

Kamufajlı Askerlerin Derin Öğrenme ile Segmentasyonu ve Kamufaj Desenine Göre Sınıflandırılması

İlkay Karatepe¹, Vasif Nabiyev²

^{1,2}Bilgisayar Mühendisliği Bölümü / Fen Bilimleri Enstitüsü, Karadeniz Teknik Üniversitesi, Türkiye

¹ORCID ID: <https://orcid.org/0000-0002-6627-1503>

²ORCID ID: <https://orcid.org/0000-0003-0314-8134>

(¹ilkaykaratepe@gmail.com, ²vasif@ktu.edu.tr)

Özet – Kamufaj, doğadaki canlıların kendilerini gizlemek için kullandığı bir tekniktir ve 19. yüzyılda uzun menzilli ateşli silahların yaygınlaşmasıyla askeri alanda da kullanılmaya başlamıştır. Çevre, iklim ve diğer faktörler göz önüne alındığında, farklı renk ve desenlerde kamufajlar geliştirilmiştir. Bu nedenle, ülkeler veya askeri birimler kendi ihtiyaçlarına uygun kamufaj desenleri geliştirmiş ve zamanla bu desenler, ulusal kimliklerinin bir parçası haline gelmiştir. Bu çalışmada 5 farklı askeri kamufaj ailesinden olan desen (Dijital, yaprak, puzzle, benekli, çöl) için kamufle askerlerin segmentasyonu ve sınıflandırılması amaçlanmıştır. Kamufaj veya desen sınıflandırması adı altında farklı çalışmalar literatürde yer almaktadır. Bu çalışmalar, kamufajlı nesnelerin segmentasyonu veya farklı türdeki kamufajlı nesnelerin sınıflandırılması şeklinde gerçekleştirilir. Çalışmamızdaki odak noktamız kamufajlı askerler olduğu için, derin öğrenme algoritmasının nesnelerin ana hatlarına değil, esas olarak kamufaj desenine göre sınıflandırılması beklenmektedir. Çalışmada 5 kamufaj ailesi için 1233 tane kamufajlı asker görüntüsü toplanmış ve poligonal etiketleme yapılmıştır. Desenler için farklı ülkelerden ortak veri seti oluşturulmuştur. Dijital desen için Türkiye, Azerbaycan, Çin ve Rusya'dan yaprak desen için Amerika'dan puzzle desen için Fransa'dan benekli desen için Almanya'dan çöl deseni için ise Irak'tan kamufajlı asker görüntüleri bulunmaktadır. Günümüzde nesne tespiti, bölütlemesi ve sınıflandırılması için yaygın olarak kullanılan Mask R-CNN algoritması, askeri kamufaj sınıflandırmasının çözümü için ele alınmış ve bu zorlu problemde derin öğrenme algoritmalarının önemi kanıtlanmıştır. Eğitim 0.008426 sınıflandırma kaybı ve 0.04284 maskeleyme kaybı ile sonuçlanmıştır.

Anahtar Kelimeler – Kamufaj, Derin Öğrenme, Mask R-CNN, Sınıflandırma, Segmentasyon

I. GİRİŞ

Kamufaj tasarımları, askeri operasyonların doğruluğunu artırmak için sürekli olarak yenilenmektedir. Günümüzde, askeri kamufaj, modern teknolojinin yardımıyla üretilmektedir ve daha karmaşık desenler ve renkler kullanılmaktadır. Bu tasarımlar, askerlerin sahada doğru bir şekilde saklanmalarını sağlarken, aynı zamanda askeri taktiklere de katkıda bulunmaktadır [1]. Askerler 1800'lerin ortalarına kadar parlak kıyafetler giyerdi [2]. Ancak, daha isabetli ve uzun menzilli ateşli silahların gelişmesiyle birlikte, askeri alanda ilk kamufaj 1848'de İngiliz Hint Ordusu'nun bir alayı tarafından sarımsı donuk uniformalar olarak

kullanılmaya başlandı [3]. 19. yüzyılın sonlarına doğru, İngiliz ordusu düz hâki renkli uniformalar benimsedi [4]. Birinci Dünya Savaşı'nda, Alman ordusu tarafından kullanılan "Fragmentation Camouflage" adlı desenli uniformalar, farklı renklerden oluşmaktaydı [4]. Kamufaj desenlerinin tasarımı 20. yüzyılın başlarında tasarımcılar ve sanatçılar tarafından gerçekleştirildi ve 1970'lerin sonlarında bilgisayar teknolojisinin gelişmesiyle birlikte kamufaj alanında kullanılmaya başlandı [4]. Günümüzde belirli özellikler etrafında gruplaşmış kamufaj desen aileleri mevcuttur. Çalışma kapsamında "dijital", "yaprak", "puzzle", "benekli" ve "çöl" kamufaj ailelerinden [5] alınan örnekler kullanılmıştır.

