# İÇERİK TABANLI UYDU GÖRÜNTÜ ARAMA MOTORU

*Kaan KALKAN[1], Zeynep GÖKCE AKER[1], Recep CAN[1]*

TÜBİTAK Uzay Teknolojileri Araştırma Enstitüsü, Ankara, kaan.kalkan@tubitak.gov.tr, zeynep.gokce@tubitak.gov.tr, recep.can@tubitak.gov.tr

## ÖZET

*Bu çalışmada, APSCO (Asya-Pasifik Uzay İşbirliği Organizasyonu) bünyesinde geliştirilen DSSP (Data Sharing Service Platform) verileri kullanılarak, bu platformdaki uydu verilerinin bir grafik arayüzü üzerinden kullanıcının seçtiği alanlara benzer alanların olduğu görüntülerin bulunduğu görüntüleri seçen bir arama motoru geliştirilmesi amaçlanmaktadır. DSSP platformundan alınan veriler daha önce sisteme entegre edilerek küçük karolama alanlarına bölünecektir. Bu alanlar temel alınarak, kullanıcıdan ekranda bu karolardan herhangi birinin seçilmesi beklenecektir. Geliştirilen sistem, sisteme entegre edilen, tüm veriler içerisindeki bu karoya benzer karoları içeren uydu görüntülerini ve bu karoların konumlarını tespit edecektir. Bu aşamada sistem seçilen karoya benzer alanları yapay öğrenme yöntemleri ile tespit ederek kullanıcıya en benzer alanların geri dönüşünü sağlayacaktır. Bu alanlar içerisinde binaların olduğu, ormanların olduğu, rüzgar tribünü veya güneş enerji santrallerinin olduğu karolar gibi örneklendirilebilir.*

**Anahtar Sözcükler:** İçerik tabanlı görüntü arama, derin öğrenme, uydu görüntüsü

## ABSTRACT

## CONTEXT-BASED SATELLITE IMAGE SEARCH ENGINE

*This study aims to develop a search engine that selects images similar to the areas selected by the user through a graphical interface using DSSP (Data Sharing Service Platform) data shared within APSCO (Asia-Pacific Space Cooperation Organization). The data received from the DSSP platform will be previously integrated into the system and divided into small image tiles. The user will be expected to select any of these tiles on the screen based on these areas. The developed system will detect satellite images that contain tiles similar to this tile in all data and the positions of these tiles. At this stage, the system will detect the areas similar to the selected tile with artificial learning methods and provide the return of the most similar areas to the user. These areas can be exemplified as tiles with buildings, forests, wind turbines or solar power plants.*

**Keywords:** context-based image retrieval, deep learning, satellite image.

## 1. GİRİŞ

Uydu görüntüleri askeri, meteoroloji, tarım, lojistik gibi çeşitli alanlar için zengin bilgi kaynakları içerir. Ancak Uydu görüntüleri kritik ve pahalı bir kaynaktır. Bunun yanında Landsat ve Sentinel gibi bazı açık erişimli uydular ve hükümetler tarafından artan uydu fırlatmaları ile erişilebilir hale gelmiştir. Uydu görüntülerinde içerik tabanlı görüntü erişim (CBIR) sistemleri, afet kurtarma, ekoloji tahmini, meteorolojik analiz ve hedef tespiti gibi birçok uygulamada önemli bir rol oynamaktadır. Temsili görsel öznitelikleri kullanarak büyük bir uydu görüntüsü veri tabanından sorgu görüntüsüne en uygun görüntüleri bulmayı amaçlar. Tipik bir CBIR sistemi üç ana modülden oluşur: öznitelik çıkarımı, öznitelik indeksleme ve öznitelik benzerlik ölçümü.

