



Meyve Bahçesinde Panorama Görüntülerden Derin Öğrenme Tabanlı Elma Tespiti

Deep Learning Based Apple Detection from Panorama Images in the Orchard

Emre Pazarlı^{1,2}, Ertuğrul Bayraktar¹

¹Mekatronik Mühendisliği Bölümü,
Yıldız Teknik Üniversitesi
emrepazarlieng@gmail.com, eb@yildiz.edu.tr

²Mekatronik Mühendisliği Bölümü,
Düzce Üniversitesi

Özet

Hassas tarımda, hasat öncesi meyve bahçelerinde meyve sayımı verim tahmini ile mahsul yükü yönetimi için önemlidir. Son yıllarda bilgisayarla görmedeki gelişmeler ve geleneksel yöntemlerin yerini derin öğrenme temelli nesne tespit algoritmalarının almasıyla hızlı ve güvenilir meyve sayımı yapılabilmektedir. Bu çalışmada sıralı elma ağaçları panorama görüntülerinde Yolo nesne tespit modeli ailesinin en yeni ve başarılı modellerinden YoloV5s ve YoloV7 algoritmaları ile meyve sayımı gerçekleştirilmiş ve model performans karşılaştırmaları sunulmuştur. Nesne tespit modellerinin eğitiminde; farklı türde elmalar içeren, geniş bir veri kümesi kullanılmıştır. Yine aynı veri kümesindeki test görüntülerinden görüntü birleştirme işlemi ile panorama görüntüler elde edilmiştir. Veri artırımı kullanılarak YoloV7 modelinde panorama görüntülerde %92,6 F1 skoru ile en yüksek başarıma ulaşılmıştır.

Abstract

Fruit counting is one of the most crucial process regarding the yield prediction and crop load management for pre-harvest orchards in precision agriculture. In recent years, deep-learning based computer vision algorithms perform the fruit counting thanks to rapid advancements resulting in the replacement of traditional approaches. Hence, faster and more reliable fruit counting can be done. This study proposes a fruit counting scheme depending on the panorama images of sequential apple trees, where we run our object detection method. We further provide the model performance comparisons as a result of attentively conducted fine-tuning of various object detection models. To fine-tune the object detection models, we exploited a large dataset containing different types of apples. Moreover, we obtained panorama images through image stitching from the test images in the same data set. We achieved an F1 score of 92.6% owing to data augmentation.

1. Giriş

Gelişen bilgisayar, görüntü işleme ve kamera teknolojileri sayesinde bilgisayarla göri, pek çok alanda olduğu gibi tarımda da yaygın biçimde kullanılmaktadır. Tarımda bilgisayarla göri

uygulamalarından bir tanesi meyve bahçelerinde meyve sayımıdır. Hızlı ve yüksek doğrulukta meyve sayımı, çiftçilerin doğru verim tahmini ile meyve bahçelerini yönetmesi için önemlidir. Hasat süreçlerini optimize etmede, arazideki olası verim değişikliğini anlayarak işgücü ve makine kullanımında, ürünün depolama ve nakliye gibi karar süreçlerinde çiftçiye avantaj sağlamaktadır.

Otomatik meyve sayımı işlemi birtakım problemi beraberinde getirmektedir. Meyvelerin dallar ve yapraklar tarafından gizlenmesi, kümelenmiş meyvelerin üst üste gelerek birbirlerini kapatması ve farklı ortam koşulları güçlük oluşturmaktadır. Geleneksel bilgisayarlı göri yöntemleri ile meyve sayımı, meyvelerin elle çıkarılan renk, şekil ve doku özelliklerine dayanmaktadır ve bu yöntemler belirli koşullar altında iyi çalışsalar da farklı ortamlarda performansları düşüktür.

Derin öğrenme; sinir ağlarını, hiyerarşik olasılık modellerini ve çeşitli denetimli ve denetimsiz öğrenme algoritmalarını kapsayan zengin bir yöntemler ailesidir. Bilgisayarlı göri alanlarında derin öğrenme yöntemlerinin kullanımı, geleneksel bilgisayarlı göri yöntemlerden daha iyi performans göstermektedir. Bu da son yıllarda araştırmacıların bilgisayarla göri alanlarında derin öğrenme yöntemlerine olan ilgilerini artırmıştır.