Farklı çalışmalar literatürde kamuflaj veya desen sınıflandırması için yapılmıştır. Doğan vd., AlexNet, Vgg16, Vgg19, ResNet50 ve GoogleNet gibi çeşitli derin öğrenme algoritmalarını kullanarak yaprak desenlerini sınıflandırmayı karşılaştırdılar [6]. Gupta vd., sisli, karlı, gece görüşlü çeşitli askerler ve ekipmanları algılamayı hedefleyen 22 sınıftan oluşan bir çalışma için 58 katmanlı Convolutional Neural Network (CNN) modeli özelleştirdiler [7]. Bayram vd., kamuflaj görüntüleri için Local Binary Patterns (LBP) kullanarak dokusal öznelikler çıkardılar ve bunları Yapay Sinir Ağları (YSA), Destek Vektör Makineleri (SVM) ve K-en Yakın Komşuluk Algoritması (KNN) ile sınıflandırdılar [8]. Yan [9]'da yaptığı çalışmada kamuflajlı nesnelerin segmentasyonu için CamoNet [10] isimli ağı kullanmıştır. Ömeroğlu vd., yüksek çözünürlüklü uydu görüntülerinde hangar tespiti yapmak için Mask R-CNN [11] algoritmasını kullanarak %85 ortalama kesinlik elde etmişlerdir [12]. Benzer şekilde, Amri vd. multispektral görüntülerde stadyum tespiti yapmak için YOLOv5 [13] ve Mask R-CNN algoritmalarını kullanmışlardır [14].

Bu çalışmada, bölgesel tabanlı nesne tespiti, bölütleme ve segmentasyon için tasarlanmış olan Mask R-CNN [11] algoritması kullanılarak 5 farklı askeri kamuflaj ailesi sınıflandırılmış ve kamuflajlı asker görüntüleri bölütlenmiştir.

II. MATERYAL VE YÖNTEM

A. Materyal

Bu çalışma için 5 farklı kamuflaj ailesi (Dijital, yaprak, puzzle, benekli, çöl) seçilerek kamuflajlı asker görüntüleri toplanıp veri seti olarak kullanılmıştır. “Dijital” desen ailesi için Türkiye-Azerbaycan, Rusya ve Çin askeri görüntüleri ortak kullanılmıştır. Her ne kadar bu ülkelerin kamuflajları farklı da olsa pixel tabanlı desen tasarımlarıdır. Bunun dışında “yaprak” deseni için Amerika’dan, “puzzle” deseni için Fransa’dan, “benekli” deseni için Almanya’dan ve “çöl” deseni için Irak askerlerinden görüntüler toplanmıştır. Fotoğrafların toplam sayısı 743’tür. Veri seti toplam 1233 adet kamuflajlı asker görüntüsünü kapsar. Fotoğraflar Mask R-CNN’nin girişi için 700x700 boyutlarına ayarlanmıştır. Eğitim dışında 115 asker görüntüsü de test için kullanılmıştır. LabelMe [15] aracıyla etiketlenmiş ve eğitime hazır hale getirilmiş

görüntüler bulunmaktadır. Desen ailelerine göre kullanılan asker görüntü sayıları Tablo 1’de ayrıntılı olarak verilmiştir. Şekil 1’de 5 farklı desen ailesine ait kamuflaj kumaş örnekleri gösterilirken, Şekil 2’de kullanılan 5 sınıfa ayrılmış kamuflajlı asker görüntülerinden örnekler sunulmaktadır.

