

FARKLI MLS NOKTA BULUTU YOĞUNLUKLARININ VE KOMŞULUK YÖNTEMLERİNİN KONTROLLÜ SINIFLANDIRMAYA ETKİSİ

Semanur SEYFELİ¹, Ali Özgün OK²,

¹Yük. Müh., Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, 06800, Çankaya, Ankara, semanur.seyveli@hacettepe.edu.tr

²Prof. Dr., Hacettepe Üniversitesi, Geomatik Mühendisliği Bölümü, 06800, Çankaya, Ankara, ozgunok@hacettepe.edu.tr

ÖZET

Mobil Lazer Tarama (MLS) sistemleri, genellikle kentsel alanların ve yol ağlarının 3-boyutlu haritalanmasında tercih edilen, hızlı, yüksek doğrulukta ve yüksek yoğunlukta bir veri toplama yöntemidir. Bu çalışmada, MLS nokta bulutu sınıflandırmasında geleneksel olarak kullanılan yöntemlerden olan 2-boyutlu ve 3-boyutlu k-en yakın komşuluk (kNN), küresel ve silindirik komşuluk yöntemleri değerlendirilmiştir. İşlem süresini azaltmak ve beraberinde farklı komşuluk hesabı yöntemlerinin düşük yoğunlukta nokta bulutlarındaki sonuca etkisini değerlendirmek amacıyla veri alt örnekleme uygulanmıştır. Bu hususta; 3 ana aşamada nokta tabanlı kontrollü sınıflandırma işlemi gerçekleştirilmiştir: (i) yerel komşuluk ilişkisinin kurulması, (ii) öznitelik bilgisinin çıkarılması ve (iii) nokta tabanlı sınıflandırma. Yöntemler, araç tabanlı MLS sistemleriyle toplanmış ve her biri 8 semantik sınıfı içeren TUM-MLS1 ve Toronto-3D nokta bulutları üzerinde test edilmiştir. Asıl ve farklı yoğunluklardaki alt örneklemlere ayrılan nokta bulutlarından, belirlenen sabit parametrelili yerel komşuluk türlerine göre geometrik ve şekil tabanlı öznitelikler çıkarılmış ve Rastgele Orman sınıflandırma yöntemi her bir noktayı etiketlemede tercih edilmiştir. Sonuç olarak, her iki nokta bulutu verisinde de %95,1 genel doğrulukta silindirik komşuluk yöntemi kullanılarak en iyi sonuç elde edilmiştir. Nokta bulutu yoğunluğunun azaltılmasıyla işlem süresi bakımından bütün yöntemlerde beklendiği üzere azalmalar gözlenmiş olup genel doğruluklarda da kayıplar olduğu gözlemlenmiştir. Ancak bu bağlamda küresel ve silindirik komşuluklara kıyasla kNN yöntemlerinin (2- ve 3-boyutlu) nokta boyutu yoğunluğundan diğer yöntemlere nazaran çok daha az etkilendiği ve düşük nokta bulutu yoğunluğunda en iyi performansın belirtilen iki farklı veri seti için 2-boyutlu kNN ile sağlandığı sonucuna varılmıştır.

Anahtar Sözcükler: MLS, alt örnekleme, öznitelik çıkarma, sınıflandırma

ABSTRACT

THE IMPACT OF DIFFERENT MLS POINT CLOUD DENSITIES AND NEIGHBORHOOD METHODS ON SUPERVISED CLASSIFICATION

Mobile Laser Scanning (MLS) systems are a fast, high-accuracy, and high-density data collection method generally preferred for 3-dimensional mapping of urban areas and road networks. In this study, 2-dimensional and 3-dimensional k-nearest neighborhood (kNN), spherical, and cylindrical neighborhood methods, which are traditionally utilized in MLS point cloud classification, were evaluated. In order to reduce the processing time and to evaluate the effect of different neighborhood calculation methods on the result in low density point clouds, data subsampling was applied. In this respect, point-based supervised classification was performed in 3 main stages: (i) establishing local neighborhood relationships; (ii) extracting feature information; and (iii) point-based classification. The methods were tested on TUM-MLS1 and Toronto-3D point clouds, which were collected with vehicle-based MLS systems and each contained 8 semantic classes. Geometric and shape-based features were extracted from the original and sub-sampled point clouds according to the local neighborhood types with fixed parameters, and the Random Forest classification method was preferred for the labeling of each point. As a result, the best results were obtained by using the cylindrical neighborhood method, with an overall accuracy of 95.1% for both point cloud data. With the reduction of the point cloud density, decreases were observed in terms of processing time, as expected in all methods, and losses in overall accuracy were also observed. However, in this context, it was concluded that kNN methods (2- and 3-dimensional) were notably less affected by the point size density than other methods compared to spherical and cylindrical neighborhoods, and the best performance was achieved with 2-dimensional kNN for the two different datasets specified at low point cloud density.

Keywords: MLS, sub-sampling, feature extraction, classification

1. GİRİŞ

LiDAR veri toplama ilkelerine dayalı olarak oluşturulan 3-boyutlu (3B) nokta bulutu verileri sıklıkla kentsel planlama, simülasyon, acil durum müdahalesi, haritalama ve görselleştirme gibi çeşitli uygulama çalışmalarında kritik bir veri kaynağı olarak önem taşımaktadır. Yerleştirildikleri platformlara göre yersel lazer tarama, mobil lazer tarama, havadan lazer tarama ve uzaydan lazer tarama olarak karşımıza çıkmaktadır. Bunlar arasında mobil lazer tarama (MLS), yüksek çözünürlüklü 3B topoğrafik verileri elde etmek için kullanılan hızlı ve nokta bulutu yoğunluğu görece yüksek bir ölçüm yöntemidir.

