme sistemi içeren otonom bir sürüş platformu ile arabalar, yayalar ve bisikletlilerden oluşmaktadır. Bu veriler nesne tespiti/takibi ve sahne akış tahminleri yapma imkanı sunulmuştur [22, 23].

Cornell RGBD veri kümesi iç mekân sahnelerinden oluşmuş; 24 ofis ve 28 ev olmak üzere toplamda 52 etiketli iç mekân sahneleri RGB değerlerine sahiptir. NB verileri, RGB-D SLAM kullanılarak [24] orijinal RGB-D görüntülerinden oluşturulmuştur. Veri kümesi, 27 nesne sınıfıyla etiketlenmiş 2495 segmente sahip olup yaklaşık olarak da 550 görünümünden meydana getirilmiştir.

VMR-Oakland veri kümesi, CMU kampüsü etrafında hareket ederek, 360 derecelik görüş alanına sahip bir karasal lazer tarama aracı ile toplanan noktaların metin tabanlı kaydedilmesi sonucu oluşturulmuştur. Her bir satıra her bir noktanın üç gerçek koordinat değeri(x,y,z) yazılmıştır [25].

Tablo 1. Yaygın kullanılan 3B NB veri kümeleri.

Tüm bu veri kümeleriyle beraber araştırmaları ve yapılacak pratik uygulamaları kısa sürede değerlendirmek için farklı Oakland [26], iQmulus [27] gibi birçok kaliteli fakat yaygın kullanılmayan veri kümeleri bulunmaktadır.

3 Boyutlu Sınıflandırma modelleri

3B nokta bulutları üç boyutlu koordinat sisteminde(x,y,z) sayısal temsil edilen noktaların geometrik özellikleri, yerel komşuluk değerlerine göre sınıflandırılmaktadır. 3B NB sınıflandırma modellerinde genellikle farklı formatta kayıt edilen verilerin sınıflandırılması yapılmaktadır. Aşağıda belirtilen tabloda 3B nokta bulutları ve özellikleri verilmiştir.

3B nesnelere 3B NB modelleri ile test edilmiştir. NB yöntemleri, sınıflandırma yöntemlerine bağlı olarak iki sınıfta incelenmiştir: projeksiyon tabanlı yöntemler ve nokta tabanlı yöntemlerdir.

Projeksiyon tabanlı yöntemler

Genellikle düzensiz bir yapıya sahiptir ve belirlenen bir yöntemle 3B NB özelliklerini çıkarmaktadır. Çoklu görünüm ve hacimsel tabanlı olarak ikiye ayrılabilir.

Çoklu görünüm yöntemi

Görünüme dayalı yöntemler, 3B nesne tanıma işlemleri üzerinde önemli bir başarıya ulaşmıştır. Su vd., nokta bulutlarının çoklu gösterimine dayalı MvCNN (Multi-View CNN) yöntemini sunmuştur [28]. Bu yöntem 3B NB'yi 2B farklı açılardan elde edilmiş anlık görüntülerden elde etmiştir. Görünüm arasında yoğun bir tanımlayıcı oluşturmak için görüntü özellikleri başka bir CNN ağından geçirilmiştir. MvCNN, nesnelere farklı görüntüleri arasındaki ilişkiler dikkate alınarak ayrılmıştır. Görünüm farklarını nicel olarak eşleştirilip fark değerlerine göre görünüm kümelerini gruplamıştır. Ayrıca görünüm tanımlayıcıların tahmin yapabilmesi için grup içi havuzlama ve gruplar arası yoğunluk değerleri alınmıştır. 3B nesnelere temsil ederken farklı görüntüler arasındaki ilişkilerin göz ardı edilmesi problemleri meydana getirmiştir. Problemleri çözmek için karşılıklı bölgeleri değişik yönlerden bağlamak ve bireysel görüntü bilgilerinin güçlendirilmesi gerekmektedir. Yang ve Wang bu probleme karşı bir ilişki Ağı (Relation Network) önermişlerdir [29].

Yu vd., görünüm tabanlı özellikleri birleştiren yöntemlerin, diğer yöntemlerden farkı, parçalar arası benzerlik olduğunu iddia etmişlerdir. Polinom çekirdeği ve çift doğrusal (bilinear pooling) havuzlama arasındaki ilişkiden yararlanarak [30] çift doğrusal havuzlama yoluyla yerel evrişimli özelliklerini bir araya getirip etkili bir 3B nesne gösterimi elde etmiştir.