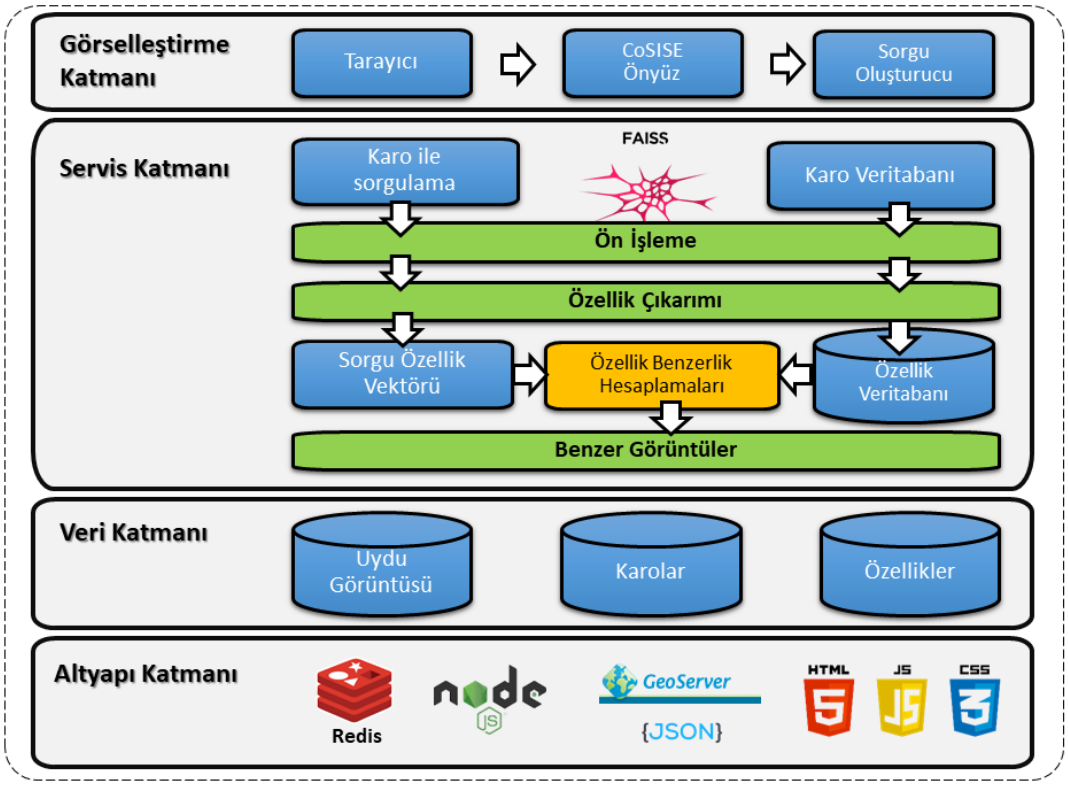
**a) Öznitelik Çıkarma**: Görüntüler, gelenekselveya derin öğrenme tabanlı öznitelik çıkarım yöntemleriyle elde edilmiş öznitelikler kullanılarak temsil edilir. Temsili öznitelikler, düşükten yüksek seviyelere kadar anlamsal bilgi sağlar.
**b) Öznitelik İndeksleme:** Veritabanı, özellikle büyük ölçekli görüntü veri kümelerinde, görüntülere erişimi hızlandırmak için her görüntünün özniteliklerini yapılandırılmış bir şekilde saklar.
**c) Benzerlik Ölçümü:** Bir benzerlik ölçüsü kullanarak sorgu görseline göre veritabanındaki en benzer görselleri sağlar. Bu metrik, iki aday görüntü arasında bir puan verir ve en benzer görüntüler bu puana göre seçilir.

Bu çalışma, Gaofen-2 Uydu görüntüleri için öznitelik çıkarma adımında en son teknoloji metodolojileri kullanarak uydu görüntüleri üzerinde görüntü benzerlik aramasını amaçlamaktadır. Her bir öznitelik vektörü bir görüntüyü temsil etmektedir. Veritabanı, ham dijital görüntülerin yüksek boyutlu sayısal değeri yerine tüm görüntüler için bu öznitelik vektörlerini depolar. Bir görsel arama sisteminde bir sorgu görüntüsü verildiğinde, benzer görseller, mesafeye dayalı benzerlik metrikleri kullanılarak aranır. En iyi k adet benzer görüntü, sırayla görüntüler arasından alınır.