Çoğu MLS sistemi; navigasyon sistemi (GNSS/IMU sistemleri, odometre), 3B lazer tarayıcılar, kameralar (dijital görüntüleme kameraları, termal kameralar vs.) ve bilgisayar gibi kontrol sisteminden oluşur ve bütün bu donanım hareket eden bir platforma yerleştirilir. Şehir bölgelerinde; 3B yüzey bilgilerini toplamak için genellikle araç tabanlı

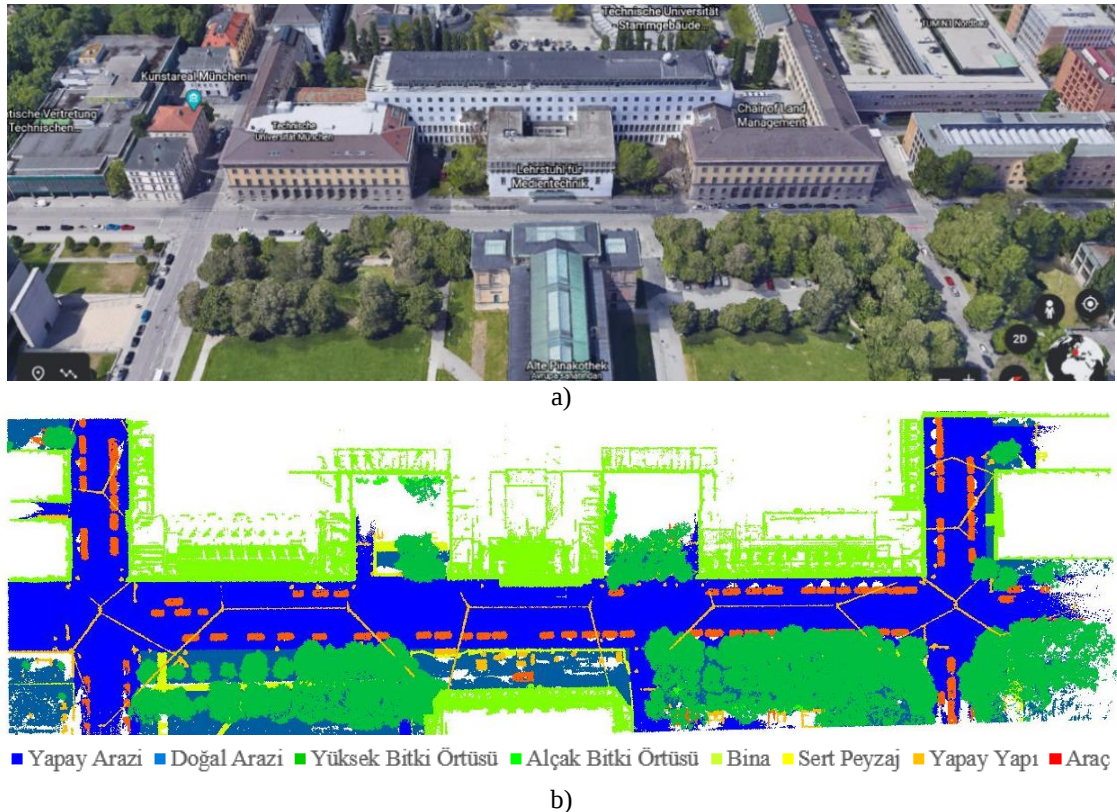
MLS sistemleri tercih edilir. Diğer sistemlerle karşılaştırıldığında, havadan ve uzaydan lazer tarama yöntemlerine göre daha yoğun ve detaylı veri üretilirken; yersel lazer tarama ile statik olarak ölçüm yapıldığından sokak düzeyinde veri toplama ise mobil sistemlerde çok daha hızlıdır. Ancak MLS sistemleri ile üretilen nokta bulutu düzensizdir, yani platforma yakın alanlarda noktalar daha sık aralıklı iken platformdan uzaklaştıkça nokta bulutu yoğunluğu azalır.

Mobil lazer tarama sonucunda; her bir noktanın hangi sınıfa ait olduğu bilinmeyen, 3B ham nokta bulutu veri seti elde edilir. Yoğun ve ham nokta bulutu için temel işleme adımlarından biri ise her noktayı ait olduğu sınıfa atamaktır. Sınıflandırma aşaması bölüt tabanlı (örn. Yokoyama vd., 2012; Xiao vd., 2016; Soilán vd., 2017), nokta tabanlı (örn. Demantké vd., 2012; Weinmann vd., 2015; Zheng vd., 2017; Wang vd., 2018; Atik vd., 2021; Seyfeli ve Ok, 2022a; Seyfeli ve Ok, 2022b) ya da derin öğrenme tabanlı (örn. Balado vd., 2019; Guo ve Feng, 2020) gibi yöntemlerle olacak şekilde birçok çalışma için ilgi odağı olmuştur. MLS ile elde edilen nokta bulutlarının çok yoğun olması hatta milyonlarca nokta içermesi sebebiyle bu noktaların veri işleme aşaması oldukça zaman alıcı ve zordur. Özellikle sınıflandırma aşamasında verinin eğitilmesi veya test sürecinde çeşitli sorunlara sebebiyet verebilmektedir.

Bu çalışmada MLS nokta bulutunun sınıflandırma performansı, iki farklı veri seti için çeşitli yoğunlukta alt örneklemelere ayrılarak 2-boyutlu (2B) ve 3B k-en yakın komşuluk (kNN), küresel ve silindirik komşuluk yöntemleri aracılığıyla değerlendirilmiştir. Kullanılan veri setleri araç tabanlı MLS sistemleriyle elde edilen ve farklı yoğunluklardaki TUM-MLS1 ve Toronto-3D nokta bulutlarından oluşmaktadır. Her iki nokta bulutu da birbirleriyle farklı sınıflar içerse de toplamda 8 semantik sınıfı içermektedir. Sınıflandırma aşaması üç temel adımda ele alınmıştır; (i) yerel komşuluk ilişkisinin kurulması, (ii) nokta bulutundan öznitelik çıkarma ve (iii) nokta tabanlı sınıflandırma. Noktaların yerel komşuluk ilişkisi kurulup öznitelik açısından zenginleştirildikten sonra rastgele orman sınıflandırıcısıyla her bir nokta bağımsız olarak sınıflandırılmıştır. Ek olarak, çok yoğun nokta bulutlarında veri işleme süresini iyileştirme ve bu yönde sonuçları test etmeye yönelik, asıl veri yoğunluğunun %1'ine kadar farklı yoğunlukta verilerin kullanılmasıyla sonuçlanan bir veri alt örnekleme yöntemi kullanılmış ve yerel komşuluk seçim yöntemleri, tam ve alt örneklenmiş nokta bulutları için uygulanmıştır.

2. MATERYAL VE YÖNTEM

Yöntemi değerlendirmek için kullanılan TUM-MLS1 nokta bulutu verisi, 18 Nisan 2016 tarihinde Münih (Almanya) şehrinde bulunan Münih Teknik Üniversitesi'nin (TUM) şehir kampüsüne aittir (Zhu vd., 2020). Test alanı manuel olarak etiketlenmiş yapay arazi, doğal arazi, yüksek bitki örtüsü, düşük bitki örtüsü, bina, sert peyzaj, yapay yapı ve araç olarak 8 semantik sınıfı içermektedir ve referans veri setinde 3 milyondan fazla nokta bulunmaktadır (Şekil 1).