Hacimsel tabanlı gösterim

Hacimsel gösterim yöntemleri; kapalı alana karşılık gelen hacimler, Öklid mesafesi ve hacimsel alanlar voksel haline getirilerek gösterilmiştir. Özellikle giriş değerleri nesnelere iç kısımlarına ait geometrik bilgileri elde edecek şekilde hedeflenmiştir. Hacimsel yöntemler genellikle girdi değerlerini voksel haline getirmektedir.

Veri Kümesi	Yapılan işlem	Sınıf	Veri Ölçeği
ShapeNet	Sınıflandırma	55	51300 model
ModelNet40	Sınıflandırma	44	12311 model
S3DIS	Segmentasyon	12	215 milyon nokta
Semantic3D	Segmentasyon	8	4 milyar nokta
ScanNet	Segmentasyon	20	2.5 milyon çerçeve
KITTI	Algılama-takip	3	802.56 nesne
Cornell RGBD	Rgb-d Sem. Seg	17	52 sahne
VMR-Oakland	Segmentasyon	27	550 görünüm.

VoxNet hacimsel gösterimi kullanan ilk NB işleyen yöntemdir. VoxelNet hacimsel gösterime sahip bir doluluk ızgara temsillerini denetimli 3B Evrişimli bir Sinir Ağı (3D CNN) ile birleştirerek 3B lazer ve sensörlerden alınan bilgileri tam olarak kullanmak için yeni bir mimari olan VoxNet'i önermiştir [31]. VoxNet, tahmin için CNN kullanır, mekânlara ait ayrıntılı bilgilerin kaybına yol açmıştır.



Şekil 3. ShapeNet 3B nesne örnekleri [16].

ShapeNet, insanlar tarafından doğrulanmış ve özelliklerin açıklamalarını sunan büyük ölçekli 3B modeller sunmuş ve bu modeller NB dağılımını öğrenmek için derin inanç ağlarını kullanmıştır. Bu yöntem, nesnelere ait 3B şekiller ve ızgaralar üzerindeki ikili değişkenleri olasılık dağılımları ile temsil etmiştir. Hacimsel yöntemler ile başarılı performans elde edilmiş; fakat bu yöntemlerin çoğu hesap karmaşası ve bellekte fazla yer kaplamasından dolayı ızgaraların çözünürlüğü sınırlı kalmıştır.

OctNet nokta bulutunu, normal ızgaralar boyunca sekiz bitlik bir dizi olarak kodlamış ve havuzlanan özellikleri aritmetik olarak hesaplamıştır. Seyrek 3B veriler üzerinde DÖ'nün bir temsili olan OctNet sunulmuştur [32]. Bu yöntem derin ve yüksek çözünürlüğe sahip 3B CNN etkinleştirmiş NB'yi hiyerarşik bölümlere ayırmak için karma bir ızgara-oktree yapısı sunarak verimliliği artırmıştır. OctNet, 3B nesne sınıflandırma, yön tahmini ve NB etiketleme dâhil olmak üzere 3B görevler üzerindeki etkisi analiz edilmiş ve avantajları belirtilmiştir. 3B şekilleri analizi etmek için Octree tabanlı Evrişimli Sinir Ağı olan O-CNN sunmuştur [33]. O-CNN 3B yüzeylerin kapladığı oktant hesaplamalarını kısıtlamış, bellek ve hesaplama maliyetini oktree derinliği artırdıkça büyümüşür. Bu da 3B CNN'yi yüksek çözünürlüklü 3B modeller için uygun hale getirmiştir. Bu bölümde anlatıldığı gibi hacimsel temsillerle ifade edilen yöntemler kalitesiz olabilir; çünkü voksellerin sadece küçük bir kısmı boş değildir ve her voksel içindeki ayrıntılı durumlar neredeyse hiç toplanmamıştır. O-CNN ise OctNet üzerine kurulan bu modelin eksik yanlarını tamamlamak için geliştirilmiştir.

Noktasal Tabanlı Yöntemler

Noktasal yöntemlerin, nokta komşuluklarına göre noktaları birleştiren projeksiyon tabanlı yöntemlerden farklı her bir noktaya ait özelliği bireysel noktalardan öğrenmiş olmasıdır.

Çoklu katman tabanlı yöntemleri

PointNet MLP oluşturulmuş ilk nokta tabanlı mimardır. Girdi ilk olarak, geometrik dönüşümler altında değişmezliği sabit tutmak için T-Net tarafından tahmin edilen afin (öteleme-döndürme-ölçeklendirme) dönüşüm matrisi ile çarpılmaktadır. Nokta kümesi bir MLP'den ve ortak hizalama ağından sonra genel özellikleri elde etmek için maksimum havuz katmanından geçirilmiştir. Bu temel yapı hem sınıflandırma hem de segmentasyon tahmini için kullanılmıştır. Sınıflandırmada, genel özellikler çıktı değerleri için bir MLP'den geçirilmiştir. Geleneksel CNN'ler, bir dizi evrişim katmanı ile farklı özellikleri temel alarak PointNet ++ geliştirilmiştir.

