






YAPAY ZEKA TABANLI GÖRÜNTÜ İŞLEME İLE AKILLI ROBOT SÜPÜRGELEDE TEHLİKE VE ENGEL ALGILAMA

*Mustafa Melikşah ÖZMEN¹, Muzaffer EYLENCE¹, Bekir AKSOY¹

¹Isparta Uygulamalı Bilimler Üniversitesi, Teknoloji Fakültesi, Mekatronik Mühendisliği Bölümü,
Isparta

(Geliş/Received: 16.10.2024, Kabul/Accepted: 05.12.2024, Yayınlanma/Published: 31.12.2024)

ÖZ

Günümüzde teknolojinin hızla gelişmesi ile birlikte yapay zekâ teknolojileri birçok alanda sıklıkla kullanılmaya başlanmıştır. Yapay zekâ teknolojileri sağlık, eğitim, mühendislik gibi birçok disiplinler arası alanda sıklıkla kullanılmaktadır. Yapay zekâ teknolojisinin önemli kullanım alanlarından birisi de mühendislik alanıdır. Özellikle makina, elektrik-elektronik ve bilgisayar sistemlerinin bir arada kullanıldığı disiplinler arası bir mühendislik alanı olan mekatronik mühendisliğinde de yapay zekâ teknolojilerinden faydalanılmaktadır. Mekatronik mühendisliğinde özellikle robotik yapay zekâ algoritmaları sıklıkla kullanılmaktadır. Yapılan çalışmada, yapay zekâ tabanlı görüntü işleme tekniklerinin robot süpürgelerde engel ve tehlike algılama işlevine uygulanması gerçekleştirilmiştir. Geleneksel sensör tabanlı sistemlerin yüksek maliyetleri ve algılama doğruluğu sınırlamaları nedeniyle, kamera ve yapay zekâ destekli bir alternatif olarak yapay zekâ tabanlı sistemler robot süpürgelerin ev ortamındaki cansız nesnelere ve tehlikeli bölgeleri algılayarak güvenli bir şekilde temizlik yapması hedeflenmiştir. Bu amaçla, sandalye, koltuk, oyuncak ve terlik nesnelere oluşan bir veri seti oluşturularak elde edilen veri seti VGG-19, AlexNet, MobileNet V2 derin öğrenme mimarileri ile eğitilmiştir. Çalışmada kullanılan üç derin öğrenme mimarisi içerisinde MobileNet V2 modeli %97.87 doğruluk oranı ile en başarılı model olarak tespit edilmiştir. Çalışmada gerçekleştirilen derin öğrenme yaklaşımı, sensör tabanlı sistemlere kıyasla maliyet etkinliği ve çevresel farkındalık açısından daha avantajlı bir çözüm sunmaktadır.

Anahtar kelimeler: Anahtar derin öğrenme, Görüntü işleme, Robotik.

HAZARD AND OBSTACLE DETECTION IN SMART ROBOT SWEEPERS WITH ARTIFICIAL INTELLIGENCE-BASED IMAGE PROCESSING

ABSTRACT

Nowadays, with the rapid development of technology, artificial intelligence technologies have begun to be used frequently in many areas. Artificial intelligence technologies are frequently used in many interdisciplinary fields such as health, education, and engineering. One of the important areas of use of artificial intelligence technology is the field of engineering. Artificial intelligence technologies are also used in mechatronics engineering, which is an interdisciplinary engineering field where machinery, electrical-electronics, and computer systems are used together. Robotic artificial intelligence algorithms are frequently used in mechatronics engineering. In the study, the application of artificial intelligence-based image processing techniques to the obstacle and danger detection function in robot vacuum cleaners was carried out. Due to the high costs and detection accuracy limitations of traditional sensor-based systems, camera and artificial intelligence-supported alternative artificial intelligence-based systems are aimed to detect inanimate objects and dangerous areas in the home environment and clean robot vacuum cleaners safely. For this purpose, a data set consisting of chair, armchair, toy, and slipper objects was created and the obtained data set was trained with VGG-19, AlexNet, MobileNet V2 deep learning architectures.

Among the three deep learning architectures used in the study, the MobileNet V2 model was determined to be the most successful model with an accuracy rate of 97.87%. The deep learning approach implemented in the study offers a more advantageous solution in terms of cost effectiveness and environmental awareness compared to sensor-based systems.

Keywords: Deep learning, Image processing, Robotics.

1. Giriş (Introduction)

Yapay zekâ teknolojisi, günümüzde otonom sistemler üzerinde kritik bir rol oynamaktadır [1]. Özellikle ev otomasyonu alanında kullanılan akıllı robot süpürgeler, çevrelerini LIDAR ve ultrasonik sensörler gibi sistemler aracılığıyla engelleri algılayarak otonom hareket kabiliyeti kazanmaktadır. [2]. Ancak bu sensörlerin yüksek maliyetleri ve bazı durumlarda algılama doğruluklarının sınırlı olması, alternatif çözümler geliştirilmesini gerekli hale getirmektedir [3]. Özellikle merdiven gibi hassas nesnelere algılanmaması veya tehlikeli bölgeleri algılayamama gibi problemler, robot süpürgelerin mevcut teknolojileriyle tam olarak çözülememektedir [4].