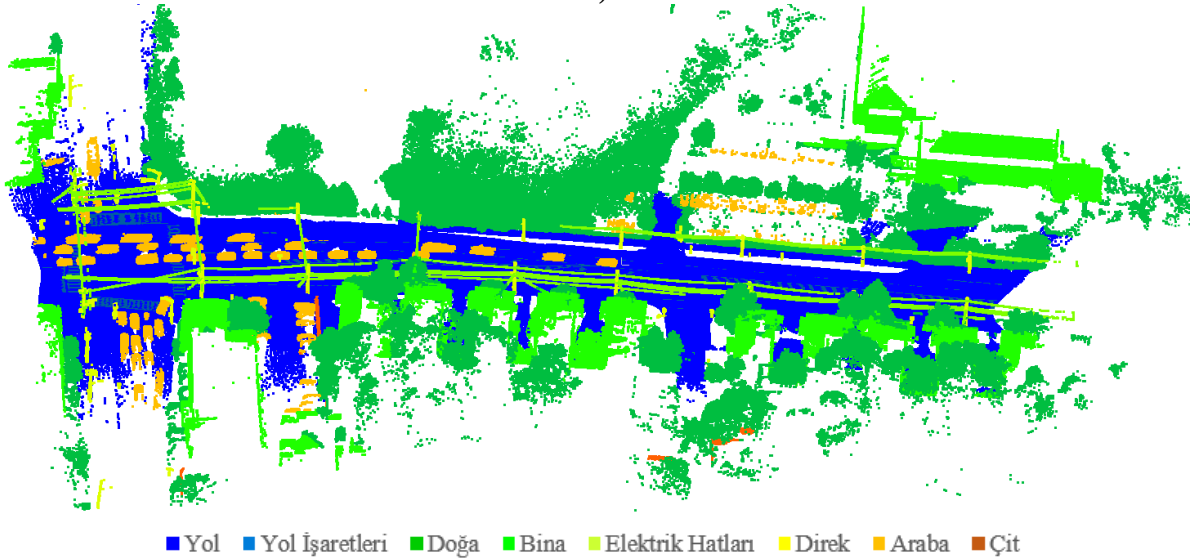


Şekil 1. TUM-MLS1 verisine ait a) Google Earth görüntüsü ve b) referans nokta bulutu (Zhu vd., 2020).

Diğer açık kaynak nokta bulutu verisi olan Toronto-3D ise, 2020 yılında Toronto'daki (Kanada) Avenue Road'dan toplanıp 4 bölüme ayrılmıştır (Tan vd., 2020). Veri seti yol, yol işaretleri, doğa, bina, elektrik hatları, direk, araba ve çit olmak üzere 8 nesne sınıfını kapsamaktadır. Bu çalışma için veri setinin sınıflandırılmamış noktaları çıkarılmış haliyle yaklaşık 6 milyondan fazla nokta içeren Şekil 2'deki L004 kısmı kullanılmıştır.



a)

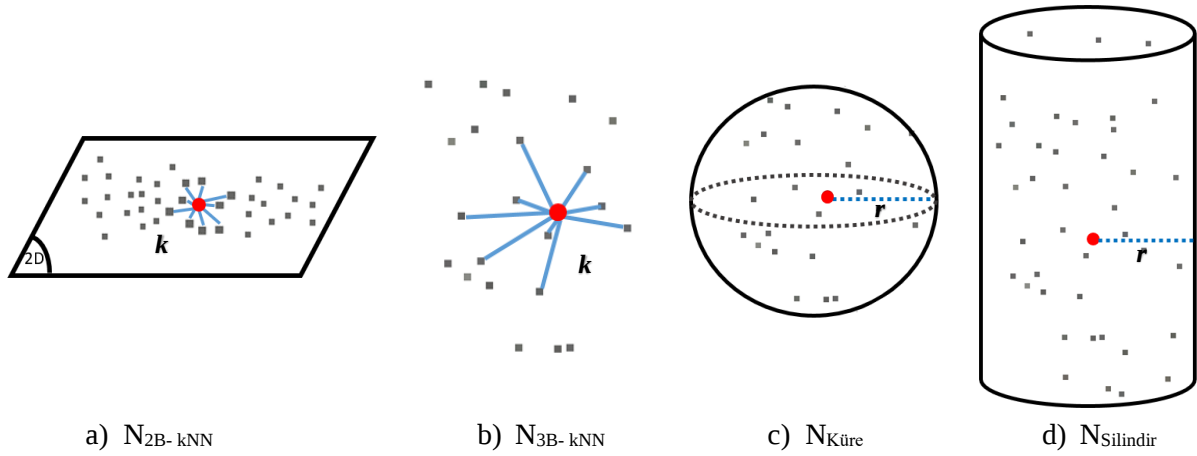


b)

Şekil 2. Toronto-3D verisinin L004 kısmına ait a) Google Earth görüntüsü ve b) referans nokta bulutu (Tan vd., 2020).

Verilerin nokta tabanlı kontrollü sınıflandırılması aşamasında; yerel komşuluk ilişkisi kurularak her bir nokta için geometrik ve şekil tabanlı öznitelikler çıkarılması ilk adımdır. Yerel komşuluk (N) kurulurken öncelikle her bir nokta merkez olarak alınır ve belirlenen komşuluk türüne göre komşuluğun içinde kalan noktalar seçilir. Nokta bulutundaki bütün noktalar için öznitelik çıkarma aşamasında ise bu komşu noktalar kullanılmaktadır. Çalışmada MLS veri setlerine Şekil 3'te belirtilen 2B ve 3B k-en yakın komşuluklar (kNN), küresel ve silindirik komşuluklar uygulanmıştır. 2B- ve 3B-kNN yönteminde ilgili noktaya belirlenen en yakın k sayıdaki nokta Öklit mesafesine göre bulunur. 2B-kNN yönteminde öncelikle tüm noktalar XY düzlemine projekte edilir ve arama yatay düzlemde yapılırken 3B-kNN yönteminde ise arama 3 boyutlu uzaydadır. Bu sebeple 3B-kNN daha çok yarıçapı değişken küreye benzerken 2B-kNN daha düzensiz bir yapıdadır. Küresel ve silindirik komşuluk yaklaşımlarında ise ilgili noktanın çevresinde r yarıçapında silindir veya küre oluşturulur ve komşuluk içinde kalan noktalar K-d ağacı algoritması ile bulunur. Küre ve dikey silindir gibi yarıçapa bağlı yöntemlerde sabit bir yarıçap değeri (r), 2B ve 3B-kNN için ise sabit bir k değeri her iki veri seti için önceki çalışmaların (Seyfeli ve Ok, 2022b) sonuçları ve veri işleme kapasitesi de değerlendirilerek seçilmiştir:

- N_{2B-kNN} : k değeri 2B Öklid mesafesine dayalı en yakın 150 noktadır.
- N_{3B-kNN} : k değeri 3B Öklid mesafesine dayalı en yakın 150 noktadır.
- $N_{Küre}$: r yarıçapı 50 cm olan küredir.
- $N_{Silindir}$: r yarıçapı 50 cm olan dikey silindirdir.



Şekil 3. Yerel komşuluk yöntemleri, a) 2B-kNN, b) 3B-kNN, c) küresel ve d) silindirik komşuluk. (k : en yakın nokta sayısı, r : yarıçap)

Çalışmada; MLS nokta bulutlarının 3B koordinat verilerinden yararlanılarak geometrik tabanlı özellikler (Çizelge 1) çıkarılmaktadır. Noktanın yüksekliği, nokta bulutundan alınan temel geometrik özelliklerden biridir. Bu sebeple yükseklik bilgisi öncelikle CloudCompare yazılımındaki Cloth Simulation Filtering (CSF) eklentisi (Zhang vd., 2016) kullanılarak normalize edilmiş ve yerden yükseklik olarak alınmıştır. Noktanın normalize yüksekliğine ek olarak yerel komşuluktaki en düşük ve en yüksek noktalar arasındaki yükseklik farkı ve tüm noktaların yüksekliklerinin standart sapması gibi yüksekliğe bağlı; yerel komşuluk içinde kalan nokta sayısı ve dikeylik gibi geometrik özellikler çıkarılmıştır.

Çizelge 1'deki şekil tabanlı özellikler ise öz değerlere bağlıdır. Temel bileşen analizi (PCA), nokta bulutunda 3B olarak nokta dağılımının yönlerini ve komşuluğun ağırlık merkezine bağlı olarak nokta dağılımının varyasyonundaki büyüklüğünü verir (Demantké vd., 2012). PCA ile komşuluk içindeki her nokta için 3x3 boyutunda kovaryans matrisi oluşturulur ve λ_1 , λ_2 ve λ_3 olmak üzere öz değerler üretilir. Öz değerler pozitifdir ve $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ olarak sıralanır. Hesaplanan öz değerlerden biri büyük ve diğerleri sıfıra yakın olduğunda yerel komşuluk çizgi şekli tanımlar. İki öz değer hemen hemen aynı değere sahipse ve diğeri sıfıra yakınsa, komşuluk bir düzlem oluşturur. Bunun dışında hepsi büyükse 3 boyutlu küresel veya belirsiz bir yüzey oluşturur.

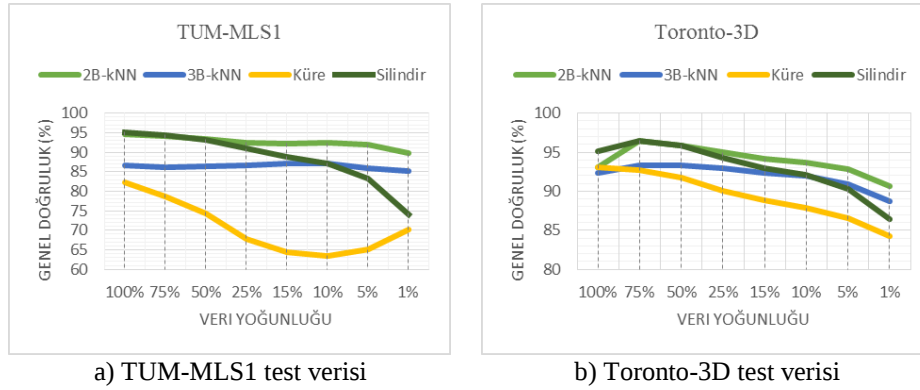
Çizelge 1. Noktaların konumuna dayalı hesaplanan ve her bir noktaya atanan yerel geometrik ve şekil öznitelikleri.

Öznitelik türü	Öznitelik Adı	Formül
Geometrik tabanlı	Normalize yükseklik	H
	Yükseklik farkı	$\Delta H = h_{\max} - h_{\min}$
	Yüksekliğin standart sapması	$\sigma H = \sqrt{\frac{1}{k} \sum_{j=1}^k (H_j - \mu H)^2}$
	Yerel komşuluktaki (N) nokta sayısı	N
	Dikeylik	$V = 1 - n_z$
Şekil tabanlı	Lineerlik	$L_\lambda = \frac{\lambda_1 - \lambda_2}{\lambda_1}$
	Düzlemsellik	$P_\lambda = \frac{\lambda_2 - \lambda_3}{\lambda_1}$
	Küresellik	$S_\lambda = \frac{\lambda_3}{\lambda_1}$
	Omnivaryans	$O_\lambda = \sqrt[3]{\lambda_1 \lambda_2 \lambda_3}$
	Anizotropi	$A_\lambda = \frac{\lambda_1 - \lambda_3}{\lambda_1}$
	Özentropi	$E_\lambda = -\sum_{i=1}^3 \lambda_i \ln(\lambda_i)$
	Öz değerler toplamı	$\Sigma \lambda = \lambda_1 + \lambda_2 + \lambda_3$
	Eğrilikteki değişim	$c_\lambda = \frac{\lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}$

Çıkarılan öznelik bilgileri ve yoğun nokta bulutları için sıklıkla tercih edilen geleneksel yöntemlerden biri olan Rassal Orman (Breiman, 2001) sınıflandırıcısı aracılığıyla noktalar otomatik olarak etiketlenmiştir. Sonraki aşamada; veri setindeki nokta yoğunluğuna bağlı olarak sınıflandırma sonucundaki değişimi analiz edebilmek için nokta bulutu %75, %50, %25, %15, %10, %5 ve %1 olmak üzere rastgele alt örneklem kümelerine ayrılmış ve komşuluk seçim yöntemleri alt örneklenmiş veri kümeleri için değerlendirilmiştir. Verinin alt örneklemelere ayrılması aşaması açık kaynaklı nokta bulutu işleme yazılımı olan CloudCompare yazılımında yapılmış, diğer işlemler ise MATLAB yazılım ortamında geliştirilmiştir.