PointWeb önceki çalışmalardan esinlenerek bir bölgeyi daha iyi temsil etmek için her noktanın özelliklerini yerel bölge özelliklerine göre belirlemiş yerel bir komşuluk ile her noktayı yoğun bir şekilde birbirine bağlamıştır [34]. PointWeb, Point-Net ++ 'dan uyarlanmış ve yerel komşulukların bağlam bilgisini kullanmak için Uyarlanabilir Özellik Ayarı (AFA) ile özelliklerin kalitesini iyileştirmiştir.

Duan vd., yerel bölgelerindeki NB'nin yapısal bağımlılıkları arasında bağlantı kurmak için yapısal ilişki ağını (structural relation network/SRN) önermiştir [35]. PointNet ++ gibi nokta kümelerindeki ağ mimarileri, yerel yapıların iç etkileşimlerini dikkate almadan ayrı ayrı yakalamış; fakat SRN modülü insanların 3B nesneleri anlamasında kritik roller oynayan geometrik ve konuma bağlı ilişkilerini modelleyerek yerel bilgilerini eş zamanlı olarak kullanmıştır. Önerilen SRN modülü basit, yorumlanabilen diğer ağlarla kolayca uyum sağlayan bir ek denetim gerektirmemektedir. Bunlara ek olarak SRN, PointNet ++ donatmak için geliştirilen SRN daha başarılı bir performans göstermiştir.

Evrişimli tabanlı yöntemler

Genellikle CNN ile yapılan çalışmalarda 2B görüntüler üzerindeki evrişim çekirdeklerini, 3B NB üzerinde de çalışmak üzere genişletilmiştir. Makalede önceden belirtildiği gibi VoxNet doğrudan 3B evrişim uygulamıştır.

A-CNN, nokta bulutlarında tıkanık bölgelerdeki gürültüleri önlemek için, A-CNN ile küreler yerine halka tabanlı bir şema önermiştir [36]. 3B NB geometrik ve anlamsal özelliklerini derin ağlar aracılığıyla analiz etmek, geometrik yapılarının düzensizliğini, seyrekliğini önerilen A-CNN ile halka şeklindeki evrişim sayesinde doğrudan 3B NB'ye evrişimi tanımlamış ve hesaplama yapmıştır. Bu yeni evrişim işleminde hesaplama halka şeklindeki yapıları ve yönleri belirleyerek her noktanın yerel komşuluk geometrisini daha başarılı bir şekilde yakalamıştır.

Liu & Fan NB verilerinde istenilen şekli yakalamak ve NB analizini yapmak için normal ızgara CNN'i düzensiz konfigürasyonla genişleten RS-CNN (Relation-Shape

CNN) önermiştir [37]. RS-CNN ağının öğrenmedeki temel mantığı noktalar arasındaki geometrik kısıtlamadan gelmiştir. Bir RS-Conv çekirdeği, girdi olarak belirli bir nokta etrafındaki komşuları yakın noktasal ilişkilerden, noktalar arasındaki üst düzey ilişkileri eşlemeyi öğrenmiş ve öğrenilen haritalama ile komşuluğundaki mekânsal yapıyı kodlamıştır. RS-CNN, NB analizini bağlamsal şekillerden öğrenmek için hiyerarşik bir mimari geliştirmesi ile sinyal işleme düzeyinde geometrik değişkenliğe, ölçeklenebilirliğe uyum sağlayan büyük ölçekli sahneler üzerinde sınıflandırma ve segmentasyon yapmak için uygulamıştır.

Evrişim çekirdeklerinden oluşan bir başka ağ olan PointCov, ağırlık ve yoğunluk fonksiyonlarından meydana gelen 3B noktaların yerel koordinatlarının doğrusal olmayan fonksiyonlardan oluşan evrişim çekirdeklerinden oluşur [38]. PointConv evrişim işleminde önemli olan sürekli gizli bir 3B evrişimin Monte Carlo tahminini bulmasıdır. Bu adım belirlenen noktaya göre MLP'den, evrişim çekirdeklerinden, yoğunluk tahmininden ve ağırlıklandırma işlemlerinden oluşmaktadır. Ayrıca, 3B evrişim, hafıza, hesaplama verimliliği ve kolay dağıtım için matris çarpımlarına ve 2B evrişime indirgenmiştir. PointCov'un en önemli katkısı, ağırlık fonksiyonlarının verimli bir şekilde hesaplanması için önerilen yeni bir formülün performansı önemli ölçüde iyileştirmesidir. Diğer yandan önerilen evrişim çekirdeği, 3B uzayda herhangi bir noktadaki dönüşüm ve permutasyon ile değişmeyen evrişimi hesaplamak için kullanılmıştır. PointConv, alt örneklenmiş özellikler bir nokta bulutundan orijinal çözünürlüğüne geri yaymak için ters evrişim operatörleri kullanmıştır. Buradaki fark evrişimi yoğunluk fonksiyonuna dayalı olarak almış ve Monte Carlo tahmini ile değiştirmiştir.