Sıralı ve bitişik meyve ağaçlarından oluşan bahçe ortamında meyve sayımı işlemi çoğunlukla hareketli görüntülerde derin öğrenme temelli nesne tespit algoritmalarının nesne takip algoritmaları ile birlikte kullanılmasıyla gerçekleştirilmektedir. Bu yöntemlerin uygulamalarında, meyvelerin çift sayımı ve arka plandaki meyvelerin yol açtığı aykırı değerler sıkça görülmektedir. Bu hatalı sayımların giderilmesi kompleks çözümler gerektirmektedir. Bu bildiride, panorama görüntüleme ile meyve sayımı önerilmiştir. Çalışmada, farklı türde elmalar içeren MinneApple [1] veri kümesi, YoloV5s ve YoloV7 modellerinin eğitiminde kullanılmıştır. Görüntü birleştirme ile panorama görüntüler elde edilip model tespiti gerçekleştirilmiş ve sonuçlar karşılaştırılmıştır.

Bu bildirinin 2. bölümünde motivasyon ve ilgili çalışmalar yer almaktadır. 3. bölümde meyve bahçesinde panoramik

görüntülerden elma tespitine yönelik tasarıma, 4. Bölümde deney çıktılarına yer verilmiş olup son olarak 5. Bölümde sonuçlar bulunmaktadır.

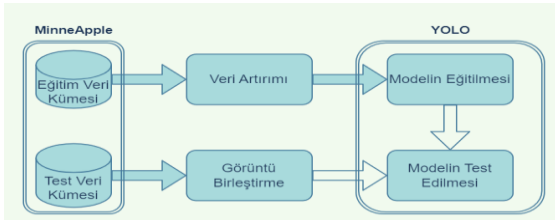
2. Motivasyon ve İlgili Çalışmalar

Meyve üretimi ülkemiz ekonomisi için önem arz etmektedir ve bu alanda yapılacak akademik ve teknolojik çalışmalar ülkemize fayda sağlayacaktır. Buradan hareketle, meyve bahçelerinde meyve sayımındaki başarı oranının yükseltilmesi ve gerçek ortamda uygulanabilirliğinin artırılması amaçlanmıştır.

Literatürde, meyve ağacı görüntülerinde meyve tespiti ve sayımı ile ilgili çok sayıda çalışma bulunmaktadır. Geleneksel yöntemler; renk tabanlı [2], [3], şekil tabanlı [4], doku tabanlı [5], [6] ve renk, şekil ile doku özelliklerinin kombine olarak kullanılması [7], [8] şeklindedir. Derin öğrenmenin kullanıldığı modern yaklaşımlar; sabit görüntülerde [9], [10], [11] ve hareketli görüntülerde [12], [13], [14], [15] meyve tespiti ile sayımı şeklinde ikiye ayrılabilir. Bunlara ek olarak, fotoğrafların birleştirilmesiyle [16], [17] veya panoramik çekilmiş görüntülerle [18] yapılan çalışmalar bulunmaktadır. Bu çalışma kapsamında çekilen tekil görüntülerden panorama görüntüler oluşturulup meyve tespit ve sayımı gerçekleştirilmiştir. Diğer bir yandan, literatürdeki güçlü modellerin ilgili probleme özgü daraltılmış veriler üzerinde ince-ayar ile eğitilerek başarımları ve güvenilirliği daha yüksek olan [19] ve [20]'deki öncül çalışmalardan da faydalanılarak, çeşitli modeller arasında sistematik karşılaştırma yapılmıştır.

3. Panoramik Görüntülerden Elma Tespiti

Bu bölümde, panoramik görüntülerden elma tespiti için önerilen yöntem açıklanmıştır. Şekil 1'de genel hatları verilen çalışma; eğitim veri kümesinde veri artırımı, test veri kümesinden görüntü birleştirme ile panorama görüntü elde edilmesi, Yolo modelinin eğitilmesi ve son olarak eğitilmiş modelin panorama görüntülerde test edilmesi adımlarından oluşmaktadır.

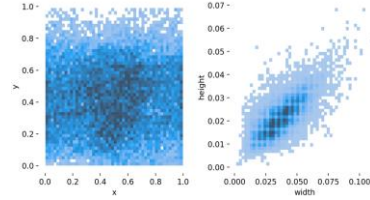


Şekil 1: Bu çalışmada yer alan adımların akış şeması.