3. BULGULAR

Yerel komşuluk yöntemlerinin ve veri yoğunluğunun sınıflandırmayı ne şekilde etkilediğini belirlemek için, yöntem farklı alt örneklemelerde üretilen nokta bulutları üzerinde test edilmiştir. Her bir alt örneklem verisi için büyük veri kümelerinde kullanışlı olan uzak-tutma çapraz doğrulama (hold-out cross validation) yöntemi aracılığıyla nokta bulutu rastgele tüm verinin %20'si eğitim %5'i doğrulama ve %75'i test verisi olarak ayrılmıştır. Böylece sınıflandırma modeli eğitim verileriyle tahmin edilirken, eğitilen modelin performansı için doğrulama verileri kullanılmış, sınıflandırma sonucu ise bağımsız olarak test verisi üzerinden değerlendirilmiştir. Sınıflandırma sonucunu değerlendirmek için hata matrisinden faydalanılmış ve hata matrisi yardımı ile her bir model için genel doğruluk, kullanıcı ve üretici doğruluğu değerleri hesaplanmıştır. Şekil 4, test verilerinin genel sınıflandırma doğruluklarını sunmaktadır.



Şekil 4. MLS nokta bulutu verilerinde yerel komşuluk tiplerinin nokta yoğunluğuna göre genel doğruluk sonuçları

TUM-MLS1'e ait asıl test veri seti genel sınıflandırma doğruluğu sonuçları 2B-kNN, 3B-kNN, küresel ve silindirik komşuluklar için sırası ile %94,5, %86,6, %82,2 ve %95,1'dir ve en iyi sonuç silindirik komşuluğa aittir. Şekil 6'ya göre kullanıcı ve üretici doğrulukları açısından, her dört yöntem için en çok LiDAR noktası içeren yapay arazi, yüksek bitki örtüsü ve bina sınıfları için yüksek doğruluk söz konusudur. Uygulanan yöntemler arasından en düşük sonuç TUM-MLS-1 test verisi için küresel komşuluktan elde edilmiştir. Özellikle sert peyzaj, yapay yapı ve alçak bitki örtüsü sınıflarına ait noktaların veri setinde diğer sınıflara göre daha seyrek olması ve daha az bilgi içermesi sonuçlar üzerinde olumsuz etki göstermiştir. Nitekim küresel komşuluk için bu sınıflar üretici doğruluğu açısından %20'nin altında kalmıştır. Bu durumun sabit olarak seçilmiş olan komşuluk boyutunun (r yarıçapı) bu sınıflar için yeterli olmamasından kaynaklandığı değerlendirilmiştir. Yapay ve doğal arazi sınıfları için ise 2B-kNN iken, diğer sınıflar için silindirik komşuluk yöntemi daha başarılıdır.

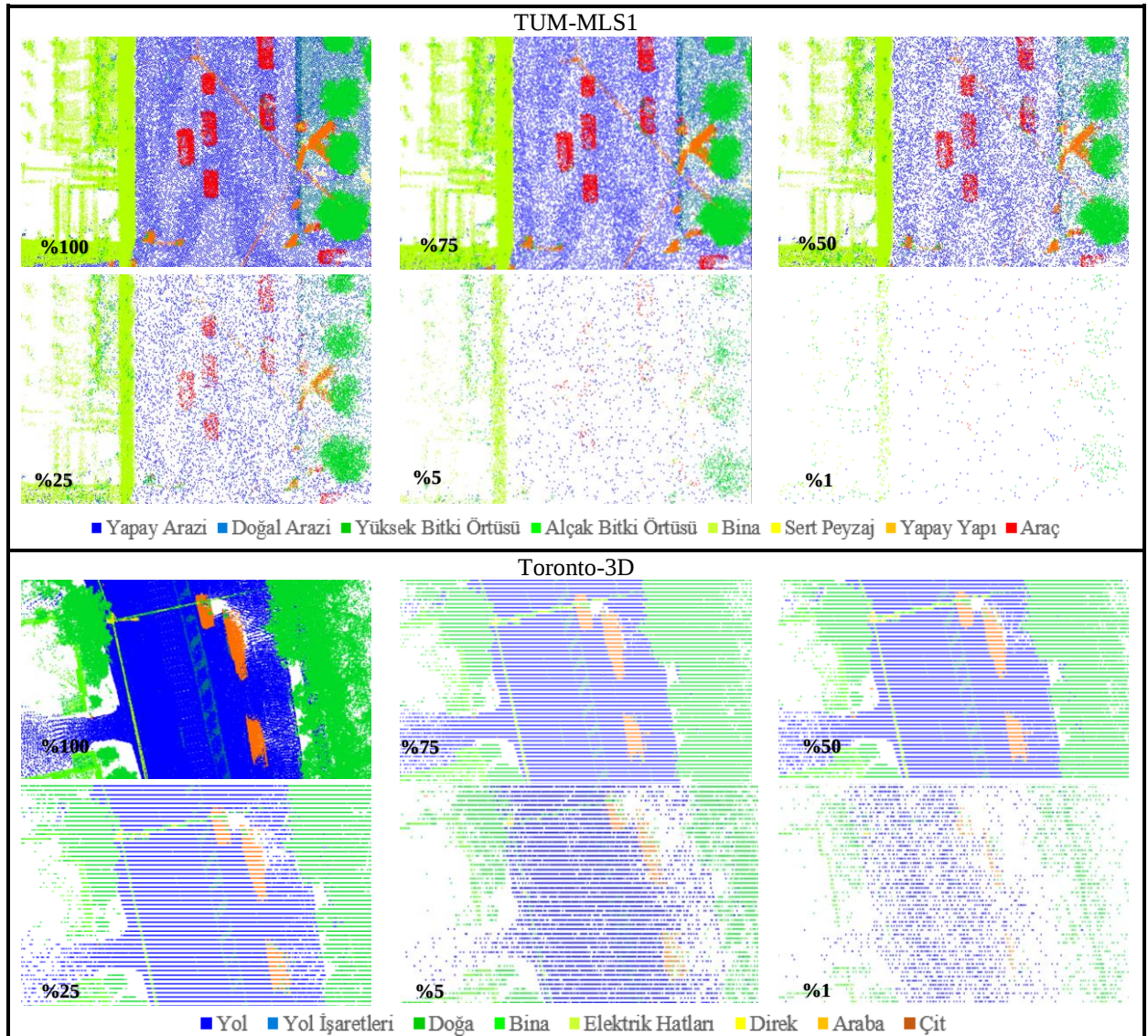
Toronto-3D asıl nokta bulutu test verisi için ise 2B- ve 3B-kNN, küresel ve silindirik komşuluklar için %93,1, %92,3, %93,1 ve %95,1 genel doğruluk ile önceki sonuçlarla benzer olarak silindirik komşuluk en iyi performans göstermiştir. Test verisindeki yol işareti sınıfının öznelik açısından yol sınıfı ile benzer olması, bu sınıfın üretici doğruluğunun bütün yöntemlerde düşük olmasına (%13-34 aralığında) sebebiyet vermektedir. Bu sınıfın ayırt edilebilirliğini arttırmak için noktalara ait (varsa) renk bilgisinin kullanılabileceği değerlendirilmektedir. Çit sınıfı için ise 2B-kNN ve silindirik komşuluk yöntemlerinin sonuçları birbirine çok yakındır. TUM-MLS1 sonuçlarına benzer şekilde Toronto-3D test verisinde Şekil 7'ye göre yol, bina ve doğa gibi nokta sayısı fazla olan sınıfların bütün yöntemlerde doğruluğu yüksektir. Silindirik komşuluk, yol, yol işaretleri, doğa, bina, direk, araba ve çit sınıflarının çıkarılmasında diğer yöntemlerden daha başarılıdır.