Robot süpürgelerin bu problemlerine çözüm olarak, kamera tabanlı görüntü işleme teknikleri hem maliyet etkinliği hem de çevresel farkındalığın artırılması açısından önemli bir çözüm sunmaktadır. Ayrıca bu çözümlere yapay zeka tabanlı modeller kullanılarak robot süpürgelerin engellerden kaçınma ve tehlikeleri algılama gibi işlevlerini iyileştirildiği görülmektedir [5,6]. Son yıllarda, robot süpürgelerde yapay zekâ destekli görüntü işleme tekniklerinin etkinliği üzerine birçok çalışma gerçekleştirilmiştir. Konu ile ilgili akademik çalışmalar incelendiğinde, Gao ve ark. (2022), robot süpürgelerde derin öğrenme tabanlı bir sistem önererek, engel tespiti ve sınıflandırmada %98'in üzerinde doğruluk oranı elde etmişlerdir [7]. Diğer bir çalışmada ise, Zhu ve ark. (2023), robot süpürgelerin yalnızca görüntü verilerini kullanarak gerçek zamanlı olarak hareketli nesnelere algılayabilen bir sistem geliştirmiştir. Çalışma, özellikle insanlarla etkileşimli ortamlarda robotların güvenliğini artırmaya yönelik önemli katkılar sunmaktadır [8].

Çalışmada, yapay zekâ destekli görüntü işleme teknikleri kullanılarak robot süpürgelerin ev ortamındaki cansız nesnelere zarar vermeden otonom şekilde hareket ederek sensörlerin algılayamadığı merdiven gibi tehlikeleri algılayarak güvenli bir temizlik işlemi gerçekleştirmesi amaçlanmaktadır. Çalışma kapsamında, evdeki farklı nesnelere algılamak için özgün bir veri seti kullanılmıştır. Hazırlanan veri seti, robotun hem engellerden kaçınmasını hem de potansiyel tehlikelerden korunmasını sağlamak için evrimsel sinir ağı tabanlı VGG-19, AlexNet, MobileNet V2 olmak üzere üç farklı model ile eğitilerek en başarılı model çalışmada kullanılmıştır.

2. Materyal ve Metot (Material and Method)

2.1. Materyal (Material)

2.1.1. Veri seti (data set)

Çalışmada, robot süpürgelerin ev ortamındaki canlı veya cansız nesnelere zarar vermeden otonom olarak hareket etmesi ve temizlik görevlerini yerine getirmesi amacıyla özgün bir veri seti oluşturulmuştur. Veri setinde yer alan görüntüler, ev ortamında bulunan sandalye, koltuk, vazo, oyuncak ve terlik gibi nesnelere görsellerinden oluşmaktadır. Veri setindeki görüntüler 240x240 boyutlarında renkli olarak kaydedilmiştir. Toplanan görüntüler üzerlerindeki gürültü ve paraziti azaltmak amacıyla Gauss filtreleme ve normalizasyon işlemlerine tabi tutulmuştur. Toplam 1500 görüntüden oluşan veri seti, 438 adet sandalye, 321 adet koltuk, 320 adet vazo, 214 adet oyuncak ve 207 adet terlikten oluşmaktadır. Çalışmaya özel olarak hazırlanan veri setinden alınmış örnek bir görüntü Şekil 1'de gösterilmiştir.



Şekil 1: Veri setinden örnek görüntü (Sample image from dataset)

2.1.2. VGG-19 Derin öğrenme mimarisi (VGG-19 Deep learning architecture)

VGG-19, derin öğrenme ve bilgisayarlı görü alanında önemli bir yere sahip olan 19 katmanlı bir konvolüsyonel sinir ağı mimarisidir. İlk olarak 2014 yılında Oxford Üniversitesi'ndeki Görsel Geometri Grubu tarafından geliştirilen VGG-19, daha derin ağların performansı artırabileceği fikriyle VGG-16'nın bir genişletmesi olarak ortaya çıkmıştır [9]. VGG-19'un geliştirilmesinin temel nedeni, ağ derinliğinin görüntü tanıma ve sınıflandırma görevlerindeki etkisini incelemektir. Fakat daha sonrasında modele fazla katman eklenerek, modelin daha karmaşık ve soyut özellikleri öğrenmesi hedeflenmiştir [10]. Bu yaklaşım, daha derin mimarilerin performansı artırabileceğini göstermiş ve derin öğrenme alanında yeni araştırmalara öncülük etmiştir [11]. Gelişim sürecinde, VGG-19 farklı optimizasyonlar ve modifikasyonlarla daha verimli hale getirilmiştir. Özellikle hesaplama maliyetini azaltmak ve hızlandırmak için model sıkıştırma ve ağırlık paylaşımı gibi teknikler uygulanmıştır [12]. Ayrıca, VGG-19'un transfer öğrenme için uygun bir mimari olması, farklı alanlardaki uygulamalarda geniş çapta kullanılmasını sağlamıştır [13]. Son yıllarda, VGG-19 üzerine yapılan çalışmalar, modelin tıbbi görüntü analizi, nesne tespiti ve doğal dil işleme gibi farklı alanlarda başarılı bir şekilde uygulanabileceğini göstermiştir [14]. Yukarıda belirtilen nedenlerden dolayı, VGG-19 modeli, esnek ve güçlü bir derin öğrenme modeli olarak önemini koruduğunu ortaya koymaktadır [15].