Tablo 1. Çalışmada Kullanılan Görüntülerin Dağılımı

Kamuflaj Deseneri	Ülke	Eğitim	Test
Dijital	Türkiye-Azerbaycan	146	12
	Çin	199	18
	Rusya	143	21
Yaprak	Amerika	131	15
Puzzle	Fransa	158	18
Benekli	Almanya	105	7
Çöl	Irak	236	24
Toplam	1233	1118	115



Şekil 1. Kamuflaj Kumaş Örnekleri: Dijital (Türkiye-Azerbaycan, Çin, Fransa), Yaprak (Amerika), Puzzle (Fransa), Benekli (Almanya), Çöl (Irak)



Şekil 2. Kamuflajlı Asker Örnekleri: Dijital, Yaprak, Puzzle, Benekli, Çöl

B. Yöntem

Bu çalışma kapsamında, LabelMe aracılığıyla poligonal şekillerle etiketlenen görüntüler eğitime hazır hale getirilmiştir. Mask R-CNN algoritması kullanılarak 50000 iterasyondan geçirilen etiketli görüntüler eğitim için kullanılmış ve 5 farklı sınıf

için örnek bölütleme (instance segmentation) başarılı bir şekilde gerçekleştirilmiştir.

Eğitim sürecinde, sabit bir öğrenme oranı (learning rate- η) kullanılmamıştır. Bunun yerine, öğrenme oranı belirli adımlarda gamma (γ) çarpanıyla küçültülerek toplam kayıp (total loss) sabit bir değer alıncaya kadar eğitim işlemi devam ettirilmiştir. Sabit ve yüksek bir öğrenme oranı kullanmak, toplam hata oranının hızlı bir şekilde azalmasına neden olabilir ancak yerel minimum noktasının atlanmasına yol açabilir. Sabit ve düşük bir öğrenme oranı kullanmak, eğitim sürecini uzatacaktır, ancak aşırı öğrenmeye (overfitting) neden olabilecektir. Problemdaki sınıfların benzer özellikler göstermesi göz önüne alındığında, küçülen bir öğrenme adımı yaklaşımı tercih edilmiştir. Bu yaklaşım, gamma çarpanı kullanılarak gerçekleştirilmiştir [16].

Çalışmada, öğrenme oranı (η) 0.004 olarak başlatılmış ve ilk 20 bin iterasyonda her 500 iterasyonda %5 düşmesi için 0.95 gamma çarpanıyla çarpılarak güncellenmiştir. 20 bin iterasyondan sonra ise her 200 adımda %1 azaltacak şekilde 0.99 gamma çarpanıyla çarpılarak güncellenmiştir.

Mask R-CNN algoritması, Faster R-CNN [17] temelinde geliştirilmiş bölge tabanlı bir evrimsel sinir ağıdır. Mask R-CNN, örnek bölütlemesi (instance segmentation) yapmak için kullanılır. Bu yöntemde, nesne bölütlemesi için sınırlayıcı kutu (bounding box) kullanılır ve aynı zamanda nesnenin maskesi de tahmin edilir. Mask R-CNN ile yapılan maskeleyme işlemi, Faster R-CNN ile tahmin edilen sınırlayıcı kutudaki hedef nesne piksellerinin de kimliklendirilmesini sağlar.

Faster R-CNN, iki temel bileşenden oluşur: İlk olarak, Residual Network (ResNet) [18] gibi bir önceden eğitilmiş bir evrimsel sinir ağı (CNN) kullanılır. İkinci olarak, algılama bölgesi önerisi ağı (region proposal network-RPN) adı verilen bir ağ kullanılır. RPN, olası nesne konumlarını belirlemek için bir resimdeki çeşitli bölgeleri inceleyerek bölge önerileri üretir. RPN (Region Proposal Network) ile önceden tahmin edilen bölgeler arasından en yüksek kesişim birleşim oranına sahip olanlar, yani IoU (Intersection over Union) değeri en yüksek olanlar, ağdaki ROI (Region of Interest) olarak adlandırılan

kesişim bölgelerine gönderilir. RPN'den gelen sınırlayıcı kutu boyutları değişken olduğundan, ROI pooling adı verilen bir kırpma işlemi yapılır ve böylece sınıflandırıcının daha doğru çalışması için sabit bir boyuta dönüştürülür. ROI'de boyut dönüşümü yapıldıktan sonra, tahmini bölgenin sınıfı da belirlenir. Faster R-CNN'de, elde edilen sonuçlar tam bağlantılı katmana (fully connected layer) iletilir ve bu katmanda sınıf ve sınırlayıcı kutu bilgileri üretilir. Ancak Mask R-CNN'de, ROI'den sonra, FC layer ile paralel olarak maske bilgisi, semantik sınıflandırma yapılarak üretilir. Bu sayede, her bir ROI için hem sınıf hem de maske bilgisi elde edilmiş olur.