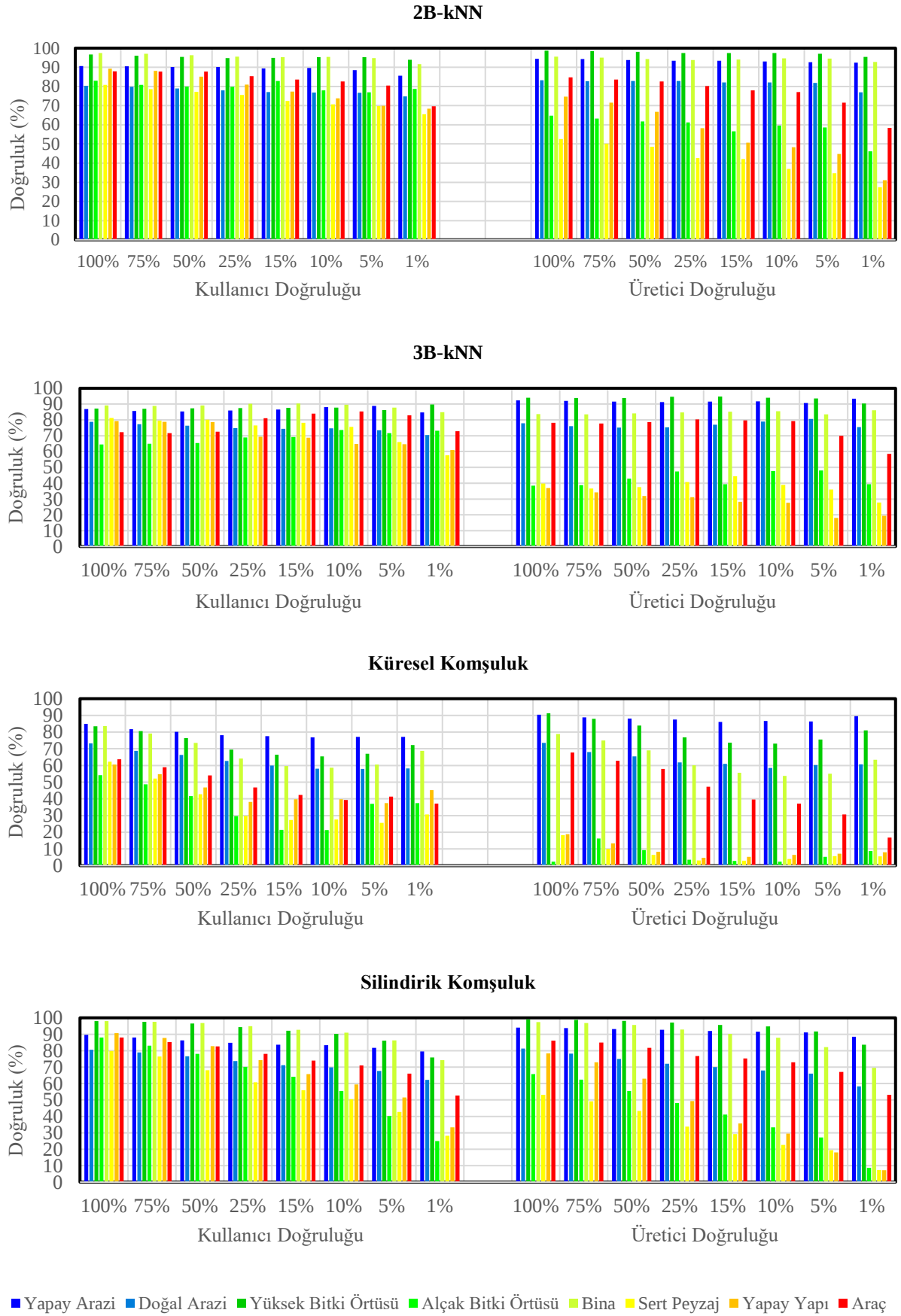
Veri işleme süreleri, verinin boyutu, kullanılan yöntemlere ve seçilen parametrelere bağlı olarak değişmektedir (Çizelge 2). Özellikle silindirik komşuluk yönteminde daha yoğun nokta bulutları için komşuluk içindeki noktaların sayısındaki artışa bağlı olarak da öznelik çıkarımı aşaması diğer yöntemlere göre daha uzun sürelidir. Bu durumda kNN yöntemlerinde komşuluk içindeki k nokta sayısının sabit olması yarıçapa bağlı yöntemlere göre daha avantajlıyken, komşuluk içinde kalan noktaların bulunmasında kullanılan algoritma da önem arz etmektedir.

Çizelge 2. TUM-MLS1 ve Toronto-3D MLS nokta bulutları için bütün işlem adımlarının gerçekleşme süreleri (sn.)