Lan vd., yerel bir bölgedeki komşu noktaların geometrik ilişkilerini modellemek için, GeoConv olarak adlandırılan kapsamlı evrişim benzeri bir işlemi her noktaya ve yerel komşusuna uygulayan Geo-CNN önerilmiştir [39]. Geo-CNN altı ortogonal baz ile alanı sekiz bölüme ayırıp ve belirli bir kadrantdaki (çemberin dörtte biri) tüm vektörleri üç bazla oluşturmuştur. Özellikler, ağırlık matrislerine karşılık gelen yöne ait matrisler boyunca bağımsız olarak çıkarılır ve geometrik vektör ile tabanlar arasındaki açıya göre toplanmıştır.

SFCNN (spherical fractal cnn) girdi NB ayrık küre koordinatlarına sahip düzenli ikozahedral kafesler üzerine yansıtılmıştır. Bu nedenle evrişim, küresel kafeslerden ve komşularından gelen birleştirilmiş özellikler üzerinde maksimum paylaşım yoluyla uygulanmıştır [40]. SFCNN, dönme değişmezliğini korur ve bozulmalara karşı dayanıklıdır. Nokta bulutlarını tanımak için küresel simetriye dayalı genel, esnek ve 3B dönüşle değişmeyen bir küreyi yaklaştırmak ve ayırmak için düzenli ikosahedral (yirmi yüzlü eşkenar üçgen) bir kafes ve fraktallerini (kırılma noktaları) tanıtarak, 3B noktaları işlemek için bir çerçeve sunmuştur [40]. Sunulan çerçeve, görüntü tanıma için evrişimli sinir ağlarından güçlü temsil özelliklerini ve

genelleme kabiliyetini miras alarak CNN'e benzer özellikler gösterir; fakat rotasyon ve bozulmalara karşı SFCNN sağlam özellikleri öğrenmek için CNN'i genişletmiştir. Küresel Fraktal Evrişimli Sınır Ağları ile (SFCNN) NB özelliklerini etkili ve sağlam bir şekilde öğrenmeyi gerçekleştirmiştir.

Grafik tabanlı yöntemler

Grafik ağları nokta bulutlarını grafik olarak algılar, grafiğin köşelerini noktasal değerlerden belirler ve grafiğin kenarlarını ise belirlenen her noktanın yakın komşuluğuna göre oluşturur.

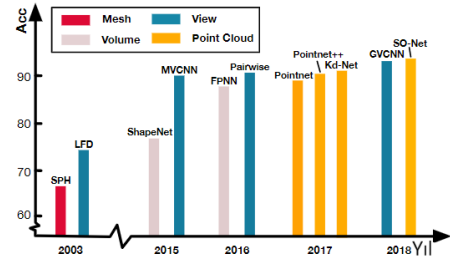
Simnovsky ve Komodakis her bir noktayı grafiğin tepe noktası olarak alıp komşu noktalar arasındaki bağlantılı kenarları bir MLP ağ filtresinden geçirdikten sonra kenar koşullu bir evrişim ağı önermiştir (EdgeCont) [41]. EdgeCont yerel komşuluk bilgilerini içerir; küresel şekil özelliklerini öğrenmek için kümeleme yapar, çok katmanlı sistemlerin özellik uzayındaki noktaların durumuna göre anlamsal özellikleri yakalar, VoxelGrid algoritması ile grafik oluşturmuştur. Sonrasında, DGCNN (Dynamic Graph CNN) ve EdgeConv'u uygulamak için bir MLP kullanılır ve her noktaya ait komşuların kenar özelliklerini simetrik olarak toplamıştır. Bu da grafik ağının her bir katmandan sonra dinamik olarak güncellenmesini sağlamıştır.

ClusterNet, nokta bulutlarındaki noktalardan değişmeyen özellikler oluşturmak için rotasyonla kesin değişmeyen bir dönüş (Rigorous Rotation-Invariant /RRI) modülü ve bir NB hiyerarşik yapılarını oluşturmak için denetimsiz ve derin bir hiyerarşik küme ağ yöntemi önermiştir [42]. Sunulan bu yöntem DGCNN'den referans alınarak nokta bulutları üzerindeki nokta ve şekil özelliklerini öğrenmek için denetimsiz eğitilmiştir.