3.1. Elma Tespiti için Kullanılan Veri Kümesi

Bu çalışmada, Minnesota Üniversitesi Robotik Sensör Ağları Laboratuvarı ekibi tarafından 2015 Haziran ve 2016 Eylül tarihleri arasında çekilen video dizilerinden çıkarılan MinneApple veri kümesi kullanılmıştır. 1280x720 boyutundaki veri kümesi, birbirinden farklı eğitim ve test veri kümelerinden oluşmaktadır. Eğitim veri kümesi yazarlar tarafından etiketlenmiş 670 görüntü, test kümesi etiketlenmesi yapılmamış 331 görüntü içermektedir. Eğitim veri kümesi görüntü başına ortalama 42 elma içermektedir ve yaklaşık %20'lik kısmı olan 135 görüntü model doğrulaması için ayrılmıştır. Yazarlar, etiketleme işlemine yalnızca ön plandaki ağaçlarda yer alan elmaları etiketlenmiş, arka planda ve yerdeki elmaların etiketlenmesinin yapılmadığını belirtmişlerdir [1]. Çalışmada,

eğitim kümesi için yeniden bir etiketleme yapılmamıştır. Açık kaynak paylaşılan etiketler sınırlayıcı kutulu yolo formatına dönüştürülmüştür. Model performansının iyileştirilmesi açısından elma görüntüsü içeren, yerde ve arka plandaki kısımlar Adobe Photoshop yardımıyla maskelenerek görüntüden çıkarılmıştır. Şekil 2'de eğitim veri kümesi görüntülerinde elma nesnelere boyut ve konumlarının dağılımı gösterilmiştir.



Şekil 2: MinneApple eğitim veri kümesindeki elmaların boyut ve konumlarının dağılımı.

3.2. Veri Artırımı

Derin öğrenme temelli nesne tespiti çok sayıda görüntüye ihtiyaç duymaktadır. İstenilen sayıda görüntünün elde edilmesinin güç olması ya da çok sayıda etiketleme işleminin maliyetli olması sebebiyle mevcut verilerden veri artırımı teknikleri kullanılarak veri sayısının suni olarak artırılması yöntemine gidilmektedir. Bu çalışmada, klasik veri artırımı tekniklerinden çevirme, döndürme, kırma ve gürültü eklemeye ilaveten parlaklık, zıtlık, doygunluk ve renk tonu değişimi kullanılarak orijinal veri kümesi boyutuna ek 3 katı kadar görüntü elde edilmiştir.



(a) Birleştirilme işlemi uygulanan tekil görüntüler.



(b) Birleştirme işlemi sonucu çıkan panoramik görüntü.

Şekil 3: Tekil görüntülerden panoramik görüntü oluşturulması.

3.3. Görüntü Birleştirme

Bu çalışmada, panoramik elma ağaçları görüntülerinde meyve sayımı için ihtiyaç duyulan görüntüler yukarıda da değinildiği üzere, MinneApple veri kümesindeki test görüntülerinden görüntü birleştirme ile elde edilmiştir. Bu görüntüler, ağaç sırası boyunca yavaş ve sabit yürüme hızıyla çekilen videoların her 30. fotoğraf karesinden oluşmaktadır ve bu fotoğraf kareleri görüntü birleştirmek için uygundur. Bu işlem için açık kaynaklı Hugin panorama yazılımının 2021.0 versiyonu kullanılmıştır. Yapılan denemelerde, birleştirilecek iki görüntü arasındaki benzerliklerin saptanabilmesi için önemli nokta (keypoint) tespitinin elle yapılması daha iyi sonuçlar vermiştir. Şekil 3'te bir örneğinin görülebileceği panorama oluşturma işleminde, bozulma en az silindirik projeksiyon kullanımıyla elde edilmiştir. Örtüşme oranı %20 ile %40 arasında 2 ile 5 arasında değişen sayıda görüntülerin birleştirilmesiyle test için 14 panoramik görüntü oluşturulmuştur.

3.4. Nesne Tespit Modeli

En önde gelen derin öğrenme tabanlı nesne tespit modellerinden biri olan Yolo, sınırlayıcı kutularla sınıflandırma ve yerleştirme için tek bir evrişimli sinir ağı (CNN) kullanır. Yolo algoritmasının temel mantığı şudur: her girdi görüntü S x S boyutlu ızgaraya böler. Gerçek çerçevedeki bir nesnenin merkez koordinatları ızgaranın içindeyse, ızgara nesnenin tespitinde sınırlayıcı kutuyu ve güven skorunu tahmin eder.