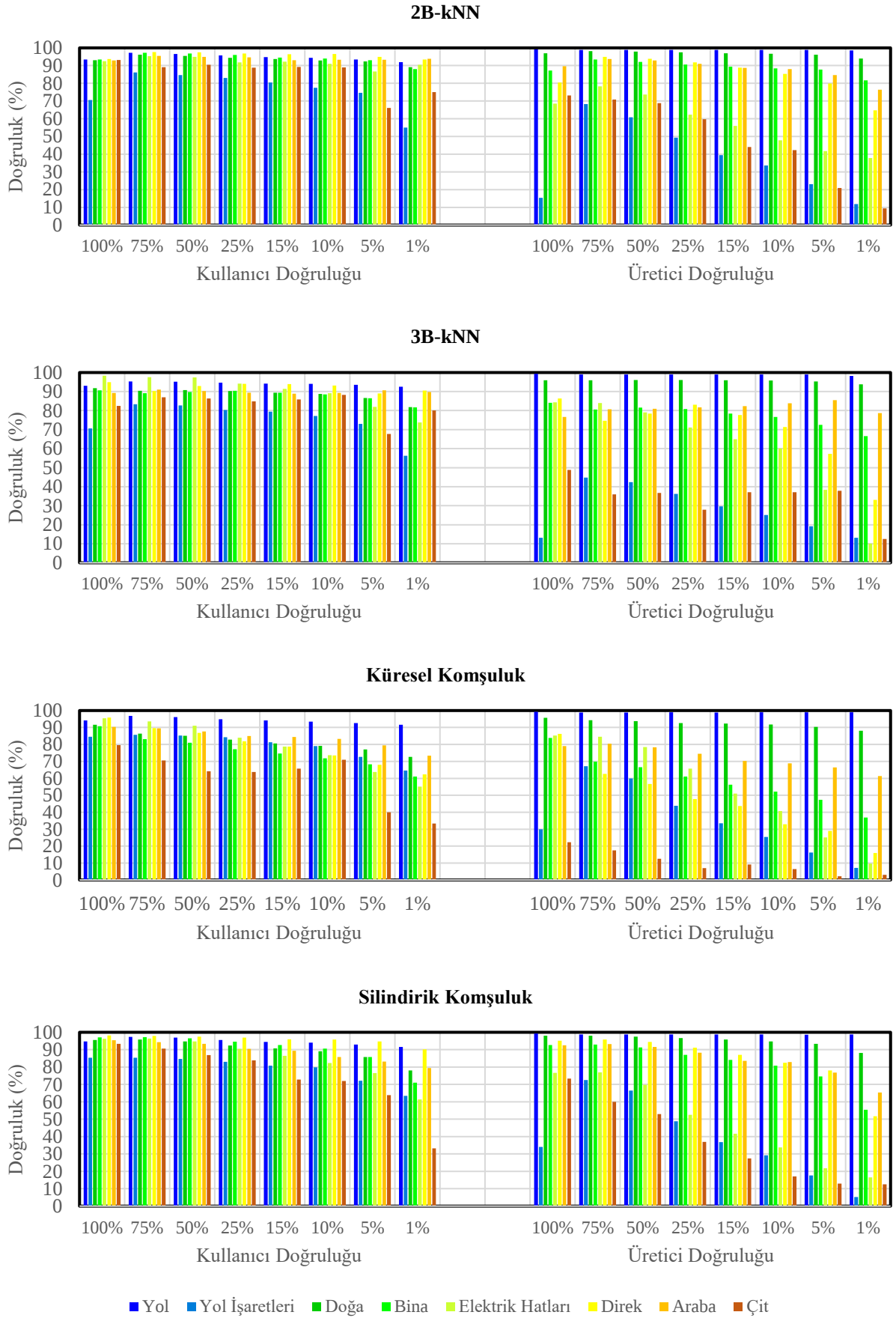
Nokta Yoğunluk Oranı	TUM-MLS1				Toronto-3D			
	N _{2B-kNN}	N _{3B-kNN}	N _{Küre}	N _{Silindir}	N _{2B-kNN}	N _{3B-kNN}	N _{Küre}	N _{Silindir}
%100	320	352	457	565	680	685	1375	3283
%75	253	297	349	392	513	542	720	981
%50	188	221	229	341	429	450	468	678
%25	99	101	103	124	193	202	216	296
%15	57	59	61	65	123	127	151	159
%10	44	46	41	41	64	82	89	95
%5	15	17	18	17	32	32	43	37
%1	4	4	4	4	8	8	8	8

Nokta yoğunluğunun azaltıldığı senaryolarda ise yarıçapa bağlı olan küresel ve silindirik komşuluk ile yapılan sınıflandırmada komşuluk içine düşen nokta sayısı azaldığı için işlem süresi kısalmasına rağmen genel olarak beklentiler yönünde sınıflandırma doğruluğunun azaldığı gözlemlenmiştir. 2B- ve 3B-kNN ise nokta yoğunluğunun azalmasından görece daha az etkilenmiş (< %5) ve en iyi sonuç ise 2B-kNN ($\approx$ %90) ile elde edilmiştir. Bu durumda, silindirik ve küresel komşuluk yöntemleriyle elde edilen sonuçlarda her bir veri seti için sabit bir yarıçap kullanmak yerine nokta yoğunluğuna bağlı olarak yarıçap parametresinin artırılmasıyla sonuçların iyileştirilebileceği değerlendirilmektedir. Nokta yoğunluğu düşürüldüğünde ise nokta bulutundaki önemli bilgilerin de kaybolduğu göz önüne alınmalıdır (örn. Şekil 5). Özellikle çalışmada kullanılan MLS verileri gibi dengesiz dağılmış sınıflar bu durumdan daha çok etkilenebilmektedir.

**Şekil 5.** Yerel silindirik komşuluk kullanılarak yapılan sınıflandırmanın asıl ve alt örneklenmiş test sonuçları.



Şekil 6. TUM-MLS1 asıl ve alt örneklenmiş nokta bulutuna uygulanan dört farklı yerel komşuluk yöntemlerine bağlı sınıflandırılması sonucunda test verisi üzerinden 8 sınıfa ait kullanıcı ve üretici doğrulukları.



Şekil 7. Toronto-3D asıl ve alt örneklenmiş nokta bulutuna uygulanan dört farklı yerel komşuluk yöntemlerine bağlı sınıflandırılması sonucunda test verisi üzerinden 8 sınıfa ait kullanıcı ve üretici doğrulukları.