2.1.3. AlexNet derin öğrenme mimarisi (AlexNet deep learning architecture)

AlexNet, derin öğrenme ve bilgisayarlı görü alanında önemli bir konvolüsyonel sinir ağı mimarisidir. 2012 yılında Krizhevsky, Sutskever ve Hinton tarafından geliştirilen AlexNet, ImageNet Büyük Ölçekli Görüntü Tanıma Yarışması'nda birinci olan bir derin öğrenme mimarisidir [16]. AlexNet'in büyük veri setleri üzerinde derin sinir ağlarının performansını artırmak ve o dönemde yaygın olan "aşırı öğrenme" ve hesaplama maliyeti sorunlarını gidermek için tasarlanmış bir modeldir [17]. AlexNet, daha derin ve geniş bir mimari kullanarak karmaşık özelliklerin öğrenilmesini sağlamıştır [18]. Model, sekiz öğrenilebilir katmandan oluşmakta ve ReLU (Rectified Linear Unit) aktivasyon fonksiyonu sayesinde eğitim süresi hızlandırılmaktadır [19]. Ayrıca, dropout ve veri artırma gibi teknikler kullanılarak modelin genelleme kabiliyeti artırılmıştır [20]. Bu mimarinin geliştirilmesi, GPU'ların paralel işlem yeteneklerinden faydalanarak büyük ölçekli ağların eğitilmesini mümkün kılmıştır [21]. AlexNet'in başarısı, derin öğrenme alanında yeni araştırmalara ve daha derin ağların geliştirilmesine öncülük etmiştir [22]. Özellikle nesne tanıma, sınıflandırma ve segmentasyon gibi görevlerde konvolüsyonel sinir ağlarının etkinliği kanıtlanmıştır [23]. Son yıllarda, AlexNet'in transfer öğrenme ve ince ayar yöntemleriyle farklı alanlara uyarlanması, modelin esnekliğini ve kullanım alanlarını genişletmiştir [24]. Bu gelişmeler, AlexNet'in derin öğrenme tarihinde önemli bir dönüm noktası olduğunu ve sonraki ağ mimarilerinin tasarımını etkilediğini göstermektedir [25].

2.1.4. MobileNet V2 derin öğrenme mimarisi (MobileNet V2 deep learning architecture)

MobileNet V2, mobil ve gömülü cihazlar gibi sınırlı hesaplama kaynaklarına sahip platformlarda derin öğrenme modellerinin etkin bir şekilde çalıştırılması için tasarlanmış bir konvolüsyonel sinir ağı mimarisidir. 2018 yılında Sandler ve arkadaşları tarafından geliştirilen MobileNet V2, "tersine çevrilmiş artıklar" ve "lineer dar boğazlar" konseptlerini tanıtarak önceki modellere göre önemli iyileştirmeler sunmuştur [26]. MobileNet V2'nin temel amacı, hesaplama maliyetini ve model boyutunu azaltırken performanstan ödün vermemesidir. Bu mimari, derinlik yönlü ayrılabilir konvolüsyonlar kullanarak parametre sayısını ve hesaplama yükünü önemli ölçüde azaltmaktadır [27]. Modelde yer alan "Tersine çevrilmiş artıklar" ise gradyanların verimli bir şekilde geri yayılmasını sağlayarak daha derin ağların eğitilmesine olanak tanımaktadır [28]. Son yıllarda yapılan araştırmalar, MobileNet v2'nin nesne tespiti, yüz tanıma ve görüntü segmentasyonu gibi farklı bilgisayarlı görü görevlerinde başarılı bir şekilde uygulandığını göstermiştir [29]. Ayrıca, transfer öğrenme teknikleriyle birleştirildiğinde, MobileNet V2 sınırlı veri setleri üzerinde bile yüksek doğruluk oranlarına ulaşmıştır [30]. Bu özellikleri sayesinde, mobil uygulamalar ve gerçek zamanlı sistemler için tercih edilen bir mimari haline gelmiştir [31]. MobileNet V2'nin geliştirilmesi, derin öğrenme modellerinin daha geniş bir cihaz yelpazesinde kullanılabilmesini sağlamış ve yapay zeka uygulamalarının yaygınlaşmasına katkıda bulunmuştur [32].

2.1.5. Performans değerlendirme metrikleri (Performance evaluation metrics)

Yapay zekâ sınıflandırma problemlerinin çözümü için oluşturulan modellerin performansını değerlendirmek için farklı metrikler kullanılmaktadır. Bu metrikler arasında doğruluk, hassasiyet, geri çağırma ve F1 skoru önemli bir yere sahiptir. Bu metrikler, özellikle sınıflandırma problemlerinde modelin başarımını nicel olarak ifade etmek için kullanılır.

Doğruluk, modelin tüm örnekler üzerinde ne kadar doğru tahmin yaptığını gösteren bir metriktir. Genel performansı ölçmek için kullanılır ve denklem 1'de verilen eşitlik kullanılarak hesaplanmaktadır [33,34].

$$\text{Doğruluk} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Denklemden TP doğru pozitif, TN doğru negatif, FP yanlış pozitif ve FN ise yanlış negatif değerleri ifade etmektedir. Doğruluk performans değerlendirme metriği, modelin genel doğruluğunu ölçmek için temel bir gösterge sağlamaktadır.