İlk Yolo versiyonlarında; karmaşık, yoğun ve üst üste çakışan küçük nesnelerin tespitindeki düşük tespit doğruluğu model mimarisindeki değişiklikler ve ilave özelliklerin eklenmesiyle büyük ölçüde giderilmiştir. Bu çalışmada da en yeni Yolo versiyonlarından YoloV5s ve YoloV7 modelleri kullanılmıştır. Eğitim öncesi uygun üst parametrelerin tanımlanması gerekmektedir. Modellerin meyve sayımı performanslarının karşılaştırılabilmesi için özdeş üst parametre değerleri seçilmiştir. Başlıca üst parametreler Çizelge 1'de verilmiştir.

Çizelge 1: Üst parametreler

İmge Boyutu	1024
Devir Sayısı	150
Yığın Adeti	8
Optimizasyon Yöntemi	SGD
İlk Öğrenme Oranı	0.01

4. Deneyler ve Çıktıları

Bu bölümde, YOLO modellerinin gerçekleştirilen çalışmadaki performans sonuçları gösterilerek karşılaştırılmıştır. Bir nesne tespit modelinin değerlendirilmesinde birçok performans kriteri vardır. Deneysel çalışmanın model eğitim performanslarının değerlendirilmesinde ortalama hassasiyet (AP) metriği, panorama test görüntüleri model başarı değerlendirmesinde ise karmaşıklık matrisi değerleri ve bu değerlerle hesaplanan hassasiyet (precision), duyarlılık (recall) ve F1 skoru metrikleri kullanılmıştır. Bu çalışmada, tüm deneyler Google Colab ortamında NVIDIA Tesla P100 16GB GPU kullanılarak gerçekleştirilmiştir.

Çizelge 2'de orijinal ve artırılmış verilerle eğitilmiş YoloV5s ve YoloV7 model ağırlıklarının ortalama hassasiyet (AP) değerleri

yer almaktadır. Veri artırımı yöntemi her iki modelde de performans iyileştirmesi sağlamış olup YoloV7 diğer modele göre meyve tespitinde daha iyi başarı sağladığı görülmektedir. Elde edilen modeller bir sonraki aşama olan test görüntülerinde meyve sayımı için kullanılmıştır.

Çizelge 2: Yolo modellerinin eğitim AP sonuçları.

Nesne Tespit Modeli	Veri Kümesi	AP(50)	AP(50:95)
Yolo-V5s	Orijinal	88	43,9
	Artırılmış	90	46,3
Yolo-V7	Orijinal	89,4	44,4
	Artırılmış	91,4	46,3

Şekil 4'deki 1 numaralı panorama görüntü örneğindeki gibi, test verileri yerde ve arka planda çok sayıda elma içermektedir. Bunların bir kısmı çürük, bir kısmı ise çok küçük piksel değerlerindedir. Bu kısımların dahil edilmesiyle doğru bir şekilde gerçek referans değer (Ground Truth) oluşturulamayacağından, çokgen biçimli bölge seçimi (ROI) ile elma sayımının yapılacağı bölgeler seçilmiştir. Şekil 5'te ROI seçimi ve seçilmiş ROI, Şekil 6 ise ROI bölgelerinde Yolo modelleri ile elma tespit sonuçları verilmiştir. Denemelerde, en iyi tahminler, 0.60 güven skoru ve 0.60 örtüşme oranı (IoU) eşik değeri seçilerek elde edilmiştir.



Şekil 4: 1 No'lu Panorama görüntüde YoloV7 ile elma tespiti.