Ana aşama olan öznitelik çıkarımına dayalı olarak, uydu görüntüsü benzerlik arama görevi için bir derin öğrenme modeli önerilmiştir. NetVLAD (Arandjelovic, 2013, 2016) en çok tercih edilen temsil modeli olduğundan, öznitelik çıkarma ağı olarak, önceden eğitilmiş model Patch-NetVLAD'ın NetVLAD kısmı kullanılmaktadır. Zaman verimliliği

ile ilgili olarak, Gaofen-2 Uydu görüntülerinin çok sayıda görüntü karosu arasında görüntü arama için Faiss yöntemi önerilmiştir.

Önerilen sistem, uydu görüntülerden oluşan bir veritabanı üzerinde görsel bir arama yapmak için bir platform olarak tanımlanabilir. Bu görsel arama motoru, bir örnek karo üzerinden (ör. tarım alanı, kentsel alan, vb.) içerik göreliliğine dayalı olarak kapsamlı bir veri tabanından benzer içeriğe sahip alanları bulmayı amaçlamaktadır. Çalışma iki ana süreçten oluşur: öznitelik çıkarma süreci ve benzerlik eşleştirme süreci. Öznitelik çıkarma işlemi sırasında görüntülerden derin öznitelik çıkarmak için uzaktan algılama görüntü veri kümesi üzerinde bir Evrişimsel Sinir Ağı (CNN) mimarisi eğitilecektir. Daha sonra, bu eğitilmiş model, tüm sorguların ve veri kümesi görüntülerinin özniteliklerini çıkarmak için kullanılacaktır. Öznitelikler, verileri ve hesaplama gereksinimlerini azaltmak için ikili değerlere dönüştürülür. Ölçeklenebilir bir veritabanı hizmeti kullanılarak çıkarılan ikili öznitelikler üzerinde doğrudan ve kesin bir arama yapılacaktır. Önerilen sistem, uydu görüntülerinde içerik tabanlı benzerlik sorgusu için hızlı uzaktan algılama arama motoru hizmeti sağlayabilir. Bu sistem, veritabanından hızlı ve ilgili bir arama kullanarak dünya yüzeyi üzerinde neredeyse gerçek zamanlı görsel arama sağlar. Sistemin iş akışı detayları Şekil 1'de verilmiştir.




**Şekil 1.** Sistemin bütüncül mimarisi

## 2. YÖNTEM

Bu bölüm, önerilen metodolojimizi ayrıntılı olarak sunmaktadır. Öznitelik çıkarma aşamasında, öncelikle tercih edilen derin öğrenme modelleri ve işlemleri açıklanmıştır. Daha sonra karo arama algoritmaları iki yaklaşım kullanılarak açıklanmıştır: Kaba Kuvvet Arama (Brute-Force) ve Faiss Arama.

### 2.1 Öznitelik Çıkarma

Öznitelik çıkarma ham görüntüyü daha kapsamlı, kompakt ve temsili sayısal değerlere dönüştürmektir. Derin öğrenmenin ortaya çıkmasıyla birlikte, Evrişimsel Sinir Ağı, öznitelik çıkarma/görüntü temsili, sınıflandırma, tanıma vb. işlemlerinde önemli rol almıştır. Uzaktan algılamada derin öğrenme yöntemleri giderek daha fazla kullanılır hale gelmiş, anlamlı ve başarılı sonuçlar üretebilir olmuştur (Zhu, 2017). Bu motivasyonla, bu projede öznitelik çıkarma aşaması için Derin Öğrenme metodolojisi önerilmiştir.

Görüntü karoları temsilleri, yer tanıma ve veri erişimi görevleri için geliştirilmiş popüler derin öğrenme modellerinden biri olan NetVLAD kullanılarak elde edilmiştir. NetVLAD mimarisi, temel CNN modeli (önceden eğitilmiş VGG16, Conv5 katmanından sonra kesilir) üzerine NetVLAD katmanı (çeşitli konvolüsyon katmanları ve softmax içeren VLAD çekirdeklerini oluşur) eklenerek oluşturulmuştur.

