

# İnsansız Araçlar için Anlamsal Bölütleme ile İmge Tabanlı Konumlandırma Image Based Localization Using Semantic Segmentation for Autonomous Driving

İbrahim ÇINAROĞLU ve Yalın BAŞTANLAR

Bilgisayar Mühendisliği

İzmir Yüksek Teknoloji Enstitüsü

İzmir, Türkiye

ibrahimcinaroglu@iyte.edu.tr, yalinbastanlar@iyte.edu.tr

**Özetçe**—Bilgisayarlı Görü alanındaki popülerliğini koruyan araştırma konularından birisi insansız araçlarda yer tespiti ve konumlandırma. Araçların konumlandırılmasında kullanılan GPS sistemlerinin bazı durumlarda faal olmadığı bilinen bir gerçektir ve bu yetersizlik imge tabanlı konumlandırma çalışmalarına hız vermiştir. Bizim çalışmamızda, araç içinden elde edilmiş Malaga şehir merkezi görüntülerinden oluşan bir veri tabanı kullanılarak imge tabanlı konumlandırma yapılmıştır. İlk olarak, anlamsal (semantik) bölütleme sonucunda elde edilen bir anlamsal betimleyici oluşturulmuş ve yaklaşık en yakın komşuluk araması tekniği de kullanılarak bir konumlandırma yapılmıştır. Ardından bu yöntemin başarısı, literatürde sıkça kullanılan yerel betimleyici tabanlı yöntemin başarısıyla kıyaslanmıştır. Ayrıca, bu iki yöntemin birleştirilmesi ile elde edilen melez bir yöntem önerilmiştir. Önerilen melez imge-tabanlı konumlandırmanın, sadece yerel betimleyici ve sadece anlamsal betimleyici kullanan yöntemden daha başarılı olduğu, dolayısıyla yerel betimleyici tabanlı yöntemlerin anlamsal betimleyiciler ile desteklenmesinin başarıyı artırdığı, deneysel sonuçlarla gösterilmiştir.

**Anahtar Kelimeler** — *imge tabanlı konumlandırma; anlamsal bölütleme; insansız araç; imge eşleştirme.*

**Abstract**—One of the research topics that maintains its popularity in the field of Computer Vision is place recognition and localization for autonomous vehicles. It is a known fact that GPS systems used for localizing vehicle cannot be activated in some cases and this inability has accelerated image based positioning studies. In our study, we performed image based localization using dataset that includes Malaga city center images. Firstly, a semantic descriptor is obtained as a result of semantic segmentation and localization was performed using the approximate nearest neighbor search. After that, success of this method was compared with the success of the local descriptor based method which is frequently used in the literature. Furthermore, a hybrid method obtained by combining these two methods is proposed. The superiority of the proposed hybrid image-based localization method, and hereby contribution of the semantic descriptor is demonstrated by experimental results.

**Keywords** — *image based localization; semantic segmentation; autonomous driving; image matching.*

## I. GİRİŞ

Yapay zekâ ve makine öğrenmesi tekniklerinin geleceğimizi şekillendirdiği bu günlerde, bu gelişmelerden nasibini alan bir çalışma alanı da insansız/sürücüsüz araçlardır. Sürücüsüz araçlarda yer tespiti ve konumlandırma için dâhili ölçüm cihazları (IMU) ile desteklenen küresel konumlandırma sistemlerinden (GPS) yararlanılmaktadır. Ancak GPS sistemlerinin, yoğun bitki örtüsü ve yapılar, tüneller, askeri alanlar gibi çeşitli çevresel faktörler ile etkisiz hale gelmektedir. GPS sistemlerinin bu önemli eksikliği ve bu cihazların tek başına güvenilir olmadığı gerçeği, imge tabanlı konumlandırma çalışmalarına hız vermiştir. Bu çalışmalarda, coğrafi konumları bilinen imgelerden oluşan bir veri tabanı olduğunda, bir sorgu imgesinin veri tabanındaki hangi imgeye en yakın olduğu bulunur. En yakın imgenin konumu, çoğu zaman sorgu imgesinin konumu olarak varsayılır.