Hassasiyet, modelin pozitif olarak tahmin ettiği örneklerin ne kadarının gerçekten pozitif olduğunu göstermek için kullanılan performans değerlendirme metriğidir. Yanlış pozitiflerin önemli olduğu durumlarda kullanılır ve denklem 2'de verilen matematiksel ifade ile hesaplanmaktadır [35,36].

$$\text{Hassasiyet} = \frac{TP}{TP+FP} \quad (2)$$

Geri çağırma, gerçek pozitif örneklerin ne kadarının model tarafından doğru bir şekilde tespit edildiğini belirlemede kullanılan performans değerlendirme metriğidir. Yanlış negatiflerin önemli olduğu durumlarda tercih edilir ve denklem 3'te verilen matematiksel ifade ile hesaplanmaktadır [37,38].

$$\text{Geri Çağırma} = \frac{TP}{TP+FN} \quad (3)$$

F1 skoru, hassasiyet ve geri çağırmanın harmonik ortalamasıdır ve bu iki metriğin dengelenmesi gerektiği durumlarda kullanılan bir performans değerlendirme metriğidir. Denklem 4'te verilen matematiksel eşitlik ile hesaplanmaktadır [39,40].

$$\text{F1 Skor} = 2 \times \frac{\text{Hassasiyet} \times \text{Geri Çağırma}}{\text{Hassasiyet} + \text{Geri Çağırma}} \quad (4)$$

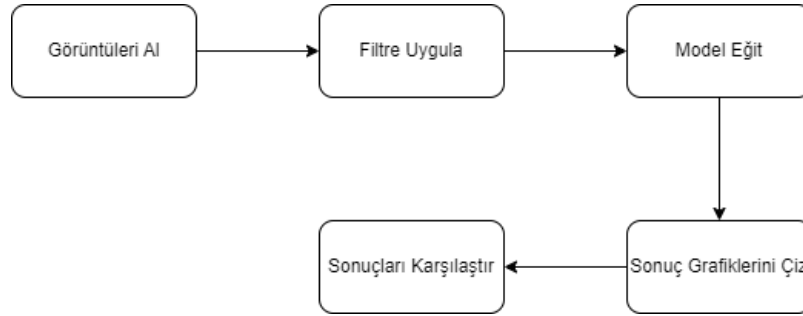
F1 skoru, dengesiz veri setlerinde modelin performansını değerlendirmesinde sıklıkla kullanılmaktadır.

Doğruluk, hassasiyet, geri çağırma ve F1 Skor performans değerlendirme metrikleri sınıflandırma modellerinin performansını daha derinlemesine anlamak ve karşılaştırmak için kullanılmaktadır.

Özellikle dengesiz veri setlerinde, doğruluk tek başına yanıltıcı olabileceğinden dolayı hassasiyet, geri çağırma ve F1 skoru daha anlamlı sonuçlar sunmaktadır [41,42].

2.2. Metot (Method)

Çalışmanın iş akışı diyagramı Şekil 2'de verilmiştir. Çalışmanın ilk aşamasında derin öğrenme modellerinin eğitimi için gerekli olan özgün görüntüler elde edilmiştir. Elde edilen görüntüler, model eğitimi öncesinde veri ön işleme adımlarına tabi tutulmuş ve eğitim süreci için uygun hale getirilmiştir. Daha sonra eğitim süreci VGG-19, AlexNet ve MobileNet V2 derin öğrenme modelleri kullanılarak gerçekleştirilmiş ve bu modellerle elde edilen sonuçlar karşılaştırmalı olarak analiz edilmiştir. Çalışmada kullanılan yöntem aşağıda daha detaylı olarak ele alınmıştır.



Şekil 2: İş Akış Diyagramı (Workflow Diagram)

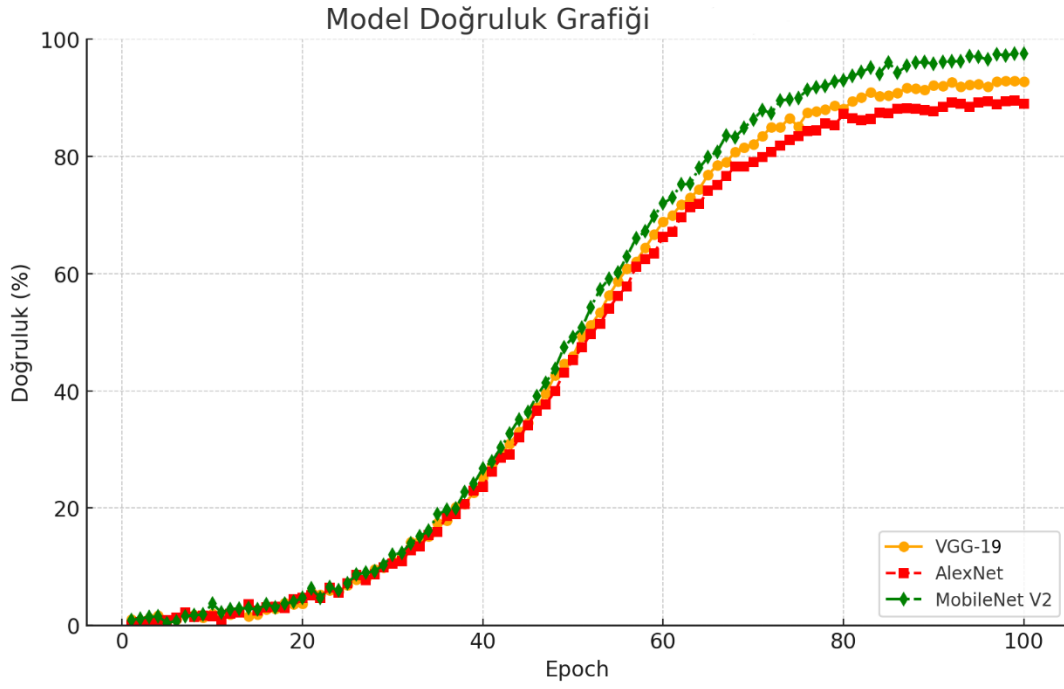
Çalışmada, robot süpürgelerin ev ortamındaki canlı veya cansız nesnelere zarar vermeden otonom olarak hareket etmesini ve temizlik görevlerini yerine getirmesini amaçlayan bir sistem geliştirilmiştir. Bu bağlamda, elde edilen özgün veri setindeki görüntüler VGG-19, AlexNet ve MobileNet V2 derin öğrenme mimarileri ile eğitilmiştir. Çalışmada kullanılan görüntüler, ev ortamında toplanan sandalye, koltuk, vazo, oyuncak ve terlik nesnelere ait görüntülerinden oluşmaktadır. Görüntüler eğitim öncesinde ilk olarak bir dizi veri ön işleme aşamasına tabi tutulmuştur.