Girdi olarak önceden eğitilmiş bir derin modele bir görüntü karosu verildiğinde, NetVLAD katmanından global NetVLAD özniteliklerini çıkarırız. Global öznitelik vektörlerine PCA uygulandıktan sonra her bir görüntü, Patch-NetVLAD'te olduğu gibi global 512-dim öznitelik vektörü ile temsil edilir ve saklanır (Hausler, 2021).

## 2.2 Karo Arama

Görüntü arama, İçerik Tabanlı Görüntü Erişim Sisteminin ikinci aşamasıdır. Öznitelik Çıkarma aşamasında, her bir görüntü 512-dim vektörler halinde veritabanında saklanır. Çalışma veritabanındaki görüntü karoları arasında en çok benzer görüntü karolarını elde etmeyi (erişmeyi) amaçlar. Benzer görüntüler, öznitelik vektörleri kullanılarak en yakın komşu yaklaşımı kullanılarak uzaklığa dayalı benzerlik metrikleri kullanılarak aranır. Bu sistemde iki farklı arama mekanizması uygulanmaktadır. Zaman verimliliği açısından bu projenin arama sistemi için Faiss tercih edilmiştir.

**Kaba Kuvvet Arama**
En yakın komşu aramasında naif yaklaşım Brute-Force (Kaba Kuvvet) algoritmasıdır. Görüntünün her bir öznitelik vektörü için veri setindeki tüm görüntüler arasındaki uzaklık hesaplanır. Literatürde, görüntüler arasındaki benzerliği hesaplamak için Öklid, Manhattan, Minkowski ve Hamming Uzaklıkları gibi çeşitli uzaklık ölçütleri kullanılmaktadır. Bu projede birincil erişim görevi için Öklid mesafesi tercih edilmiştir. Küçük veri setleri için **Kaba Kuvvet Arama** algoritması tercih edilebilirken, büyük veri setlerinde oldukça maliyetli olmaktadır.

**Faiss Arama**
Görüntüyü bir veritabanında hızlı aramak için Faiss, Facebook AI Research Group tarafından geliştirilmiştir (https://faiss.ai/). Faiss, kümeleme yaklaşımı kullanarak verimli arama sağlayan bir kütüphanedir. Faiss kümeleri, sorgu görüntü karosunu tüm veri kümesindeki her bir görüntü karosu ile karşılaştırmak yerine, önce tüm veritabanındaki görüntü karo vektörlerini arar ve ardından ilgili kümedeki sorgu vektörünü arar. Kaba kuvvet algoritmanın aksine Faiss, arama sistemini hızlandırarak daha hızlı sonuçlara ulaşmamızı sağlar.

# 3. DENEYLER

Bu bölümde, deneylerin metrik ve görsel sonuçları gözlemsel yorumlarla birlikte sunulmaktadır. Veri setleri öncelikle detaylarıyla birlikte açıklanmaktadır. Sonrasında içerik tabanlı görüntü arama motorumuzun doğruluk analizi için kullanılan metrikler tanımlanmıştır. Son olarak deneysel detaylar, derin öğrenme modeli ve performans sonuçlarıyla birlikte sunulmuştur.

### 3.1. VERİ

Deneylerin kapsamlı bir değerlendirmesi için çeşitli veri kümelerini kullanılmıştır: UCMerced ve Gaofen-2.

**UCMerced:** Sınıf başına 100 görüntü içeren 21 sınıflı bir arazi kullanımı uzaktan algılama görüntü veri kümesidir. Görüntüler, ülke çapındaki çeşitli kentsel alanlar için USGS Ulusal Harita Kentsel Alan Görüntüleri (National Map Urban Area Imagery) koleksiyonundan alınmıştır. Bu RGB görüntülerin piksel çözünürlüğü 0,3m'dir.