Örgü tabanlı yöntemler

Nokta bulutları arasında ilişki bulunan noktaların birleştirilmesiyle örgü (mesh) yapısı oluşturmaktadır [43]. NB verileri ile oluşturulan nesnelere bu örgü yapılarının üçgenel olarak birleştirilmesi ile oluşturulur. Ayrıca nesnelere 3B NB olarak ifade edilen verilerin örgü yapısı yüzey bilgisi olarak ta kullanılmaktadır.

Genel anlamda 3B iki tür veri kümesi vardır: iç mekân sahneleri [6], [44] ve dış mekân kentsel sahneleri [8]. İç mekân sahnelerindeki NB ya yoğun derinlik haritalarından dönüştürülür ya da 3B noktasal örgülerden elde edilir. Dış mekân sahneleri, NB seyrek olduğu otonom sürüşler için tasarlanmıştır. Dış mekân sahnelerinden toplanan NB için hareketli ve sabit platform üzerinde kinematik ve geometrik parametrelerin konumları ve açıları eniyileme yapılarak [45] daha başarılı veri kümeleri elde edilebilir.



Şekil 4. Mesh dağılımı[49].

Benzer noktaları birleştirmek için oluşturulan benzerlik matrisi büyük bellek tüketimine neden olmuştur. Liu vd., her bir vokselin anlamsal puanını ve yakın voksellerin tahmini [45] için alt manifold seyrek evrişim ağından faydalanmıştır [46]. Son olarak, tahmini yakınlığa göre ve ağ yapısına göre gruplamak için bir kümeleme algoritması sunmuştur.

PointNet++, noktanın anlamlı etiketini ve ayrık maskelerini oluşturmak için temel bir yapı olmuştur [2]. Liang vd., ayırıcı kısımları öğrenmek için bu yapıya duyarlı bir kayıp önermiştir [47]. Bu öneri, benzer özellikleri ve noktalar arasındaki geometrik ilişkileri esas almıştır. Ayrıca dikkate duyarlı bir grafik CNN, yakın noktasal bilgiler ile eğitilerek iyileştirme yapılmıştır. VRMesh Studio NB otomatik sınıflandırma, özellik çıkarma ve düz yüzey örgü çıkarmayı sağlamıştır.



Şekil 5. a) NB b) Voksel c) Üçgen ağ(mesh) [48].

Analiz

Bu çalışmada incelenen makalelerden, NB verilerinin birçok zorluğuna rağmen başarılı çevrimici ve çevrimdışı olarak çalışan uygulamaların gelişime açık olduğu gösterilmiştir. NB'yi gerçek zamanlı ve doğru bir şekilde gruplara ayırmak için yeni modeller geliştirilmiştir. Yeni model ve algoritmaların gelişimine katkı sağlamak 3B NB'yi DÖ yöntemleri ile sınıflandırmak için kapsamlı bir araştırma makalesi sunulmuştur. Araştırılan yöntemlerin nokta bulutlarını sınıflandırılırken başarılı tahmin yöntemleri ile beraber bölge büyütme, model geliştirme, matematiksel model ve geometrik tekniklerin özellikleri ile analiz edilmiştir. NB sınıflandırma işlemlerinde nesnelere yerleştirilirken model seçim boyutunun zorluğu, gürültüye duyarlılığı ve karmaşık sahnelerdeki çalışmaları incelenmiş en iyi sonuçlar tartışılmıştır.

Tablo 2. 3B veri kümelerinin sınıflandırma sonuçları (% ortalama doğruluk) gösterilmiştir.

Model	ModelNet-40	ShapeNet	ScanNet
PointNet	89.2	83.7	73.9
PointNet++	90.7	-	84.5
PointWeb	92.3	-	85.9
SRN-PointNet++	91.5	-	-

PointConv	92.5	82.8	-
RS-CNN	93.6	86.1	-
DGCNN	92.2	86.2	-
So-Net	90.9	84.6	-

3B NB'ye ait özellikler sınıflandırmada DÖ tekniklerinin karmaşık sahnelerde hazır makine öğrenme algoritmalarından daha başarılı olduğu görülmüştür. Hazır makine öğrenimi teknikleri genellikle yavaş ve sonuçlar seçilen öznel çıkarma işlemine dayanmaktadır. Sonuç olarak 3B nokta bulutlarını sınıflandırmada DÖ yöntemlerinden PointConv, DGCNN ve ClusterNet önerilmiştir. Zengin bir kaynakça içeriğine sahip 3B NB sınıflandırma araştırmasının bu önemli konu hakkında değerli bilgiler sağlayacağına ve yeni araştırmalar için yol gösterici bir araştırma olacağına inanıyorum.