İmge tabanlı yer tespiti çalışmalarında, imgelerin ifade edilmesinde sıklıkla kullanılan yöntemlerin başında, imgedeki ilgi noktalarından oluşturulan yerel betimleyiciler (SIFT[4], SURF[21], FAST[5] vb.) gelmektedir [1]. Bununla birlikte, sıklıkla kullanılan diğer bir öznitelik çıkarım tekniği de imgenin bir bütün olarak ifade edildiği genel betimleyicilerdir (Gist vb.). İmgelerdeki ilgi noktalarını karşılaştırmaya dayalı teknikler ve gelişmiş versiyonları (VLAD [24], Disloc, [6]), iki boyutlu (2B-2B) imgeler üzerinde yapılan imge tabanlı kıyaslamalarda kullanılırken; öteki taraftan konumun üç boyutlu modeli çıkarılarak 2B noktalar ile üç boyutlu (2B – 3B) noktalar arasındaki eşlemeye bakan çalışmalar da mevcuttur. Bizim çalışmamızda da benimsediğimiz 2B-2B eşleştirmeli yaklaşımın, 2B-3B yaklaşımına göre daha maliyetsiz olduğu ve hemen hemen aynı derecede başarı gösterdiği çeşitli çalışmalar ile gösterilmiştir [2][3]. Bunlara ek olarak, her iki yaklaşım (2B-2B, 2B-3B) için de uygulanması gereken öznitelik eşleştirme yöntemleri incelendiğinde, Yaklaşık En Yakın Komşuluk (ANN) arama yönteminin yüksek boyutlu veri eşleştirme için daha verimli çalıştığı çeşitli çalışmalarda [12][14][15] vurgulanmıştır.

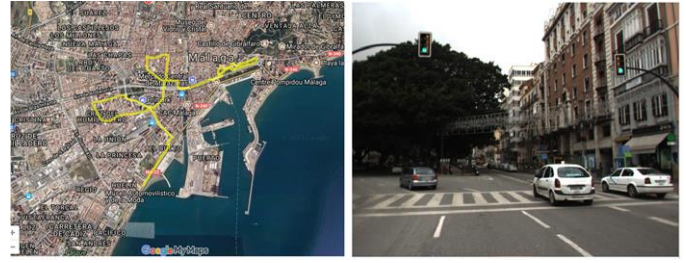
Görüntüdeki anlamsal bilginin konumlandırmayı iyileştirmek için kullanıldığı çalışmalar mevcuttur. Singh ve Kosecka [10] imgelere anlamsal bölütleme uyguladıktan sonra bölütlemeyen bir betimleyici çıkarmıştır. Bu betimleyiciler sahneleri cadde arası, bina önü veya kavşakta olması gibi gruplara ayırmakta kullanılmıştır. Mousavian v.d. [7] yerel betimleyici tabanlı çalışmalarında, çıkardıkları anlamsal etiketleri kullanarak insan yapımı yapılardan gelmeyen yerel öznitelikleri (ağaç vb.) elemişlerdir. Özniteliklerin ağırlıklandırılması insan-yapımı yapılarda artırılırken, doğal yapılarda sağlıklı eşleşme şansı az olduğu için azaltılmıştır. Yine anlamsal ipuçlarının değerlendirilmesi adına ama 2B-3B eşleştirmeye yönelik bir örnekte [8], eşleştirmenin doğruluğu semantik açıdan da kontrol edilmiştir. Semantik bilginin popülerlik kazandığı diğer bir alan olan 3B sahnelere yönelik olarak, Schonberger v.d. [9]'de semantik içerik için bir sözlük hazırlanmış ve sahne anlamsal BoW olarak ifade edilmiştir. Bizim çalışmamızda ise 2B-2B eşleştirmeye dayalı bir yöntem önerilmiş, yukarıdaki çalışmalardan farklı olarak imgelerden çıkarılan anlamsal etiketler ile bir betimleyici oluşturulup doğrudan konumlandırma için kullanılmıştır.