**Gaofen-2:** CNSA-APSCO Veri Paylaşım Hizmeti Platformu (DSSP) tarafından sunulan Gaofen-2 (GF-2) uydu görüntüleri, görüntü benzerlik arama deneyleri için kullanılmıştır. Veri kümesi her biri 256x256 boyutunda ve 3 bantlı (R-G-B) JPG görüntüsü olan 861.954 görüntü karosundan oluşur. (Şekil 2)

**Şekil 2.** Örnek GF-2 Uydusu Karoları

### 3.2. DOĞRULUK ANALİZİ (METRİK)

Görüntü Erişim sisteminde, doğruluk analizi için mAP (Ortalama Hassasiyet - mean Average Precision) metriği kullanılmıştır. Doğruluk analizi için etiketli veri kümesine ihtiyaç duyulduğundan dolayı UCMerced veri kümesi bu deney için kullanılmıştır. mAP metriğini net bir şekilde ifade edebilmek için aşağıda sırasıyla Kesinlik (Precision), Average Precision (Ortalama Kesinlik) ve son olarak mAP metrik formülleri tanımlanmıştır. Metrik hesaplama yapılırken uzaklık tabanlı skora göre sıralanmış olan görüntü listesinin en benzer K görüntüyü ele alır ve mAP sonuçlarını hesaplanmaktadır.

**Kesinlik (P@K):** Bu metrik bir sorgu görüntüsünden elde edilen en benzer K görüntü arasından kaç tane görüntünün ilgili olduğunu ölçer. İlgili görüntü sorgu görüntüsüyle aynı etikete sahip olan görüntü demektir.

$$P@K = \frac{G@K}{K}$$

$$G@K: K\ görüntü\ içerisindeki\ ilgili\ görüntü\ sayısı$$

**Ortalama Kesinlik (AP@K):** Bir sorgu görüntüsü verildiğinde model tarafından elde edilen ilgili görüntülerin yüksek doğrulukta sıralanıp sıralanmadığını ölçer. Ortalama Kesinlik metriğini, bir sorgu görüntüsünde farklı K değerleri için P@K değerine bağlı olarak hesaplanır.

$$AP@K = \sum_{k=1}^{K} \frac{P@k * rel(k)}{K}$$

$$rel(k) = 1\ eğer\ k.görüntü\ ilgili\ ise$$

$$rel(k) = 0\ eğer\ k.görüntü\ ilgili\ değil\ ise$$

**Mean Average Precision (mAP@K):** Bütün sorgu görüntülerinde ortalama AP@K değerini ölçer. N veri kümesindeki toplam sorgu görüntü sayısıdır.

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i$$

### 3.3. DENEY DETAYLARI VE DEĞERLENDİRME

Nitel ve nicel değerlendirmeler için UCMerced veri kümesi üzerinde yapılan deneyler sonrasında, ana veri kümesi olan Gaofen-2 üzerinde nitel değerlendirmeler için deneyler devam ettirilmiştir. Her iki veri kümesi için de aynı derin öğrenme model mimarisi ve aynı arama mekanizması kullanılmıştır. Veri kümelerindeki her bir karo için öznitelikler çıkarıldıktan sonra, bu öznitelik vektörleri üzerinde uzaklık tabanlı Faiss Arama algoritması uygulanmıştır. Öznitelik çıkarım aşamasında, veritabanındaki her bir karo 512-boyutlu görsel öznitelik vektörleri ile temsil edilmektedir. Bu öznitelik vektörleri, önceden eğitilmiş derin öğrenme modelinin, Patch-NetVLAD (Hausler, 2021), NetVLAD katmanındaki aktivasyon değerlerine karşılık gelmektedir. Veri kümelerinin büyüklüklerinden kaynaklı olarak, Gaofen-2 veri kümesinin öznitelik vektörleri ortalama 1,8 GB bellek gerektirirken, UCMerced ver kümesi için 4,5 MB gerektirir.