İlk olarak, görüntülerden gürültü ve paraziti gidermek için Gauss filtresi uygulanmıştır. Bu işlem, görüntüleri daha net ve temiz hale getirerek modelin daha güvenilir sonuçlar üretmesine katkıda bulunmuştur. Gürültüyü gidermek, modelin yanlış pozitif veya yanlış negatif sonuçlar verme olasılığını azaltmak için önemli bir veri ön işleme aşamasıdır. Gauss filtresinden sonra, normalizasyon işlemi veri kümesindeki görüntülere uygulanmıştır. Normalizasyon, her görüntüdeki piksel değerlerini belirli bir aralığa ölçekleyerek modelin eğitim süresini kısaltmış ve ayrıca modelin öğrenme yeteneğini artırmıştır. Bu işlem ayrıca modelin farklı ışık ve kontrast koşulları altında daha genel performans göstermesine yardımcı olmuştur.

Görüntü ön işleme aşamasından geçirilen görüntüler VGG-19, AlexNet ve MobileNet V2 derin öğrenme modelleri ile 100 tekrar (epoch) boyunca eğitilmiş ve öğrenme oranı 0,0001 olarak belirlenmiştir. Bu süreçte, modelin aşırı öğrenme yapmasını önlemek için uygun eğitim stratejileri benimsenmiştir. Eğitim sırasında, VGG-19, AlexNet ve MobileNet V2 derin modelleri aynı veri kümesi ve parametrelerle eğitilmiştir, böylece üç modelin performansları karşılaştırılmıştır. Eğitim sürecinde, doğruluk, kayıp değerleri ve model performansını izlemek için farklı metrikler kullanılmıştır.

3. Bulgular (Results)

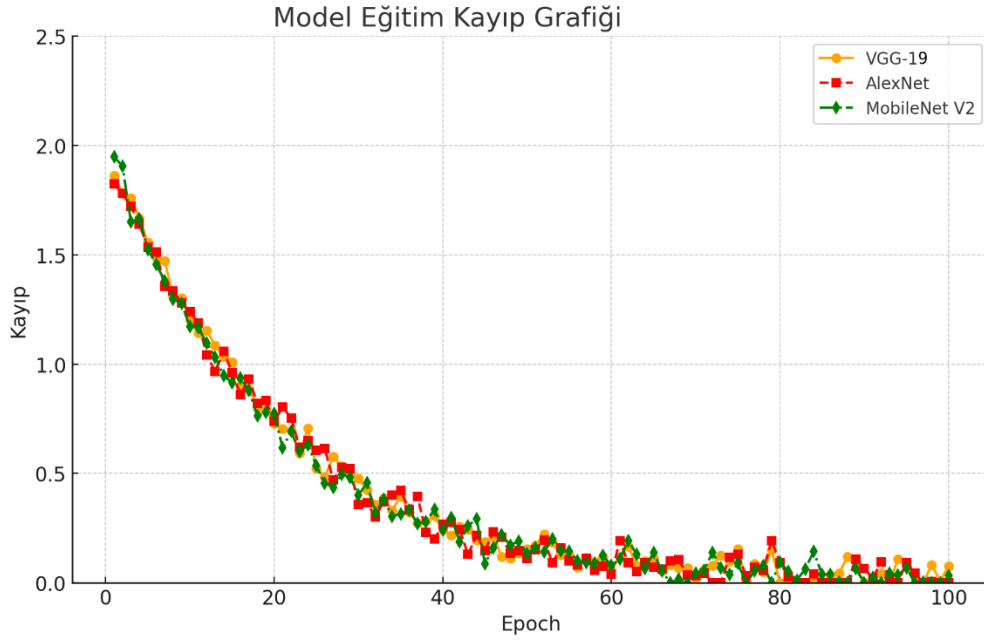
VGG-19, AlexNet ve MobileNet V2 derin öğrenme modellerinin eğitimleri sonucu elde edilen model doğruluk grafiği Şekil 3'te verilmiştir.