Mask R-CNN'nin genel yapısı Şekil 4'te gösterilir. Mask R-CNN'de, kayıp fonksiyonu (L) sınıflandırma kayıpları (L_{cls}), maskeleyme kayıpları (L_{mask}) ve çerçeve kayıplarının (L_{box}) toplamından oluşur [19]. Kayıp fonksiyonu Denklem (1) ile gösterilmiştir.

$$L = L_{cls} + L_{mask} + L_{box} \quad (1)$$

III. DENEYSEL ÇALIŞMALAR

Çalışmamız, Facebook Yapay Zekâ Araştırması (FAIR) tarafından geliştirilen Detectron2 [20] ile gerçekleştirilmiştir. Bu işlemci, Intel(R) Xeon(R) CPU @ 2.30GHz işlemcili, Tesla T4 ekran kartlı ve 12 GB birincil bellekli bir bilgisayarda gerçekleştirilmiştir. Eğitim için 5 farklı sınıftan toplam 743 fotoğraf toplanmıştır. Veri setinde toplam 1233 kamuflajlı asker bulunmaktadır. Veri setinde her bir sınıf, %90 eğitim ve %10 test olarak rastgele ayrılmıştır.

IV. BULGULAR

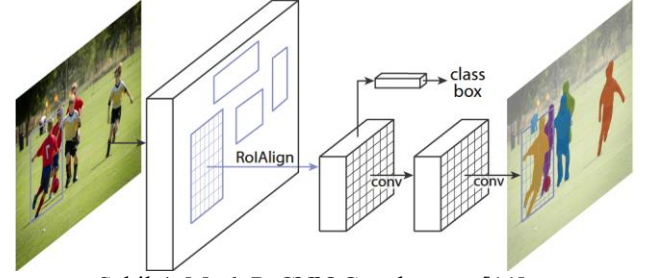
Modelin test görüntülerine uygulanması ile ortalama ~%97 doğrulukla tahmin ettiği görülmüştür. 50000 iterasyonla ulaşılan hata oranları Tablo 2'de gösterilmiştir. Eğitim sonunda oluşan eğitim metrikleri Şekil 5'te ve test görüntüleriyle elde edilen segmentasyon sonuçları ise Şekil 6'da verilmiştir. Ayrıca Şekil 3 gösteriyor ki yapılan çalışmada sadece kamuflajlı insanlar tespit edilmiştir.

Tablo 2. Kayıp Oranları

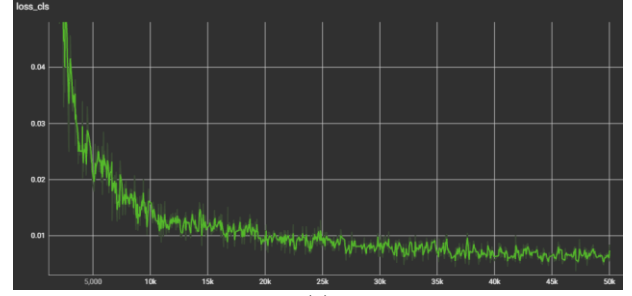
	Toplam Kayıp	Sınıflandırma Kaybı	Maskeleme Kaybı	Çerçeve Kaybı
En İyi Değer	0.051518	0.003698	0.03223	0.01559
En Son Değer	0.076306	0.008426	0.04284	0.02504



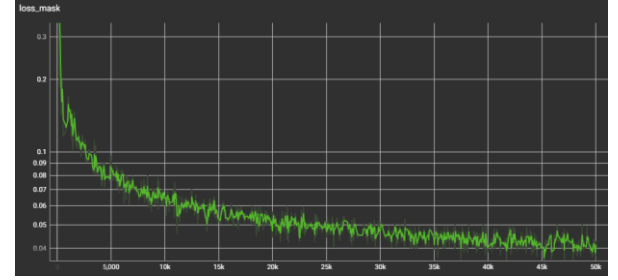
Şekil 3. Kamuflajlı İnsanların Segmentasyonu (Sadece kamuflaj olan insanlar algılanmaktadır)



Şekil 4. Mask R-CNN Genel yapısı [11]



(a)