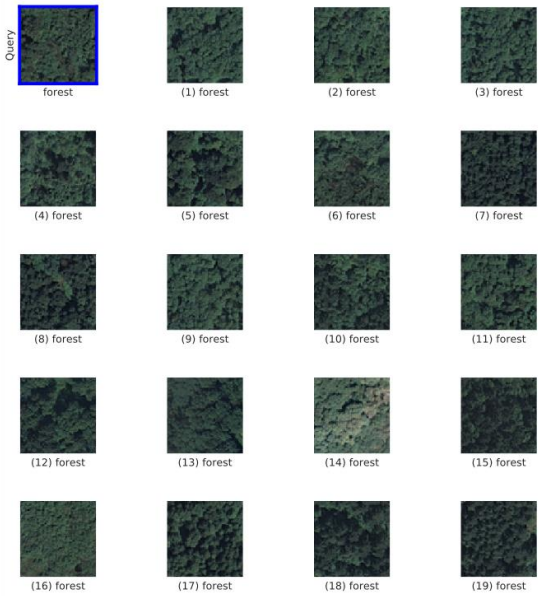
Faiss Arama algoritmasının ile GPU (GeForce RTX 2080) birlikte kullanılmasından sonra arama kısmı, Kaba Kuvvet Arama algoritmasına göre oldukça hızlanmıştır. Faiss algoritmasında parametre olarak, 50 küme merkezi (cluster centres), 10 hücre (cells) kullanılmıştır. Görüntü karo arama deneylerine göre, Kaba Kuvvet Arama algoritmasına göre Faiss Arama algoritması yaklaşık 250 kat daha hızlı benzer görüntü karolarını elde etmiştir. Her bir görüntü sorgusu için yaklaşık 861,954 görüntü karosu arasından ortalama 0.4 saniye içinde sonuç dönmektedir.

Görüntü Erişim modelinin niteliksel ve niceliksel performansını değerlendirmek için etiketlenmiş veri kümesi olarak UCMerced ve genel geçerliği kabul edilmiş olan Ortalama Hassasiyet (mAP) metriği kullanılmıştır. Doğruluk analizi için kullanılan metriğe göre, Görüntü Erişim sistemimiz UCMerced veri kümesi üzerinde tüm görüntüler içim %85,22 mAP@10 ve %80.05 mAP@20 değerlerine ulaşmaktadır.

Aşağıda UCMerced ve GF-2 veri kümelerinden bazı örnek görüntüler gösterilmektedir. (Şekil 3-7)