Şekil 3: Model Doğruluk Grafiği (Model Accuracy Graph)

Şekil 3'te, VGG-19, AlexNet ve MobileNet V2'nin üç farklı derin öğrenme modeli olan doğruluk performansı gösterilmektedir. Yatay eksen tekrar (epoch) olarak adlandırılan eğitim sürecindeki adımları ifade ederken, dikey eksen modellerin doğruluk yüzdesini göstermektedir. Şekil 3, her üç modelin de eğitim süreci boyunca doğruluklarının arttığını ve belli bir noktada doygunluğa ulaştığını göstermektedir. VGG-19, ilk dönemlerde diğer iki modelle benzer bir performans sergilerken, 20. epoch'tan itibaren daha istikrarlı bir doğruluk artışı göstermekte ve 60. epoch civarında en yüksek doğruluk oranına ulaşmaktadır. MobileNet V2, başlangıçta belirgin bir üstünlük sağlamamakla birlikte, genel olarak eğitim süresince en yüksek doğruluk değerlerine ulaşarak VGG-19 derin öğrenme mimarisi doğruluk anlamında biraz daha üstünlük sağladığı tespit edilmiştir. AlexNet derin öğrenme mimarisi ise diğer iki modelin gerisinde kalarak daha düşük doğruluk oranlarına ulaşmıştır.

Elde edilen sonuçlara göre, MobileNet V2'nin genel anlamda en yüksek performansı sergilediğini, VGG-19'nın da benzer bir başarıya ulaştığını, ancak AlexNet'in performans açısından diğer iki modelin gerisinde kaldığı tespit edilmiştir. Oluşan bu farklılıklar, modellerin mimarilerine ve hesaplama kapasitelerine bağlı olarak değişiklik gösterebilmektedir. MobileNet V2, düşük hesaplama gereksinimlerine sahip olmasına karşın yüksek performans sunarak verimlilik açısından önemli bir avantaj sağlamaktadır. Şekil 4'te çalışmada kullanılan derin öğrenme modellerinin eğitim kayıp grafikleri verilmiştir.



Şekil 4: Model Kayıp Grafiği (Model Loss Chart)

Şekil 4 incelendiğinde VGG-19, AlexNet ve MobileNet V2 modellerinin eğitim sırasında kayıp hataların zaman içinde nasıl azaldığı görülmektedir. Çalışmada kullanılan her üç model de başlangıçta yüksek bir kayıp değeriyle (yaklaşık 2.0) başlamakta ve ilk 20 tekrar (epoch) süresince modellerin kayıplarının hızlı bir şekilde azaldığı, ardından daha yavaş bir düşüşle 60. tekrar (epoch) civarında hemen hemen en düşük kayıp değerlerine ulaştıkları gözlemlenmektedir. Eğitim sürecinin sonlarına doğru ise kayıpların neredeyse sıfıra yaklaştığı gözlemlenmektedir. VGG-19, MobileNet V2 ve AlexNet arasındaki kayıp eğrileri genel olarak birbirine oldukça yakındır. Başlangıç aşamasında MobileNet V2 ve VGG-19 hafif bir üstünlük gösterirken, AlexNet özellikle 20. tekrar (epoch) sonrası benzer bir performans sergilemektedir. 60. tekrardan (epoch) sonra ise her üç modelde de kayıpların neredeyse sabitlendiği ve 0.2'nin altına düştüğü gözlemlenmektedir.

Şekil 4'teki grafik sonuçları, modellerin doğruluğunun arttığını ve hatalarının azaldığını göstermektedir. Genel anlamda, her üç modelin de eğitim sırasında hatalarını başarıyla azaltarak etkili bir öğrenme gerçekleştirdiği gözlemlenmiştir. MobileNet V2 ve VGG-19 küçük bir farkla daha düşük kayıplara sahip olsa da genel performansları arasında büyük farklar bulunmamaktadır. VGG-19, AlexNet ve MobileNet V2 derin öğrenme modellerinin eğitimi sonucu elde edilen sonuçların doğruluk, hassasiyet, geri çağırma ve F1 skor performans değerlendirme metrikleri ile değerlendirilmesi sonucu elde edilen sonuçlar Tablo 1'de gösterilmiştir.

Tablo 1: Performans Değerlendirme Sonuçları (Performance Evaluation Results)

Model	Doğruluk	Hassasiyet	Geri Çağırma	F1 Skor	Eğitim Süresi
VGG-19	%93.51	%93.59	%93.51	%93.51	12dk45sn
AlexNet	%90.11	%90.11	%91.12	%90.21	15dk25sn
Mobilenet V2	%97.87	%97.87	%97.00	%97.82	11dk21sn