Sonuç

3B NB işleme alanında makale sayılarının yetersiz olduğu araştırılmış ve bu alanda literatürü zenginleştirmek için; Özellikle 3B NB ile 3B verilerin sınıflandırılma yöntemleri detaylı olarak araştırmacılar sunulmuştur. Bu çalışmada verilerin 3B bir ortamda noktasal olarak temsil edilmesi ve bu yapılandırılmamış noktaların eğitilmesi bu alanı diğer alanlardan ayırmıştır. Ek olarak bu alana özgü derin öğrenme tabanlı yeni modellerin geliştirilmesi ile 3B bir uzayda nesnelere farklı yöntemlerle sınıflandırılması literatüre yeni kaynaklar sağlamıştır. Gelecek çalışmalarda yapılandırılmamış noktaların doğru bir şekilde nokta tabanlı yeni modül ve uygun yöntemlerle birleştirilmesi sonucunda bellek kullanımı ve hızlı performansa sahip modeller oluşturulabilir.

Standart bir DÖ mimarisinde, evrişim tabanlı ağlar, 3B nokta bulutlarında üstün performans gösterebilir. Düzensiz verileri işlerken ayırık ve sürekli evrişim ağlarını kullanırken dikkat edilmelidir. Düzensiz NB verilerinin güçlü yapısı nedeniyle, grafik tabanlı ağlar son zamanlarda daha fazla dikkat çekmiştir. Bununla beraber, alandaki grafik tabanlı ağların 3B alanda farklı özelliklere sahip grafik yapılarını genişletmek ve geliştirmek hala zordur.

Etik kurul onayı ve çıkar çatışması beyanı

Hazırlanan makalede etik kurul izni alınmasına gerek yoktur.

Hazırlanan makalede herhangi bir kişi/kurum ile çıkar çatışması bulunmamaktadır.

Yazar Katkıları

Yazılan derleme makalesinde özellikle nokta bulutu alanında geliştirilen modellerin katkıları ve özellikleri kıyaslanmıştır.

Teşekkür

Bu çalışmanın gerçekleştirilmesinde, değerli bilgilerini benimle paylaşan, saygıdeğer danışman hocam; Dr.Öğr.Üyesi A. Burak İNNER'e, teşekkür ederim.

Kaynaklar

- [1] Qi, C. R., Su, H., Mo, K., & Guibas, L. J. (2017'a). PointNet: Deep learning on point sets for 3D classification and segmentation. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-January*, 77–85. <https://doi.org/10.1109/CVPR.2017.16>
- [2] Qi, C. R., Yi, L., Su, H., & Guibas, L. J. (2017). PointNet++: Deep hierarchical feature learning on point sets in a metric space. *ArXiv, (Nips)*.
- [3] Mccarthy, T., Fotheringham, A. S., Charlton, M., Winstanley, A., & Malley, V. O. (2007). Integration of LiDAR and stereoscopic imagery for route corridor surveying. *ANational Centre for Geocomputation, National University of Ireland, Maynooth*.
- [4] Chang, A. X., Funkhouser, T., Guibas, L., Hanrahan, P., Huang, Q., Li, Z., ... Yu, F. (2015). *ShapeNet: An Information-Rich 3D Model Repository*. Retrieved from <http://arxiv.org/abs/1512.03012>.
- [5] Wu, Z., Song, S., Khosla, A., Yu, F., Zhang, L., Tang, X., & Xiao, J. (2015). 3D ShapeNets: A deep representation for volumetric shapes. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 07-12-June, 1912–1920*. <https://doi.org/10.1109/CVPR.2015.7298801>.
- [6] Dai, A., Chang, A. X., Savva, M., Halber, M., Funkhouser, T., & Nießner, M. (2017). ScanNet: Richly-annotated 3D reconstructions of indoor scenes. *ArXiv*.
- [7] Geiger, A., Lenz, P., Stiller, C., & Urtasun, R. (2013). Vision meets robotics: The KITTI dataset. *International Journal of Robotics Research, 32(11)*, 1231–1237.
- [8] Kesten, R., Usman, M., Houston, J., Pandya, T., Nadhamuni, K., Ferreira, A., ... Shet, V. (2019). Lyft Level 5 AV Dataset 2019. Retrieved from <https://level5.lyft.com/dataset/2020>.
- [9] Zhao, H., Jiang, L., Fu, C. W., & Jia, J. (2019). Pointweb: Enhancing local neighborhood features for point cloud processing. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2019-June*, 5560–5568. <https://doi.org/10.1109/CVPR.2019.00571>.
- [10] Duan, Y., Zheng, Y., Lu, J., Zhou, J., & Tian, Q. (2019). Structural relational reasoning of point clouds. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2019-June*, 949–958. <https://doi.org/10.1109/CVPR.2019.00104>
- [11] Rusu, R. B., & Cousins, S. (2011). 3D is here: Point Cloud Library (PCL). *Proceedings - IEEE International Conference on Robotics and Automation, (May)*. <https://doi.org/10.1109/ICRA.2011.5980567>