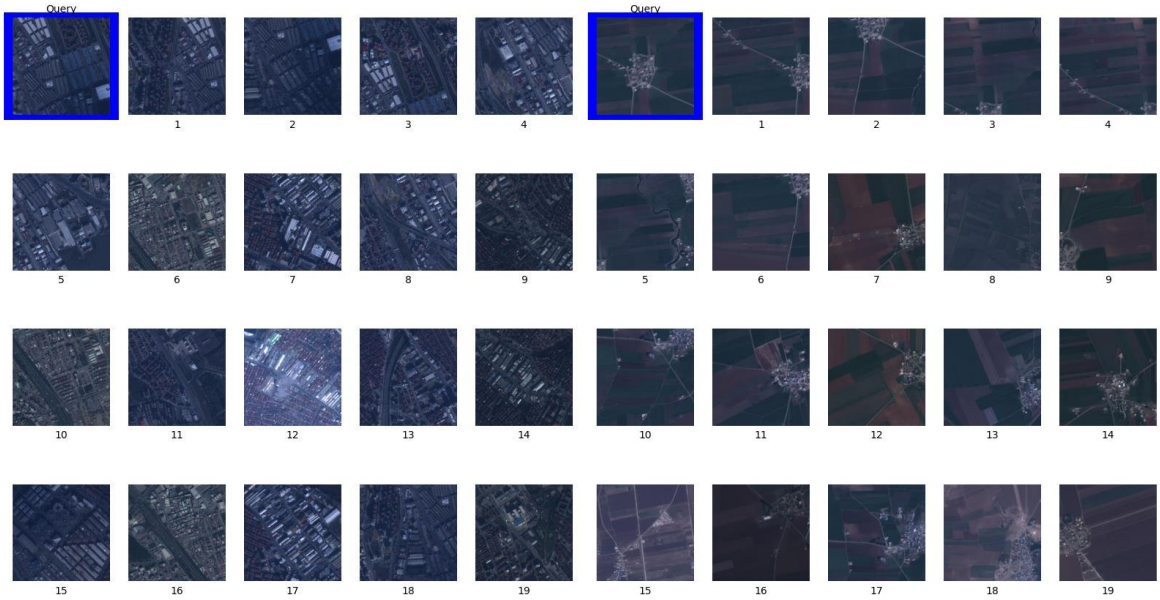
**UC-Merced Veri Setinden Örnekler**

**Şekil 3a)** Uçak araması sonuçları (en iyi 19) **b)** Karavan parkı araması

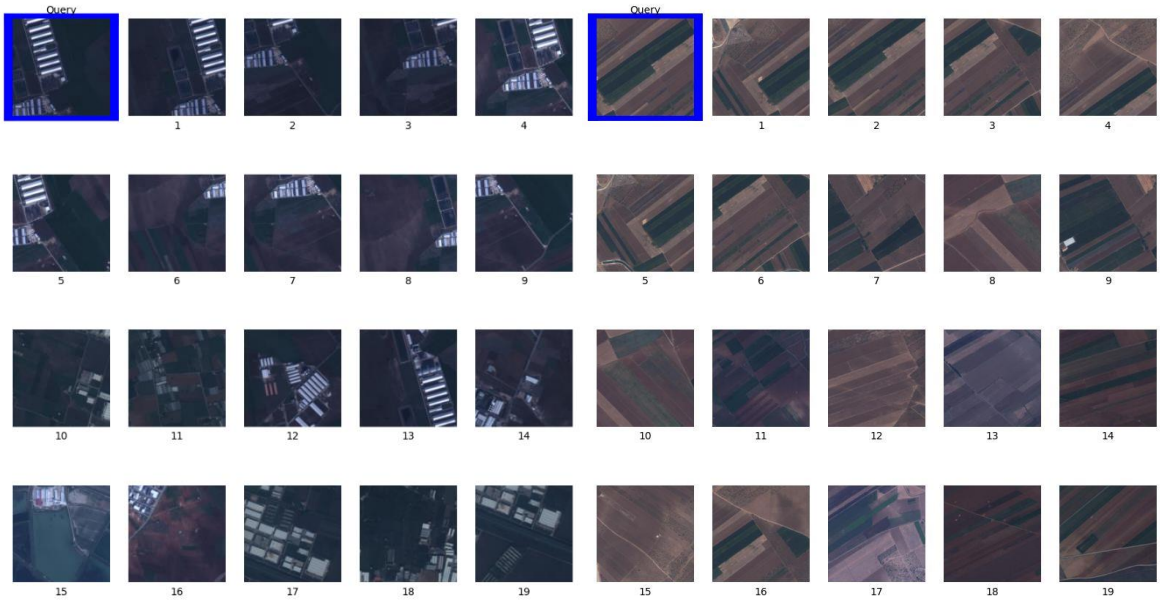



**Şekil 4.** Orman araması

**GF-2 Veri Setinden Örnekler**

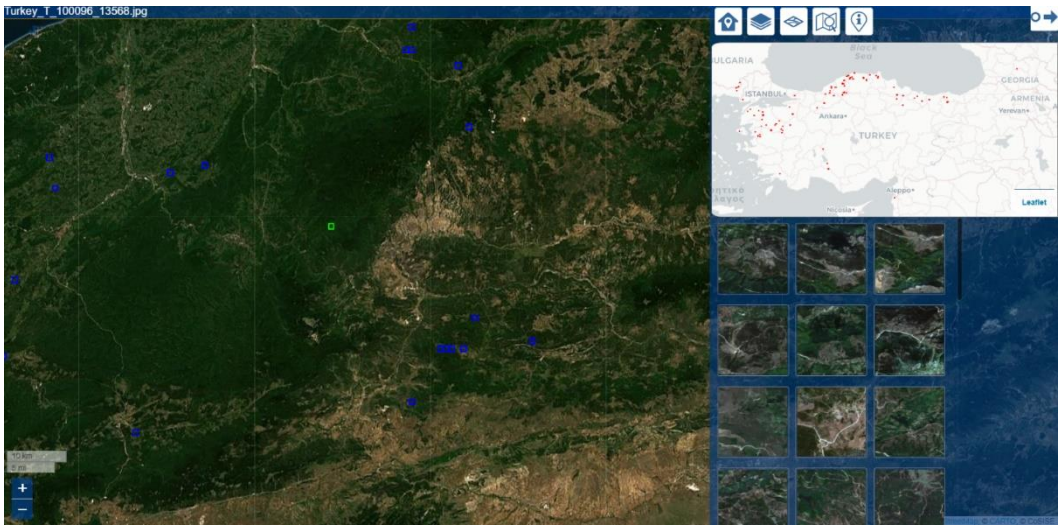



**Şekil 5a)** Kentsel alanda arama örneği **b)** Köy merkezinde arama örneği




**Şekil 6a)** Sera alanları arama örneği **b)** Benzer tarımsal alanları arama örneği

Geliştirilen sistem için, aranmak istenen alana tıklandığından benzer çıkan karo sonuçlarını ve lokasyonları gösteren bir arayüz geliştirilmiştir (Şekil 7).




**Şekil 7.** Geliştirilen sistem arayüzü

## 4. SONUÇLAR VE TARTIŞMA

Bu çalışma, öznitelik çıkarımı adımı için NetVLAD ve uzaktan algılama görüntüleri arama adımı için Faiss modelini temel alan bir İçerik Tabanlı Görüntü Arama Sistemi önermektedir. Her görüntü karosu derin öğrenme modelinin NetVLAD katmanını kullanan bir öznitelik vektörü ile temsil edilir. Görüntü arama, en yakın komşu yaklaşımına dayalıdır. Faiss uygulamasıyla kümeleme mekanizmasına dayalı olarak en iyi 100 benzer görüntü hızlı bir şekilde alınır. Önerilen model, Gaofen-2 ve UCMerced görüntü karolarında test edilmiştir. UCMerced veri seti, mAP metriği ile model performansını