Çalışmamızda araç içi görüntüleri ve GPS verilerini içeren Malaga kentsel veri seti kullanılmıştır. Veri seti imgeleri, önce yerel betimleyici (YB) kullanılarak, daha sonra da özgün bir şekilde oluşturduğumuz anlamsal betimleyici (AB) kullanılarak ifade edilmiştir. Anlamsal betimleyici, imgeler üzerinde belirlemiş olduğumuz 7 nesne sınıfının elle belirlenmesi ve piksel seviyesindeki sınıflandırma bilgisinin vektöre aktarılması ile oluşturulmuştur. Böylece, imgeler üzerinde yaklaşık en yakın komşuluk araması (ANN) yöntemi kullanılan yerel betimleyici tabanlı YB-ANN, anlamsal betimleyici tabanlı AB-ANN yöntemleri ve bu iki yöntemin eşit ağırlıklı olarak birleştirilmesi ile Melez-ANN yöntemi konumlandırma için önerilmiştir. Kullanılan verisetinin özellikleri ve amaca yönelik yapılan düzenlemeler II. Bölümde ele alınmıştır. Daha sonra III. Bölümde, önerilen öznitelik ve eşleştirme yöntemleri geniş bir şekilde açıklanmıştır. Önerilen üç yöntemin veri seti üzerinde uygulanması ve elde edilen deneysel sonuçlar IV. Bölümde işlenmiştir. Son bölümde ise, sonuçlar yorumlanarak gelecekte yapılabilecek olası çalışmalardan bahsedilmektedir.

## II. MALAGA KENTSEL VERİ SETİ

İnsansız araçlarda imge tabanlı konumlandırma çalışmamızda kullanılmak üzere, mevcut imge veri setleri arasından Malaga Stereo ve Lazer Kentsel Veri Seti [11] seçilmiştir. Malaga veri seti; şehir merkezinde 36.8 km'lik rotada dolaşan, bir adet stereo kamera (Bumblebee2) ve beş adet lazer algılayıcı ile donatılmış bir aracın içinden, yüksek çözünürlük (1024x768) ve yüksek sıklıkla (20fps) elde edilen imgelerden oluşan, çeşitli Bilgisayarlı Görü çalışmalarında kullanılmak üzere hazırlanmış referans bir veri setidir (Şekil 1).

Malaga veri seti ile birlikte, bu çalışmamızda metrik hata hesaplaması yapmamız için ihtiyacımız olan, her biri bir imgeye denk gelen 1571 GPS kaydı mevcuttur. Bu 1571 imgeye ait olan GPS (enlem-boylam) bilgilerinin uyumluluğu ve doğruluğu, derece ondalık cinsine çevrilerek Google Haritalardan teyit edilmiştir. Veri setimizde bulunan, ardışık her 3 imgeden üçüncüsü test imgesi olarak seçilmiş, geri kalan-



Şekil 1. Malaga şehir merkezi veri seti rotası (sol) ve veri setine ait örnek bir imge (sağ).

lar eğitim seti olarak ayrılmıştır. Dolayısı ile Bölüm IV' te uygulanan deneyler GPS verileri bulunan 523 test ve 1048 eğitim imgesi üzerinde gerçekleştirilmiştir.

## III. YÖNTEM

### A. Yaklaşık En Yakın Komşular ile İmge Eşleştirme

İmge tabanlı konumlandırma kısaca ifade edilecek olursa; gelen test (sorgu) imgesine en çok benzeyen imgenin coğrafi-etiketli binlerce referans imge arasından etkili bir şekilde seçilmesidir. Her bir imgeden binlerce öznitelik vektörü geldiği ve bu eşleştirmenin özünde, bu büyük vektör yığın setlerinin her bir elemanı arasında yapıldığı düşünülecek olursa yapılacak olan imge eşleştirme sürecinin ne kadar maliyetli olduğu ortaya çıkmaktadır. Dolayısıyla, etkili bir imge eşleştirme algoritması kullanılması şarttır. İlgili çalışmalar [20][16][18] incelendiğinde, yüksek boyutlu uzaylarda verimli bir şekilde eşleştirme yapılması için, kesin en yakın komşu (exact nearest neighbor) aramak yerine, yaklaşık en yakın komşuları (approximate nearest neighbor-ANN) aramak tercih edilmektedir. Bununla birlikte, imge eşleştirme çalışmalarında genellikle kullanılan ANN algoritmaları ise ağaç yapılı olarak kullanılan *k boyutlu ağaçlar* (*k-d ağaçlar*) yöntemidir. *k-d ağaçlar* yönteminde, büyük veri uzayı hiper-düzlemler ile bölünerek, ikili bir ağaç oluşturulur ve *k en yakın komşu* (*k nearest neighbor-KNN*) araması yapılır. Zaman içinde *k-d ağaçlar*'ın varyasyonları önerilmiştir [19][17]. Muja ve Lowe [12][13][14], çalışmalarında önerdikleri *çoklu rastgele k-d ağaçlar* (multiple randomized *k-d trees*) yönteminin imge eşleştirmedeki başarısını göstermişlerdir. Bütün bu bilgiler ışığında çalışmamızda, öznitelik vektörlerinin eşleştirilmesinde *çoklu rastgele k-d ağaçlar* yöntemini kullanan *Yaklaşık En Yakın Komşular için Hızlı Kütüphane* (Fast Library for Approximate Nearest Neighbors-FLANN) [14] algoritması kullanılmıştır.