4. SONUÇ

Bu araştırma MLS veri sınıflandırması için silindirik, küresel, 2B- ve 3B-kNN olmak üzere dört farklı yerel komşuluk yöntemini iki farklı veri setinde test etmiştir. Her bir yöntem için tüm verinin %20'si eğitim %5'i doğrulama olarak Rassal Orman modelinin eğitiminde, %75'i ise sınıflandırma sonucunu bağımsız olarak test edilmesinde değerlendirilmiştir. Kullanılan nokta öznelikleri ve sınıflandırma algoritmasına göre, silindirik komşuluk yönteminin her iki veri seti için de asıl veri yoğunluğunda görece fazla işlem süresine rağmen en iyi sonucu verdiği görülmüştür. Ayrıca çalışmada, veri yoğunluğunu değiştirmek için veri alt örnekleme yöntemi uygulanmış ve açıklanan yöntemle veri alt örnekleme, işlem süresini azaltmada önemli derecede etkili olmuştur. Asıl veriden oluşturulan %1'e kadar yoğunluktaki alt örnekleme veri setlerinden, yerel komşuluk yöntemlerinden en yakın noktaya bağlı kurulan 2B- ve 3B-kNN komşulukların doğruluk seviyeleri daha az etkilenirken yarıçapa bağlı silindirik ve küresel komşuluklar daha çok etkilenmiştir. Nitekim bu senaryoda en iyi sonuçlar her iki veri seti için de 2B-kNN ile elde edilmiştir. Diğer yöntemlerin sonuçlarını iyileştirmek adına komşuluk boyutlarının yoğunluğa bağlı olarak uyarlanabileceği değerlendirilmektedir. Ancak veri alt örnekleme özelliği yoğunluk bakımından dengesiz sayılabilecek veri setlerinde örnekleme sayısı diğerlerine göre daha az olan sınıflara ait önemli bilgilerin kaybolmasına yol açabildiği de göz önünde bulundurulmalıdır. Gelecekteki çalışmalarımız derin öğrenme dâhil olmak üzere MLS sınıflandırması için farklı stratejilere ve analizlere odaklanacaktır.

TEŞEKKÜR

TUM-MLS1 veri kümesinin telif hakkı Fraunhofer IOSB'ye aittir ve CC BY-NC-SA 4.0 Lisansı altında kullanıma sunulmuştur. Toronto-3D veri seti ise, Waterloo Üniversitesi, Mobil Algılama ve Geodata Bilim Laboratuvarı'na aittir ve CC BY-NC 4.0 Lisansı altındadır. Veri setlerini sağladıkları için Dr. Marcus Hebel, Joachim Gehring ve Weikai Tan'a teşekkür ederiz.

KAYNAKLAR

- Atik, M. E., Duran, Z. ve Seker D. Z.,** 2021. Machine Learning-Based Supervised Classification of Point Clouds Using Multiscale Geometric Features. *ISPRS International Journal of Geo-Information*, 10(3), 187.
- Balado, J., Martínez-Sánchez, J., Arias, P., Novo, A.,** 2019. Road Environment Semantic Segmentation with Deep Learning from MLS Point Cloud Data. *Sensors*, 19(16), 3466.
- Breiman, L.,** 2001. Random Forests, *Machine Learning*, 45, 5–32.
- Demantké, J., Mallet, C., David, N. ve Vallet, B.,** 2012. Dimensionality Based Scale Selection in 3D LiDAR Point Clouds. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38, 97–102.
- Guo, Z. ve Feng, C. C.,** 2018. Using Multi-Scale and Hierarchical Deep Convolutional Features for 3D Semantic Classification of TLS Point Clouds. *International Journal of Geographical Information Science*, 34(4), 661-680.
- Seyfeli, S. ve Ok, A.O.,** 2022a, Classification of Mobile Laser Scanning Point Cloud in an Urban Environment using kNN and Random Forest. *in: Ben Ahmed, M., Boudhir, A.A., Karaş, İ.R., Jain, V., Mellouli, S. (Eds.), Innovations in Smart Cities Applications Volume 5. Springer International Publishing, Cham*, 963–973.
- Seyfeli, S. ve Ok, A.O.,** 2022b, Classification of Mobile Laser Scanning Data with Geometric Features and Cylindrical Neighborhood. *Baltic Journal of Modern Computing*, 10(2), 209-223.
- Soilán, M., Riveiro, B., Martínez-Sánchez, J. ve Arias, P.,** 2017. Segmentation and Classification of Road Markings Using MLS Data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 123, 94–103.
- Tan, W., Qin, N., Ma, L., Li, Y., Du, J., Cai, G., Yang, K. ve 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)*, Seattle, WA, USA, 797–806.
- Wang, Y., Chen, Q., Liu, L., Li, X., Sangaiah, A. K. ve Li, K.,** 2018. Systematic Comparison of Power Line Classification Methods from ALS and MLS Point Cloud Data. *Remote Sensing*, 10(8), 1222.

Weinmann, M., Jutzi, B., Hinz, S., Mallet, C., 2015. Semantic Point Cloud Interpretation Based on Optimal Neighborhoods, Relevant Features and Efficient Classifiers, *ISPRS Journal of Photogrammetry and Remote Sensing*, 105, 286-304.

Xiao, W., Vallet, B., Schindler, K. ve Paparoditis, N., 2016. Street-Side Vehicle Detection, Classification and Change Detection Using Mobile Laser Scanning Data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 166–178.

Yokoyama, H., Date, H., Kanai, S. ve Takeda, H., 2012. Pole-Like Objects Recognition from Mobile Laser Scanning Data Using Smoothing and Principal Component Analysis. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38, 115–120.

Zhang, W., Qi, J., Wan, P., Wang, H., Xie, D., Wang, X. ve Yan, G., 2016. An Easy-to-Use Airborne LiDAR Data Filtering Method Based on Cloth Simulation, *Remote Sensing*, 8(6), 501.

Zheng, M., Lemmens, M. ve van Oosterom, P., 2017. Classification of Mobile Laser Scanning Point Clouds from Height Features. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W7, 321–325.

Zhu, J., Gehring, J., Huang, R., Borgmann, B., Sun, Z., Hoegner, L., Hebel, M., Xu, Y. ve Stilla, U., 2020. TUM-MLS-2016: An Annotated Mobile LiDAR Dataset of The TUM City Campus for Semantic Point Cloud Interpretation in Urban Areas. *Remote Sensing*, 12 (11), 1875.