Şekil 5: ROI seçimi (solda), seçilen maskelenmiş ROI görüntüsü. (sağda)

Çizelge 3: Birleştirme sonucu elde edilen farklı panoramik görüntülerde ince-ayar yapılan modellerin veri kümesi durumuna göre sınıflandırma ölçütleri üzerinden karşılaştırılması. (DP: Doğru Pozitif, YP:, Yanlış Pozitif, YN: Yanlış Negatif)

Panorama Görüntü No	Gerçek Değer	YOLO-V5s Orijinal Veri Kümesi			YOLO-V5s Artırılmış Veri Kümesi			YOLO-V7 Orijinal Veri Kümesi			YOLO-V7 Artırılmış Veri Kümesi			
		DP	YP	YN	DP	YP	YN	DP	YP	YN	DP	YP	YN	
1	50	43	5	7	50	11	0	42	3	8	49	9	1	
2	49	42	3	7	46	5	3	43	4	6	46	8	3	
3	63	56	1	7	61	1	2	57	2	6	57	3	6	
4	43	40	0	3	43	3	0	41	0	2	42	3	1	
5	43	34	6	9	40	9	3	37	7	6	42	7	1	
6	56	45	2	11	49	5	7	51	2	5	54	4	2	
7	73	55	1	18	60	4	13	65	1	8	70	3	3	
8	33	26	1	7	30	5	3	27	1	6	31	5	2	
9	66	60	6	6	63	9	3	59	2	7	63	10	3	
10	104	93	5	11	97	6	7	95	6	9	100	9	4	
11	77	65	5	12	69	9	8	68	6	9	71	8	6	
12	89	76	5	13	83	7	6	78	5	11	85	9	4	
13	84	75	2	9	78	8	6	74	1	10	80	9	4	
14	141	125	4	16	130	8	11	130	3	11	131	11	10	
Toplam	14	971	835	46	136	899	90	72	867	43	104	921	98	50



Şekil 6: Elma ağaçları üzerinde Yolo modellerinin nesne tespit sonuçları.

Çizelge 4: Panorama görüntülerde Yolo-tabanlı modellerin performansları.

Model	Hassasiyet	Duyarlılık	F1 Skoru
YOLO-V5s Orijinal Veri Kümesi	94,8	86	90,2
YOLO-V5s Artırılmış Veri Kümesi	90,9	92,6	91,7
YOLO-V7 Orijinal Veri Kümesi	95,3	89,3	92,2
YOLO-V7 Artırılmış Veri Kümesi	90,4	94,9	92,6

Panorama test verilerinde model performans değerlendirmesinde Çizelge 3'te sınıflandırma ölçütleri üzerinden karmaşıklık matrisi değerleri ve Çizelge 4'te hassasiyet, duyarlılık ve F1 skoru metrikleri verilmiştir. Görüldüğü üzere, artırılmış verilerle eğitilen modellerde duyarlılık yüksek, hassasiyet düşüken orijinal verilerle eğitilen modellerde ise hassasiyet yüksek, duyarlılık daha düşüktür. Yani, elma tespitinde daha çok veriyile eğitilen modeller daha çok hatalı tahminde bulunmasına karşın daha az sayıda elmayı görüntüde yakalayamamıştır. Hassasiyet ve duyarlılık değerlerinin harmonik ortalaması şeklinde ifade edilen F1 skoru metriği açısından en iyi performans, %92,6 ile çoğaltılmış veriler ile ince-ayarla eğitilen YoloV7 modeliyle elde edilmiştir.

5. Tartışma ve Sonuçlar

Sıralı meyve ağaçlarında meyve sayımına yönelik bu çalışma kapsamında önerilen panorama görüntüleme ile meyve sayımı yapılmıştır. Farklı türde elmalar içeren MinneApple veri kümesinin kullanıldığı bu çalışmada, YoloV5s ve YoloV7 derin öğrenme modelleri ince-ayarla eğitilmiştir. Görüntü birleştirme ile elde edilen panorama görüntülerde, ROI ile meyve sayımının yapılacağı bölgeler seçilip arka planda ve yerdeki elmaların sayıma dahil olması önlenmiştir. Veri artırımı yönteminin model performansına etkisi incelenmiş ve her iki modelde de veri artırımı performans artışı sağlamıştır. Panorama görüntülerde meyve sayımı işleminde %92,6 F1 skoru ile yüksek başarımla elde edilmiştir.

Gelecekteki çalışmalarda, veri kümeleri çeşitlendirilip bahçe ortamında çekilecek panorama görüntülerden modelin testi gerçekleştirilecektir. Buna ek olarak, meyve tespitine yönelik YoloV7 tabanlı model geliştirilecektir.

6. Kaynaklar

- [1] N. Hani, P. Roy ve V. Isler, "MinneApple: A Benchmark Dataset for Apple Detection and Segmentation," in *IEEE Robotics and Automation Letters*, Cilt No: 5, Sayı: 2, Sayfa No: 852-858, 2020.