### B. Yerel Betimleyicinin İmge Eşlemesine Uyarlanması

Çalışmamızda, *Hızlandırılmış Güçlü Öznitelikler* (Speeded-Up Robust Features-SURF) [21] yerel betimleyicisi kullanılmıştır. SURF yerel betimleyicisi çoğu diğer betimleyicilere göre hem daha hızlı hem de görüntüdeki dönüşümlere karşı daha dirençlidir. SURF betimleyicisi her imge üzerinde, imgeden imgeye değişen sayılarda ilgi noktası (1x64) bulmaktadır. FLANN algoritması ile de, bu ilgi noktası vektörleri seviyesinde eşleştirme yapılabilmektedir. Bu eşleştirmenin imgeler seviyesinde yapılabilmesi için, hangi ilgi noktasının hangi imgeden geldiğini düzenleyen bir indeksleme uygulanmıştır. Bu indeksleme sayesinde, girdi imge ile en çok sayıda ilgi noktası benzerliği taşıyan veri seti imgesi belirlenip, eşleştirilmiştir.

### C. İmgelerden Anlamsal Betimleyici Çıkarımı

Bu bölümde, anlamsal bilgi çıkarımı için imge bölütlemesinin nasıl yapıldığı ve bu anlamsal bilginin öznitelik vektörüne nasıl aktarıldığı anlatılmaktadır. Bölütleme işleminin ilk adımı olarak, imgedeki nesnelere çokgenler şeklinde etiketlememizi sağlayan imge etiketleme aracı LIBLABEL [22] kullanılmıştır. Böylece, 1571 imge için imgedeki her bir pikselin hangi nesne sınıfına ait olduğunu belirlenerek, imgenin önceden belirlenen nesne sınıflarına göre bölütlenmesi sağlanmıştır.

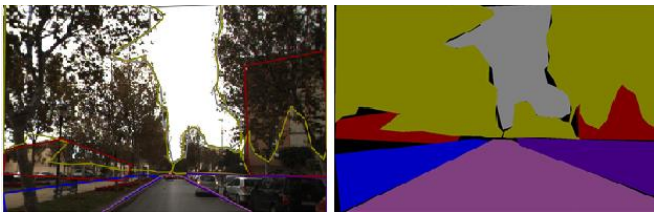
Veri setimiz üzerindeki bir imgeye LIBLABEL uygulayarak yapılan anlamsal bölütleme Şekil 2'de görülmektedir. Çalışmamızda anlamsal bilginin vektöre çevrilmesinde 7 nesne sınıfı (*Bina, Araba, Yol, Kaldırım, Gökyüzü, Ağaç, Duvar*) hesaba katılmıştır. Bu nesne sınıfları, Malaga veri setimizde çok yer aldığı ve dolayısı ile anlamsal bilginin aktarılmasına diğer nesne sınıflarına nazaran daha çok katkı sunacağı için tercih edilmiştir.

Elde edilen anlamsal bölütleme bilgisini öznitelik vektörü haline getirmek için, her bir nesne sınıfına atanan piksellerin miktarının bütün piksellere olan oranları  $[1 \times 7]$ 'lik vektöre gömülmüştür. Böylece, bu vektörün her bir hücresinde 7 nesne sınıfına ait olan tekrar bilgisi 0-1 arasında değişen ondalık sayılar ile kaydedilmiştir. Bu işlem bütün imgeye uygulandıktan sonra, aynı işlem imgenin 4 eşit kare parçaya ayrılmış alt bölümleri üzerinde tekrar edilmiş ve her 4 parçadan yine  $[1 \times 7]$ 'lik vektörler elde edilmiştir. Böylece, imgedeki nesnelere göre olan pozisyonlarının ayırt edici özelliği temsil edilerek; örneğin binanın sağda, ağaçların solda olduğu sahneden tam tersi sahneyi ayırt edilebilme yeteneği sağlanmıştır. Sonuç olarak imgenin bütününden ve 4 alt parçasından gelen  $[1 \times 7]$ 'lik vektörler, Şekil 3'te de tasvir edildiği gibi arka arkaya eklenerek bütün bir imge  $[1 \times 35]$ 'lik anlamsal öznitelik vektörü ile ifade edilmiştir.