değerlendirmek için uygun şekilde kullanımıştır ve %85,22 mAP@10 ve %80.05 mAP@20'ye ulaşmıştır, Gaofen-2 nitel sonuçlar ve sistem entegrasyonu elde etmek için ana veri setimizdir. Bu çalışmada, ilk modeller ve deneyler sunulmaktadır, ilerleyen çalışmalarda, yeni iyileştirmeler araştırılacak ve yayınlanacaktır.

## TEŞEKKÜR

Yazarlar bu çalışmayı destekleyen ve verileri sağlayan APSCO (Asia-Pasific Space Coorperation Organization) ve TÜBİTAK Uzay Teknolojileri Araştırma Enstitisü'ne teşekkürlerini sunar.

## KAYNAKLAR

**Arandjelovic, R., Gronat, P., Torii, A., Pajdla, T., Sivic, J.,** 2016. NetVLAD: CNN architecture for weakly supervised place recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 5297-5307).

**Arandjelovic, R., Zisserman, A.,** 2013. All about VLAD. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (pp. 1578-1585).

**Hausler, S., Garg, S., Xu, M., Milford, M., Fischer, T.,** 2021. Patch-netvlad: Multi-scale fusion of locally-global descriptors for place recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 14141-14152).

**Zhu, X. X., Tuia, D., Mou, L., Xia, G. S., Zhang, L., Xu, F., Fraundorfer, F.,** 2017. Deep learning in remote sensing: A comprehensive review and list of resources. IEEE Geoscience and Remote Sensing Magazine, 5(4), 8-36.