Tablo 1 incelendiğinde, MobileNet V2 modeli, %97.87 doğruluk oranı ile en iyi performansı sergilemektedir. Bu modelin hassasiyet oranı %97.87, geri çağırma oranı ise %97.00'dür. Bu yüksek değerler, MobileNet V2'nin sınıflandırma işlemlerinde doğru tahminler yaptığını göstermektedir. Ayrıca, %97.82 F1 skoru, modelin genel denge performansını yansıtmaktadır. MobileNet V2, diğer modellere göre daha kısa bir eğitim süresine (11 dakika 21 saniye) sahip olup, bu da performansının yanı sıra zaman açısından da avantaj sağladığını göstermektedir. VGG-19 modeli, %93.51 doğruluk oranı ile ikinci en iyi performansı sunmaktadır. Hassasiyet değeri %93.59, geri çağırma değeri %93.51 ve F1 skoru yine %93.51 ile oldukça başarılı bir sonuç göstermiştir. Eğitim süresi ise 12 dakika 45 saniye olup, bu modelin dengeli bir performans sergilediğini, ancak MobileNet V2'ye kıyasla biraz geride kaldığını tespit edilmiştir. AlexNet modeli, %90.11 doğruluk oranı ile diğer modellere kıyasla

daha düşük bir performans sergilemektedir. Hassasiyet değeri %90.11, geri çağırma oranı %91.12 ve F1 skoru %90.21 ile diğer iki modele göre daha az başarılı bir sonuç göstermektedir. Eğitim süresi ise 15 dakika 25 saniye olup, bu modelin sınıflandırma görevlerinde hem performans hem de zaman açısından diğer iki modele başarısının daha az olduğu gözlemlenmiştir.

4. Sonuç (Conclusion)

Çalışmada, derin öğrenme tabanlı görüntü işleme ve derin öğrenme yöntemleri kullanılarak robot süpürgelerin engel ve tehlikeleri algılaması sağlanmıştır. Çalışmada robot süpürgelerin kamera verileri kullanarak çevrelerindeki cansız nesnelere ve merdiven gibi tehlikeli bölgeleri algılaması hedeflenmiştir.

Çalışmanın bulguları doğrultusunda, MobileNet V2 modeli %97.87 doğruluk oranı ile en başarılı performansı göstermiştir. Ek olarak, MobileNet V2'nin düşük hesaplama maliyetleri ve yüksek performansı, bu modelin özellikle gömülü sistemlerdeki uygulamalar için uygunluğunu göstermektedir. Eğitim süresi ve işlem maliyetlerine dair analizler, MobileNet V2'nin çalışmada kullanılan diğer derin öğrenme modellere kıyasla zaman ve kaynak verimliliği açısından daha başarılı olduğu görülmüştür. Bu durum, özellikle sınırlı donanımsal kaynaklara sahip sistemlerde bu modelin tercih edilme olasılığını artırmaktadır.

Çalışma kapsamında geliştirilen derin öğrenme tabanlı bu yaklaşım, LIDAR ve ultrasonik sensörlerin yüksek maliyetleri göz önüne alındığında, daha uygun maliyetli bir çözüm sunarken, robot süpürgeci çevresel farkındalığını ve güvenliğini artırmayı hedeflemektedir. Çalışma kapsamında geliştirilen MobileNet V2 derin öğrenme tabanlı sistemin, robotun çevresindeki nesnelere ve tehlikeli bölgeleri algılamasını sağlayarak daha güvenli bir kullanım sunmaktadır. Ayrıca gerçekleştirilen çalışma ile, evde kullanılan akıllı robotların hem cansız nesnelere zarar vermesi engellenecek hem de akıllı robot süpürgecinin kendisini tehlikelerden koruması sağlanacaktır.

İlerleyen çalışmalarda yüksek doğruluk oranlarının önemli olduğu göz önüne alındığında, bu oranı artırmaya yönelik optimizasyon çalışmaları önerilmektedir.

5. Teşekkür (Acknowledgement)

Bu çalışma 6. Uluslararası Mühendislikte Yapay Zeka ve Uygulamalı Matematik Konferansı'nda (ICAIAME 2024) özet metin olarak sunulmuştur.

6. Kaynaklar (References)

- [1] H. He, J. Gray, A. Cangelosi, Q. Meng, T.M. McGinnity, J. Mehnen, The challenges and opportunities of artificial intelligence for trustworthy robots and autonomous systems, in: Proceedings of the 2020 3rd International Conference on Intelligent Robotic and Control Engineering (IRCE), IEEE, August 2020, pp. 68–74.
- [2] A. Pandey, A. Kaushik, A.K. Jha, G. Kapse, A technological survey on autonomous home cleaning robots, Int. J. Sci. Res. Publ. 4 (4) (2014) 1–7.
- [3] N. Lopac, I. Jurdana, A. Brnelić, T. Krljan, Application of laser systems for detection and ranging in the modern road transportation and maritime sector, Sensors 22 (16) (2022) 5946.
- [4] Y.H. Lee, T.S. Leung, G. Medioni, Real-time staircase detection from a wearable stereo system, in: Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012), IEEE, November 2012, pp. 3770–3773.
- [5] G.A. Affonso, A.L. De Menezes, R.B. Nunes, D. Almonfrey, Using artificial intelligence for anomaly detection using security cameras, in: Proceedings of the 2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), IEEE, October 2021, pp. 1–5.
- [6] J. Vélez, W. McShea, H. Shamon, P.J. Castiblanco-Camacho, M.A. Tabak, C. Chalmers, et al., An evaluation of platforms for processing camera-trap data using artificial intelligence, Methods Ecol. Evol. 14 (2) (2023) 459–477.