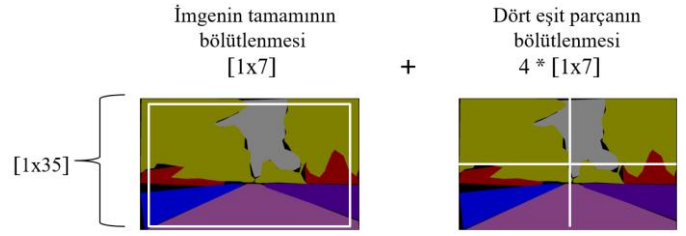
### D. Anlamsal ve Yerel Betimleyicilerin Birleştirilmesi: Melez Yöntem

Önceki bölümlerde anlatıldığı üzere, imgeleri yerel betimleyici ve anlamsal betimleyici vektörleri ile ifade ettikten sonra, her iki yöntem de FLANN algoritması ile uyumlu hale getirilerek sırası ile yerel betimleyici tabanlı ANN (YB-ANN) ve anlamsal betimleyici tabanlı ANN (AB-ANN) imge tabanlı eşleştirme yöntemleri oluşturulmuştur. Bu yöntemlerin ikisinde de  $k=10$  olarak çalıştırılmıştır, başka bir deyişle verilen sorgu imgesine karşılık en yakın 10 komşu imge sonuç olarak döndürülmüştür.

YB-ANN ve AB-ANN yöntemlerinin güçlü yönlerinden yararlanmak için, bu iki yöntem karar aşamasında birleştirilerek, Melez-ANN yöntemi oluşturulmuştur (Şekil 4). Her iki yöntemden gelen 10 en yakın komşu imge sonuçları, sonuç imgesinin sorgu imgesine olan uzaklık değerleri ile eşit olacak



Şekil 2. LIBLABEL ile anlamsal bölütleme uygulanmış veri seti imgesi.



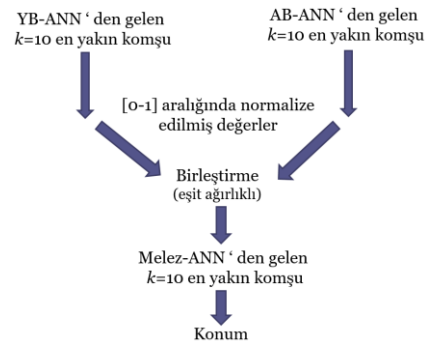
Şekil 3. Bölütlenmiş bütün bir imgeden çıkarılan ve 7 nesne sınıfının yüzdesel dağılımını betimleyen anlamsal öznitelik vektörü  $[1 \times 35]$ .

şekilde ağırlıklandırılarak (Şekil 4) o imgenin nihai uzaklığı belirlenmektedir. FLANN algoritması, döndürülen en yakın imge ile sorgu imgesi arasındaki uzaklık değerini Öklid uzaklığı cinsinden sağlamaktadır. Ayrıca, bir yöntemin döndürdüğü en yakın komşular arasında yer alıp diğer yöntem tarafından döndürülmeyen imgelerin uzaklık değeri 0 olarak kabul edilerek ağırlıklandırmaya dâhil edilir. Sonuç olarak, her iki yöntemden gelen toplam 20 sonuç, sorgu imgesine olan uzaklıklarına göre ağırlıklandırılarak yakından uzağa doğru sıralanır ve bu listedeki ilk 10 imge Melez-ANN yönteminin döndürdüğü sonuçlar olarak kabul edilir.