- [12] Nguyen, A., & Le, B. (2013). 3D point cloud segmentation: A survey. *IEEE Conference on Robotics, Automation and Mechatronics, RAM - Proceedings*, 225–230. <https://doi.org/10.1109/RAM.2013.6758588>
- [13] Shamir, A. (2006). Segmentation and shape extraction of 3D boundary meshes. *Eurographics*, (September), 137–149.
- [14] Pu, S., & Vosselman, G. (2006). Automatic extraction of building features from terrestrial laser scanning. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 36.
- [15] Xu, Danfei ; Anguelov, Dragomir; Jain, A. (2018). PointFusion: Deep Sensor Fusion for 3D Bounding Box Estimation. (arXiv:1711.10871v2 [cs.CV] UPDATED). *The IEEE Conference on Computer Vision and Pattern Recognition*, 244–253. Retrieved from <http://arxiv.org/abs/1711.10871>
- [16] Chang, A. X., Funkhouser, T., Guibas, L., Hanrahan, P., Huang, Q., Li, Z., ... Yu, F. (2015). *ShapeNet: An Information-Rich 3D Model Repository*. Retrieved from <http://arxiv.org/abs/1512.03012>
- [17] Xiang, Y., Mottaghi, R., & Savarese, S. (2014). Beyond PASCAL: A benchmark for 3D object detection in the wild. *2014 IEEE Winter Conference on Applications of Computer Vision, WACV 2014*, 75–82. <https://doi.org/10.1109/WACV.2014.6836101>
- [18] Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2010). SUN database: Large-scale scene recognition from abbey to zoo. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 3485–3492. <https://doi.org/10.1109/CVPR.2010.5539970>
- [19] Armeni, I., Sener, O., Zamir, A. R., Jiang, H., Brilakis, I., Fischer, M., & Savarese, S. (2016). 3D Semantic Parsing of Large-Scale Indoor Spaces Supplementary Material. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1534–1543
- [20] Hackel, T., Savinov, N., Ladicky, L., Wegner, J. D., Schindler, K., & Pollefeys, M. (2017). Semantic3D.Net: a New Large-Scale Point Cloud Classification Benchmark. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4(1W1), 91–98. <https://doi.org/10.5194/isprs-annals-IV-1-W1-91-2017>
- [21] Fritsch, J., Kuhn, T., & Geiger, A. (2013). A new performance measure and evaluation benchmark for road detection algorithms. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, (Itsc), 1693–1700. <https://doi.org/10.1109/ITSC.2013.6728473>
- [22] Geiger, Andreas, Lenz, P., & Urtasun, R. (2012). Are we ready for autonomous driving? the KITTI vision benchmark suite. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 3354–3361. <https://doi.org/10.1109/CVPR.2012.6248074>
- [23] Schoenberg, J. R., Nathan, A., & Campbell, M. (2010). Segmentation of dense range information in complex urban scenes. *IEEE/RSJ 2010 International Conference on Intelligent Robots and Systems, IROS 2010 - Conference Proceedings*, 2033–2038. <https://doi.org/10.1109/IROS.2010.5651749>
- [24] Xiong, X., Munoz, D., Bagnell, J. A., & Hebert, M. (2011). 3-D scene analysis via sequenced predictions over points and regions. *Proceedings - IEEE International Conference on Robotics and Automation*, 2609–2616. <https://doi.org/10.1109/ICRA.2011.5980125>
- [25] Munoz, D., Bagnell, J. A., Vandapel, N., & Hebert, M. (2009). Contextual classification with functional max-margin markov networks. *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2009, 2009 IEEE*, 975–982. <https://doi.org/10.1109/CVPRW.2009.5206590>
- [26] Vallet, B., Brédif, M., Serna, A., Marcotegui, B., & Paparoditis, N. (2015). TerraMobilita/iQmulus urban point cloud analysis benchmark. *Computers and Graphics (Pergamon)*, 49, 126–133. <https://doi.org/10.1016/j.cag.2015.03.004>
- [27] Su, H., Maji, S., Kalogerakis, E., & Learned-Miller, E. (2018). GVCNN: Group-View Convolutional Neural Networks for 3D Shape Recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 264–272. <https://doi.org/10.1109/CVPR.2018.00035>
- [28] Yang, Ze, & Wang, L. (2019). Learning relationships for multi-view 3D object recognition. *Proceedings of the IEEE International Conference on Computer Vision, 2019-October*, 7504–7513. <https://doi.org/10.1109/ICCV.2019.00760>
- [29] Yu, T., Meng, J., & Yuan, J. (2018). Multi-view Harmonized Bilinear Network for 3D Object Recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2, 186–194. <https://doi.org/10.1109/CVPR.2018.00027>
- [30] Maturana, D., & Scherer, S. (2015). VoxNet: A 3D Convolutional Neural Network for real-time object recognition. *IEEE International Conference on*