- [7] X. Gao, J. Liu, W. Chen, Deep learning-based obstacle detection and classification for autonomous vacuum robots, *J. Artif. Intell. Res.* 75 (2022) 45–58, <https://doi.org/10.1016/j.jair.2022.06.001>.
- [8] Y. Zhu, T. Wang, X. Li, Real-time object recognition for robotic vacuum cleaners using convolutional neural networks, *Robot. Auton. Syst.* 125 (2023) 102420, <https://doi.org/10.1016/j.robot.2023.102420>.
- [9] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, in: *Proceedings of the International Conference on Learning Representations*, 2015.
- [10] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.
- [11] C. Szegedy, W. Liu, Y. Jia, et al., Going deeper with convolutions, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1–9.
- [12] S. Han, J. Pool, J. Tran, W. Dally, Learning both weights and connections for efficient neural networks, *Adv. Neural Inf. Process. Syst.* 28 (2015) 1135–1143.
- [13] J. Yosinski, J. Clune, Y. Bengio, H. Lipson, How transferable are features in deep neural networks? *Adv. Neural Inf. Process. Syst.* 27 (2014) 3320–3328.
- [14] G. Litjens, T. Kooi, B.E. Bejnordi, et al., A survey on deep learning in medical image analysis, *Med. Image Anal.* 42 (2017) 60–88.
- [15] M. Tan, Q.V. Le, EfficientNet: rethinking model scaling for convolutional neural networks, in: *Proceedings of the 36th International Conference on Machine Learning*, 2019, pp. 6105–6114.
- [16] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, *Adv. Neural Inf. Process. Syst.* 25 (2012) 1097–1105.
- [17] Y. Bengio, Learning deep architectures for AI, *Found. Trends Mach. Learn.* 2 (1) (2009) 1–127.
- [18] M.D. Zeiler, R. Fergus, Visualizing and understanding convolutional networks, in: *Proceedings of the European Conference on Computer Vision*, 2014, pp. 818–833.
- [19] V. Nair, G.E. Hinton, Rectified linear units improve restricted Boltzmann machines, in: *Proceedings of the 27th International Conference on Machine Learning*, 2010, pp. 807–814.
- [20] N. Srivastava, G. Hinton, A. Krizhevsky, et al., Dropout: a simple way to prevent neural networks from overfitting, *J. Mach. Learn. Res.* 15 (1) (2014) 1929–1958.
- [21] R. Raina, A. Madhavan, A.Y. Ng, Large-scale deep unsupervised learning using graphics processors, in: *Proceedings of the International Conference on Machine Learning*, 2009, pp. 873–881.
- [22] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, *Nature* 521 (7553) (2015) 436–444.
- [23] R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 580–587.
- [24] M. Oquab, L. Bottou, I. Laptev, J. Sivic, Learning and transferring mid-level image representations using convolutional neural networks, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1717–1724.
- [25] C. Szegedy, W. Liu, Y. Jia, et al., Going deeper with convolutions, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1–9.
- [26] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L.-C. Chen, MobileNetV2: inverted residuals and linear bottlenecks, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510–4520.
- [27] A.G. Howard, M. Zhu, B. Chen, D. Kalenichenko, et al., MobileNets: efficient convolutional neural networks for mobile vision applications, *arXiv preprint arXiv:1704.04861*, 2017.

- [28] F.N. Iandola, S. Han, M.W. Moskewicz, et al., SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size, arXiv preprint arXiv:1602.07360, 2016.
- [29] Z. Huang, X. Wang, L. Wu, Lightweight MobileNetV2 based on depthwise separable convolution for mobile device, *Int. J. Comput. Intell. Syst.* 11 (1) (2018) 1100–1108.
- [30] M. Tan, Q.V. Le, EfficientNet: rethinking model scaling for convolutional neural networks, in: *Proceedings of the 36th International Conference on Machine Learning*, 2019, pp. 6105–6114.
- [31] Y. Chen, T. Yang, X. Zhang, et al., DetNAS: backbone search for object detection, *Adv. Neural Inf. Process. Syst.* 31 (2018) 6638–6648.
- [32] Z.-Q. Zhao, P. Zheng, S.-T. Xu, X. Wu, Object detection with deep learning: a review, *IEEE Trans. Neural Netw. Learn. Syst.* 30 (11) (2019) 3212–3232.
- [33] D.M.W. Powers, Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation, *J. Mach. Learn. Technol.* 2 (1) (2011) 37–63.
- [34] R. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection, in: *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, 1995, pp. 1137–1143.
- [35] M. Sokolova, G. Lapalme, A systematic analysis of performance measures for classification tasks, *Inf. Process. Manag.* 45 (4) (2009) 427–435.
- [36] J. Davis, M. Goadrich, The relationship between precision-recall and ROC curves, in: *Proceedings of the 23rd International Conference on Machine Learning*, 2006, pp. 233–240.
- [37] C.J. Van Rijsbergen, *Information Retrieval*, 2nd ed., Butterworth-Heinemann, 1979.
- [38] C.D. Manning, P. Raghavan, H. Schütze, *Introduction to Information Retrieval*, Cambridge University Press, 2008.
- [39] N. Chinchor, MUC-4 evaluation metrics, in: *Proceedings of the 4th Message Understanding Conference*, 1992, pp. 22–29.
- [40] J. Opitz, S. Burst, Macro F1 and macro F1, arXiv preprint arXiv:1911.03347, 2019.
- [41] H. He, E.A. Garcia, Learning from imbalanced data, *IEEE Trans. Knowl. Data Eng.* 21 (9) (2009) 1263–1284.
- [42] N. Japkowicz, S. Stephen, The class imbalance problem: a systematic study, *Intell. Data Anal.* 6 (5) (2002) 429–442.