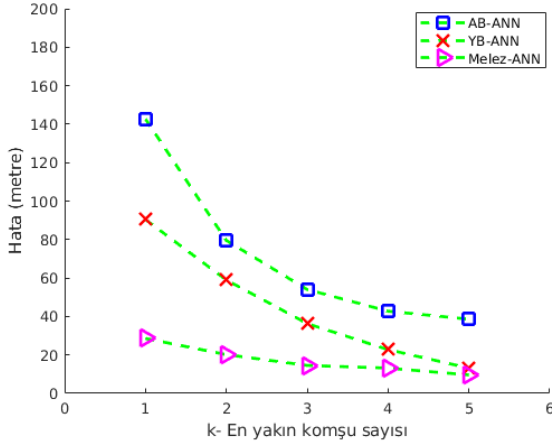
### IV. DENEYSEL SONUÇLAR

Eğitim seti (1048 imge) üzerinde, bütün imgelerin YB ve AB öznitelikleri bir araya getirilerek indekslenmiş, daha sonra oluşturulan bu öznitelikler matrisi FLANN algoritmasına verilerek eşleştirme modeli oluşturulmuştur. Daha sonra, test imgeleri (523 imge) için de aynı yöntemle öznitelik matrisleri oluşturulmuştur. Böylece, FLANN algoritması her bir test imgesi için oluşturulan model üzerinden arama yapmakta, 10 en yakın komşu imgeyi sonuç olarak döndürmektedir. Daha sonra, her iki yöntem için de uzaklık değerleri normalize edilerek Bölüm III. D'de ayrıntılı olarak anlatılan birleştirme yöntemi uygulanmış, Melez-ANN yöntemi için de 10 en yakın komşu elde edilmiştir.

Çalışmamızda önerilen üç imge tabanlı konumlandırma yönteminin (YB-ANN, AB-ANN ve Melez-ANN) veri seti üzerindeki başarılarının değerlendirilmesi için GPS tabanlı metrik bir hata hesaplama yöntemine ihtiyaç duyulmuştur. Bunun için GPS bilgileri radyan biriminde verilen enlem ve boylam değerleri ile küre üzerindeki iki nokta arasındaki büyük daire uzaklığı Haversine [23] formülü kullanılarak metre cinsinden hesaplanmıştır. Deney sonuçlarının (Şekil 5) değerlendirilmesinde, yöntemlerden elde edilen 10 imgeden ilk 5'i hesaba katılmıştır. Hata hesaplanmasını bir örnekle açıklayacak olursak: Şekil 5'te görüleceği üzere  $x$  eksenindeki  $k$



Şekil 4. Melez-ANN yönteminin oluşumu



Şekil 5. Melez-ANN yönteminin, diğer iki yöntemle karşı üstünlüğünün GPS tabanlı hata hesaplaması ile gösterimi.

komşu sayısı değeri 3 iken, döndürülen ilk üç en yakın komşudan hata değeri en düşük olan baz alınmıştır. Böylece, döndürülen en yakın k komşu arasından sorgu imgesine en yakın olanı doğru sonuç olarak kabul edilmiştir.

Karşılaştırılan üç yöntemin sonuçları Şekil 5'te tasvir edilmiştir, 5 en yakın komşuluk değerlerine karşılık gelen GPS tabanlı hatalar metre biriminde çizdirilmiştir. Şekilden açıkça görüldüğü üzere, YB-ANN yöntemi AB-ANN yöntemine göre daha iyi iken, anlamsal ve yerel özniteliklerin birleştirildiği Melez-ANN yöntemini diğer iki yöntemle karşılaştırıldığında da daha başarılı olmuştur.

## V. SONUÇLAR VE DEĞERLENDİRME

Bu çalışmada temel olarak, anlamsal bilginin kullanılmasının imge tabanlı konumlandırmaya olan etkisi gözlemlenmiştir. Bu amaç doğrultusunda, insansız sürüş senaryosunu yansıtan, şehir merkezi araç içi görüntüleri ve GPS verilerini içeren, Malaga kentsel veri seti kullanılmıştır. Veri seti imgeleri, önce yerel betimleyici (YB) kullanılarak, daha sonra da anlamsal betimleyici (AB) kullanılarak ifade edilmiştir. Önerilen bu iki yöntemle konumlandırma yapılmış, ardından bu iki yöntemi eşit ağırlıklı olarak birleştiren melez yöntemin konumlandırma başarısı ölçülmüştür.

Sonuç olarak bu üç yöntemin konumlandırma performansları, GPS tabanlı uzaklık hesabına dayalı hata hesaplama yöntemi ile kıyaslanmıştır. Önerilen melez yöntemin diğer önerilen iki yöntemden daha başarılı olduğu ve anlamsal ifadenin imge tabanlı konumlandırmaya olan olumlu etkisi deneysel sonuçlarla gösterilmiştir.

Elde edilen sonuçlar ışığında, önerilen yöntemlerin daha fazla çevresel değişikliğin olduğu farklı araç konumlandırma veri setleri üzerinde, farklı yerel ve anlamsal betimleyiciler kullanılarak denemesi gelecek çalışmalar olarak görülmektedir.