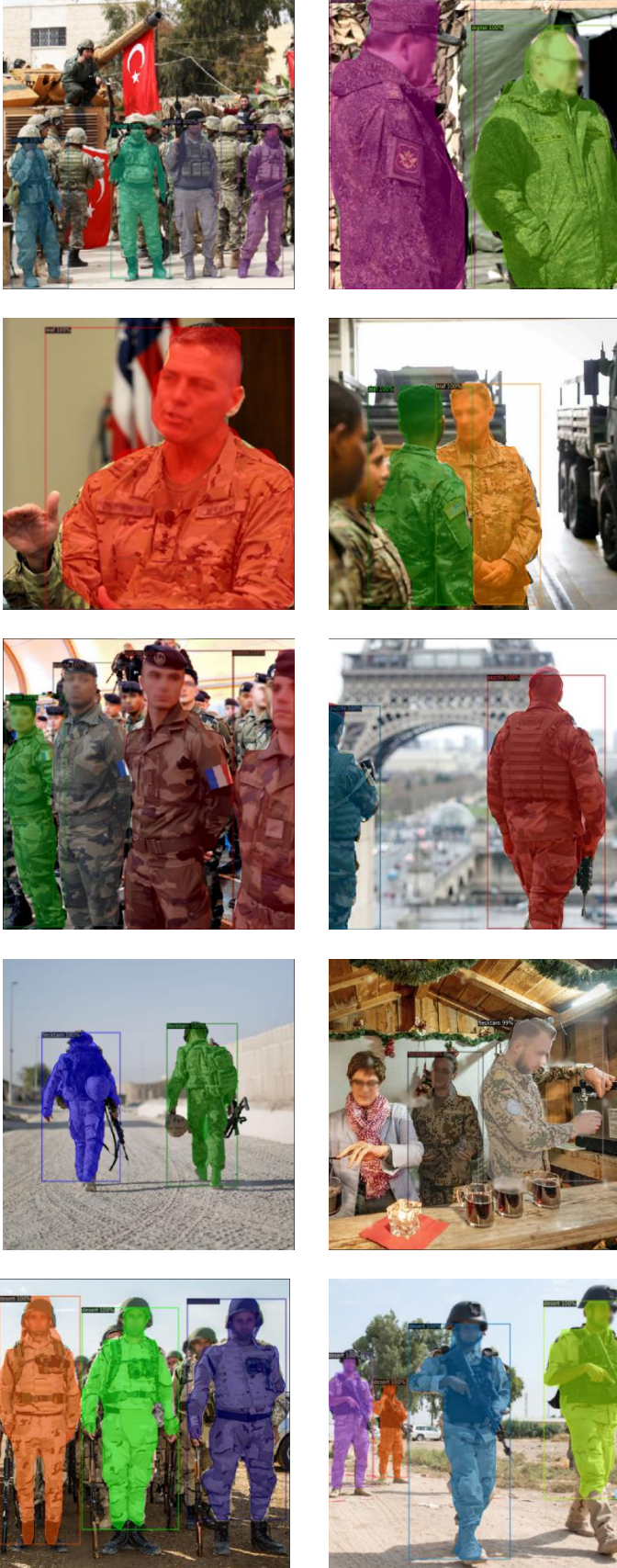
(b)

Şekil 5. Eğitim metrikleri: (a) Sınıflandırma Kaybı, (b) Maskeleme Kaybı

V. TARTIŞMA

Literatürdeki çalışmalara göre, kamuflajlı nesnelerin bölütlenmesi için görüntü işleme yöntemleri kullanılmakta ve genellikle farklı eşik değerleri seçimi için çeşitli yöntemler tercih edilmektedir. Kamuflajlı nesne sınıflandırmalarında ise sınıflandırılan nesnelerin desenlerden ziyade genel hatlarına odaklanmak daha etkilidir, çünkü bu nesneler birbirinden oldukça farklıdır. Askeri kamuflaj desenlerinin matematiksel karmaşıklığı düşünüldüğünde, görüntü işleme yöntemleriyle sınıflandırılması zor bir problemidir. Çalışmamızda, sınıflandırılacak nesnelerin sadece askerler olması, kullanılan algoritmanın desenleri baz alarak sınıflandırma yapabildiğini sağlamıştır. Tablo 2’de verilen eğitim metriklerine bakıldığında, Mask R-CNN'nin bu karmaşık sınıflandırma problemi için başarılı bir yöntem olduğu görülmektedir. Ayrıca Şekil 3’e göre sadece kamuflajlı insanların tespiti yapılmaktadır. Bu da gösterir ki segmentasyon ve sınıflandırmada öğrenilen kriter kamuflaj deseninin kendisidir. Veri setinde normal insanların

etiketlenmemesi yani eğitime dahil edilmemesi de bunu kanıtlar.