- Intelligent Robots and Systems, 2015-Decem, 922–928. <https://doi.org/10.1109/IROS.2015.7353481>.
- [31] Riegler, G., Osman Ulusoy, A., & Geiger, A. (2017). Octnet: Learning deep 3d representations at high resolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3577-3586).
- [32] Wang, Y., Fathi, A., Kundu, A., Ross, D. A., Pantofaru, C., Funkhouser, T., & Solomon, J. (2020). *Pillar-Based Object Detection for Autonomous Driving. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12367 LNCS, 18–34. https://doi.org/10.1007/978-3-030-58542-6_2
- [33] Zhao, H., Jiang, L., Fu, C. W., & Jia, J. (2019). Pointweb: Enhancing local neighborhood features for point cloud processing. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2019-June*, 5560–5568. <https://doi.org/10.1109/CVPR.2019.00571>
- [34] Duan, Y., Zheng, Y., Lu, J., Zhou, J., & Tian, Q. (2019). Structural relational reasoning of point clouds. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2019-June*, 949–958. <https://doi.org/10.1109/CVPR.2019.00104>
- [35] Komarichev, A., Zhong, Z., & Hua, J. (2019). A-cnn: Annularly convolutional neural networks on point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 7421-7430)*.
- [36] Liu, Y., Fan, B., Xiang, S., & Pan, C. (2019). Relation-shape convolutional neural network for point cloud analysis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 8895-8904)*.
- [37] Wu, W., Qi, Z., & Fuxin, L. (2019). Pointconv: Deep convolutional networks on 3d point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 9621-9630)*.
- [38] Lan, S., Yu, R., Yu, G., & Davis, L. S. (2019). Modeling local geometric structure of 3d point clouds using geo-cnn. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 998-1008)*.
- [39] Rao, Y., Lu, J., & Zhou, J. (2019). Spherical fractal convolutional neural networks for point cloud recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 452-460).
- [40] Simonovsky, M., & Komodakis, N. (2017). Simonovsky, Komodakis - 2017 - Dynamic Edge-Conditioned Filters in Convolutional Neural Networks on Graphs.pdf. *Cvpr*, 3693–3702.
- [41] Hassani, K., & Haley, M. (2019). Unsupervised multi-task feature learning on point clouds. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 8160-8171).
- [42] Bischoff, B.S., et al. 2002. OpenMesh—a generic and efficient polygon mesh data structure. In OpenSG Symposium.
- [43] S. Song, S. P. Lichtenberg, and J. Xiao, “Sun RGB-D: A RGB-D scene understanding benchmark suite,” in CVPR, 2015.
- [44] İnnner, A. Burak (2013). Sweart Platform Benzetim ve Eniyileme Yazılımının Gerçekleştirilmesi. Doktora Tezi, Kocaeli Üniversitesi Fen Bilimleri Enstitüsü, Kocaeli
- [45] C. Liu and Y. Furukawa, “MASC: Multi-scale affinity with sparse convolution for 3D instance segmentation,” arXiv preprint arXiv:1902.04478, 2019.
- [46] B. Graham, M. Engelcke, and L. van der Maaten, “3D semantic segmentation with submanifold sparse convolutional networks,” in CVPR, 2018.
- [47] Z. Liang, M. Yang, and C. Wang, “3D graph embedding learning with a structure-aware loss function for point cloud semantic instance segmentation,” arXiv preprint arXiv:1902.05247, 2019.
- [48] Hoang, L., Lee, S. H., Kwon, O. H., & Kwon, K. R. (2019). A Deep Learning Method for 3D Object Classification Using the Wave Kernel Signature and A Center Point of the 3D-Triangle Mesh. *Electronics*, 8(10), 1196.
- [49] Feng, Y., Feng, Y., You, H., Zhao, X., & Gao, Y. (2019, July). Meshnet: Mesh neural network for 3d shape representation. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 33, No. 01, pp. 8279-8286).