- [2] Zhou, Rong, vd. "Using colour features of cv.'Gala'apple fruits in an orchard in image processing to predict yield." *Precision Agriculture*, Cilt No: 13, Sayı: 5, Sayfa No: 568-580, 2012.
- [3] Dorj, Ulzii-Orshikh, Malrey Lee, ve Sang-seok Yun. "An yield estimation in citrus orchards via fruit detection and counting using image processing." *Computers and Electronics in Agriculture*, Cilt No: 140, Sayfa No: 103-112, 2017.
- [4] Lin, Guichao, vd. "Fruit detection in natural environment using partial shape matching and probabilistic Hough transform." *Precision Agriculture*, Cilt No: 21, Sayfa No: 160-177, 2020.
- [5] Z. S. Pothen ve S. Nuske, "Texture-based fruit detection via images using the smooth patterns on the fruit," *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016, Sayfa No: 5171-5176.
- [6] Chaivivatrakul, Supawadee, ve Matthew N. Dailey. "Texture-based fruit detection." *Precision Agriculture*, Cilt No: 15, Sayı: 6, Sayfa No: 662-683, 2014.
- [7] Kurtulmus, Ferhat, Won Suk Lee, ve Ali Vardar. "Green citrus detection using 'eigenfruit', color and circular Gabor texture features under natural outdoor conditions." *Computers and Electronics in Agriculture*, Cilt No: 78, Sayı: 2, Sayfa No: 140-149, 2011.
- [8] Zhang, Chunlong, Kunlin Zou, ve Yue Pan. "A method of apple image segmentation based on color-texture fusion feature and machine learning." *Agronomy*, Cilt No: 10, Sayfa No: 972, 2020.
- [9] Bargetti, Suchet, ve James Underwood. "Deep fruit detection in orchards." *2017 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2017, Sayfa No: 3626-3633.
- [10] Sa, Inkyu, vd. "Deepfruits: A fruit detection system using deep neural networks." *Sensors*, Cilt No: 16, Sayfa No: 1222, 2016.
- [11] Tian, Yunong, vd. "Apple detection during different growth stages in orchards using the improved YOLO-V3 model." *Computers and electronics in agriculture*, Cilt No: 157, Sayfa No: 417-426, 2019.
- [12] Vasconez, Juan Pablo, vd. "Comparison of convolutional neural networks in fruit detection and counting: A comprehensive evaluation." *Computers and Electronics in Agriculture*, 2020.
- [13] Liu, Xu, vd. "Robust fruit counting: Combining deep learning, tracking, and structure from motion." *2018 IEEE/RSJ international conference on intelligent robots and systems (IROS)*. IEEE, 2018, Sayfa No: 1045-152.
- [14] Osman, Youssef, Reed Dennis, ve Khalid Elgazzar. "Yield Estimation and Visualization Solution for Precision Agriculture." *Sensors*, Cilt No: 21, Sayı: 19, Sayfa No: 6657, 2021.
- [15] Zhang, Wenli, vd. "Deep-learning-based in-field citrus fruit detection and tracking." *Horticulture Research*, 2022.
- [16] Mekhalfi, Mohamed Lamine, et al. "Vision system for automatic on-tree kiwifruit counting and yield estimation." *Sensors*, Cilt No: 20, Sayı: 15, Sayfa No: 4214, 2020.
- [17] Zhang, Yanchao, vd. "Complete and accurate holly fruits counting using YOLOX object detection." *Computers and Electronics in Agriculture*, 2022.
- [18] Turečková, Alžběta, vd. "Slicing aided large scale tomato fruit detection and counting in 360-degree video data from a greenhouse." *Measurement*, 2022.
- [19] Bayraktar, E., Yigit, C.B. ve Boyraz, P., 2017, February. Tailoring the AI for Robotics: Fine-tuning Predefined Deep Convolutional Neural Network Model for a Narrower Class of Objects. In *Proceedings of the 2017 International Conference on Mechatronics Systems and Control Engineering*, 2017, Sayfa No: 10-14.
- [20] Bayraktar, E., Yigit, C.B. ve Boyraz, P., 2017, May. Robotic arm control by fine-tuned convolutional neural network model. In *2017 25th Signal Processing and Communications Applications Conference (SIU)*, 2017, Sayfa No: 1-4.