## KAYNAKLAR

[1] Lowry, S., Sünderhauf, N., Newman, P., Leonard, J. J., Cox, D., Corke, P., & Milford, M. J. Visual place recognition: A survey. *IEEE Transactions on Robotics*, 32(1), 1-19, 2016.

[2] Camposco, F., Cohen, A., Pollefeys, M., & Sattler, T. Hybrid Camera Pose Estimation. In *CVPR* (pp. 136-144), 2018.

[3] Sattler, T., Torii, A., Sivic, J., Pollefeys, M., Taira, H., Okutomi, M., & Pajdla, T. Are Large-Scale 3D Models Really Necessary for Accurate Visual Localization?. In *2017 IEEE Conference on CVPR*.

[4] D. Lowe, "Object recognition from local scale-invariant features," in *IEEE Conference on ICCV*, 1999, vol. 2, pp. 1150-1157.

[5] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *ECCV*, Springer, 2006, pp. 430-443.

[6] R. Arandjelović and A. Zisserman. DisLocation: Scalable descriptor distinctiveness for location recognition. In *Proc. ACCV*, 2014.

[7] Mousavian, A., Kosecka, J., Lien, J.M., Semantically Guided Location Recognition for Outdoors Scenes, *ICRA*, 2015.

[8] Toft, C., Stenborg, E., Hammarstrand, L., Brynte, L., Pollefeys, M., Sattler, T., Kahl, F. Semantic Match Consistency for Long-Term Visual Localization, *ECCV* 2018.

[9] Schönberger, J.L., Pollefeys, M., Geiger, A., Sattler, T., Semantic Visual Localization, *CVPR* 2018.

[10] Singh, G., Kosecka, J., Acquiring Semantics Induced Topology in Urban Environments, *ICRA* 2012.

[11] Blanco-Claraco J.L., Moreno-Dueñas F.A., Gonzalez-Jimenez J., "The Málaga Urban Dataset: High-rate Stereo and Lidars in a realistic urban scenario", *IJRR*, Feb, 2014, vol. 33, no. 2, 207-214.

[12] Muja, M., and Lowe, D. G., Fast approximate nearest neighbors with automatic algorithm configuration. In *VISSAPP*, 2009, (pp. 331-340).

[13] Muja, M., and Lowe, D. G., Scalable nearest neighbor algorithms for high dimensional data. *PAMI*, 2014, *IEEE Transactions on*, 36.

[14] Muja, M., and Lowe, D. G., FLANN- Fast Library for Approximate Nearest Neighbors, 2017, <https://www.cs.ubc.ca/research/flann/>, (Access Date: February, 8, 2019)

[15] Majdik AL, Verda D, Albers-Schoenberg Y, Scaramuzza D., Air-ground matching: appearance-based GPS-denied urban localization of micro aerial vehicles. *J Field Robot.*;32(7):1015-39, 2015,

[16] Andoni, A. and Indyk, P., Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions. *Commun. ACM*, 2008, 51(1):117-122,.

[17] Jo J, Seo J, Fekete JD, A Progressive k-d tree for Approximate k-Nearest Neighbors, *Workshop on Data Systems for Interactive*, 2017,

[18] Bentley J. L., Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 1975, 18(9):509-517.

[19] Silpa-Anan C. and Hartley R., Optimised kd-trees for fast image descriptor matching. In *2008 IEEE Conference on CVPR*, pp. 1-8, June.

[20] Beis J. S. and Lowe D. G., Shape indexing using approximate nearest neighbour search in high-dimensional spaces. In *CVPR*, 1997.

[21] Bay, H., Tuytelaars, T., & Van Gool, L. Surf: Speeded up robust features. In *ECCV 2006*. (pp. 404-417). Springer, Berlin, Heidelberg.

[22] Geiger, A., Lauer, M., Wojek, C., Stiller, C., & Urtasun, R.. 3d traffic scene understanding from movable platforms. *IEEE transactions on pattern analysis and machine intelligence*, 2014, 36(5), 1012-1025.

[23] Inman, James. *Navigation and Nautical Astronomy: For the Use of British Seamen* (3 ed.). London, UK: W. Woodward, C. & J. Rivington. Retrieved 2015-11-09, 1835.

[24] Jegou, H., Perronnin, F., Douze, M. Sanchez, J., Perez, P., Schmid, C., Aggregating local image descriptors into compact codes. *IEEE PAMI*, 34(9):1704-1716, 2012.