Şekil 6. Mask R-CNN ile Örnek Segmentasyon Sonuçları
(Her sıra bir desen ailesi olacak şekilde: dijital, yaprak, puzzle, benekli, çöl desenleri)

VI. SONUÇLAR

Sonuçlar, CNN algoritmalarının sadece farklı ana hatlara sahip nesnelerin sınıflandırılması için değil, matematiksel karmaşıklığı yüksek sınıflandırma problemlerinde de etkili olduğunu göstermektedir. İleri çalışmalarda sınıf sayısı artırılarak farklı problemler üzerinde de çalışmalar yapılabilir.

KAYNAKLAR

- [1] (2023) How Military Camouflage Works. [online]. Available: <https://science.howstuffworks.com/military-camouflage1.htm>
- [2] Talas, Laszlo & Baddeley, Roland & Cuthill, Innes. (2017). Cultural evolution of military camouflage. *Philosophical Transactions of The Royal Society B Biological Sciences*. 372. 20160351. 10.1098/rstb.2016.0351.
- [3] Hodson-Pressinger S. 2004 Khaki uniform, 1848-49: first introduction by Lumsden and Hodson. *J. Soc. Army Hist. Res.* 82, 341-347.
- [4] Singh, Rohit. (2014). Study of Graphic Camouflage Patterns on Battle Uniform and improving the pattern used by Indian Army.
- [5] (2023) The Camouflage encyclopedia. [online]. Available: https://www.camopedia.org/index.php/Main_Page
- [6] Doğan, F., & Türkoğlu, I. (2018). Derin Öğrenme Algoritmalarının Yaprak Sınıflandırma Başarımlarının Karşılaştırılması.
- [7] A. Gupta and U. Gupta, "Military Surveillance with Deep Convolutional Neural Network," 2018 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT), 2018, pp. 1147-1152, doi: 10.1109/ICEECCOT43722.2018.9001381.
- [8] E. Bayram and V. Nabyev, "Classification of Camouflage Images Using Local Binary Patterns (LBP)," 2021 29th Signal Processing and Communications Applications Conference (SIU), 2021, pp. 1-4, doi: 10.1109/SIU53274.2021.9478040.
- [9] Yan, Jinnan. "Camouflaged Object Segmentation in Images." Master's thesis, University of Dayton, 2019.
- [10] Fan, D.P., Ji, G.P., Sun, G., Cheng, M.M., Shen, J., & Shao, L. (2020). Camouflaged Object Detection. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [11] K. He, G. Gkioxari, P. Dollár and R. Girshick, "Mask R-CNN," 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2980-2988, doi: 10.1109/ICCV.2017.322.
- [12] A. Nur Ömeroğlu, N. Kumbasar, E. Argun Oral and I. Y. Ozbek, "Mask R-CNN Algoritması ile Hangar Tespiti Hangar Detection with Mask R-CNN Algorithm," 2019 27th Signal Processing and Communications Applications Conference (SIU), 2019, pp. 1-4, doi: 10.1109/SIU.2019.8806552.
- [13] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection,"

- 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.
- [14] M. B. Amri, D. Yedjour, M. El Amin Larabi and K. Bakhti, "Stadium Detection From Alsat-2 and Google-Earth Multispectral Images using YOLOv5 and Mask R-CNN," 2022 4th International Conference on Pattern Analysis and Intelligent Systems (PAIS), 2022, pp. 1-4, doi: 10.1109/PAIS56586.2022.9946887.
- [15] A. Torralba, B. C. Russell and J. Yuen, "LabelMe: Online Image Annotation and Applications," in Proceedings of the IEEE, vol. 98, no. 8, pp. 1467-1484, Aug. 2010, doi: 10.1109/JPROC.2010.2050290.
- [16] (2023) Deep Learning Wizard. [online]. Available: https://www.deeplearningwizard.com/deep_learning/boosting_models_pytorch/lr_scheduling/
- [17] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, 1 June 2017, doi: 10.1109/TPAMI.2016.2577031.
- [18] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.
- [19] R. Yayla and B. Şen, "Region-based Segmentation of Terrain Fields in SAR Images," 2020 28th Signal Processing and Communications Applications Conference (SIU), 2020, pp. 1-4, doi: 10.1109/SIU49456.2020.9302432.
- [20] A. K Yuxin Wu, Francisco Massa, Wan-Yen Lo and Ross Girshick, Detectron2, July 2022.