

Image Processing and Deep Learning Based Smart Door Lock System Using Face Recognition

Ali Yankı TEKOL¹, Mehmet ELMACI^{2*}, Veysel ASLANTAŞ³

¹Graduate School of Natural and Applied Sciences, Erciyes University, Kayseri, Turkey

^{2*, 3}Department of Computer Engineering, Erciyes University, Kayseri, Turkey

Abstract: The scope of this study is to design and implement a door lock system by means of face recognition using image processing and deep learning algorithms, which adds important dimensions to security systems. This system, which can operate completely contactless, is configured to allow passage only when the images obtained by the camera are verified using deep learning algorithms. The proposed system is designed to run on the B model of the Raspberry Pi 4 board, one of the widely used development boards in the market. HAAR-Cascade, HOG (Histogram of Oriented Gradients) and MTCNN (Multi-Task Cascaded Convolutional Neural Networks) algorithms, which are widely used in image processing, were used during the development of the study. As a result of the literature research, the effects of different optimization parameters on the results of ResNet50, VGG16 models Keras VGG-Face based deep learning library, which can produce stable and high performance for this problem, were examined. In addition, the performance of the pre-trained InceptionResNetV1 model in the FaceNet-PyTorch library, which is widely used for the same problem, was also analyzed. Experimental studies show that the HOG features perform best when used with the ResNet50 neural network model and trained with the Adam optimization algorithm. The combination of these three choices is the most successful among the other experiments.

Keywords: Face recognition, Deep learning, Raspberry pi, Artificial neural networks, Face detection

Görüntü İşleme ve Derin Öğrenme ile Yüz Tanıma Tabanlı Akıllı Kapı Kilit Sistemi

Özet: Bu çalışmanın amacı, güvenlik sistemlerine önemli katkılar sağlayan bir kapı kilit sistemi tasarlamak ve gerçekleştirmektir. Bu sistem, görüntü işleme ve derin öğrenme algoritmalarını kullanarak yüz tanıma işlemiyle çalışmaktadır. Tamamen temassız çalışan sistem, yalnızca kamera ile elde edilen görüntülerin derin öğrenme algoritmalarıyla doğrulanması durumunda geçişe izin verecek şekilde yapılandırılmıştır. Önerilen sistem, piyasada yaygın olarak kullanılan Raspberry Pi 4 model B geliştirme kartı üzerinde çalışacak şekilde tasarlanmıştır. Çalışmanın geliştirilmesi sürecinde görüntü işleme alanında yaygın olarak kullanılan HAAR-Cascade, HOG (Histogram of Oriented Gradients) ve MTCNN (Multi-Task Cascaded Convolutional Neural Networks) algoritmalarından yararlanılmıştır. Literatür araştırması sonucunda, bu problem için stabil ve yüksek performans gösteren ResNet50 ve VGG16 modelleri, Keras VGG-Face tabanlı derin öğrenme kütüphanesi üzerinde farklı optimizasyon parametreleriyle incelenmiştir. Ayrıca, FaceNet-PyTorch kütüphanesinde yer alan önceden eğitilmiş InceptionResNetV1 modelinin performansı da değerlendirilmiştir. Yapılan deneysel çalışmalar, HOG özelliklerinin ResNet50 sinir ağı modeli ile birlikte kullanılıp Adam optimizasyon algoritması ile eğitildiğinde en başarılı performansı sergilediğini göstermektedir. Bu üç seçimin bir araya getirildiği kombinasyon, diğer deneyler arasında en iyi sonucu vermiştir.

Anahtar Kelimeler: Yüz tanıma, Derin öğrenme, Raspberry pi, Yapay sinir ağları, Yüz algılama

Reference to this paper should be made as follows (bu makaleye aşağıdaki şekilde atıfta bulunulmalı):
Tekol, A.Y., Elmacı, M., Aslantaş, V., 'Image Processing and Deep Learning Based Smart Door Lock System Using Face Recognition', Electronic Letters on Science & Engineering, vol. 20(1), (2024), pp. 11-36.

1. GİRİŞ

Güvenlik, insan yaşamının önemli kavramlarından biri olup, güvenliğin önemi ve beraberinde alınması gereken tedbirler gün geçtikçe artmaktadır. Bunun başlıca sebebi, teknolojilerin gelişmesi

*Corresponding author; mehmetelmaci@erciyes.edu.tr

ile paralel olarak sistemsal ihlal ve açıkların, dolayısıyla hırsızlık gibi suçların artmasıdır [1]. Güvenliğin gerekli olduğu her alanda, güvenli girişler büyük önem taşımaktadır. Bu bölgelerde kullanılan geçiş sistemleri, orada bulunan önemli bilgilerin yetkisiz erişimlerden korunmasını sağlamaktadır. Genel olarak kapı kilitleri, daha teknolojik anlamda kartla geçiş, şifreli geçiş, parmak izi veya yüz tanıma ile geçiş gibi yöntemlerle giriş güvenliği sağlanabilmektedir [2]. Yüz tanıma, bir kişinin kimliğini yüz görüntüsü kullanarak tanıma ve onaylama yöntemidir. Yüz tanıma sistemleri; fotoğraflardan, videolardan veya gerçek zamanlı olarak elde edilen görüntülerde yer alan kişileri tanımlamak için kullanılmaktadır [3]. Yüz tanıma işlemi, biyometrik tanımanın bir alt dalıdır. Diğer biyometrik tanıma yöntemleri arasında ses tanıma, parmak izi tanıma ve göz retinası ya da iris tanıma yer almaktadır. Bu teknoloji çoğunlukla güvenlik alanında kullanılmakla birlikte, kurumsal olarak da farklı alanlarda giderek daha fazla uygulanmaktadır. Güvenliğin yanı sıra, biyometri ve bilgisayar sistemleri-insan etkileşimi dâhil olmak üzere birçok alanda kapsamlı bir şekilde uygulanmakta ve geliştirilmeye devam edilmektedir [4]. Özellikle görüntü işleme alanında yaygın olarak kullanılan, yapay zekânın bir alt dalı olan derin öğrenme, son yıllarda artan popülaritesiyle dikkat çekmektedir [5, 6]. Görüntü işleme alanında gerçekleştirilen çok sayıda çalışma sonucunda bu sahada kayda değer performans artışları gözlemlenmiştir. Bir derin öğrenme yöntemi olan Evrişimli Sinir Ağları (Convolutional Neural Networks, CNN) modelleri, biyometrik tanımanın yanı sıra bitki hastalıklarının sınıflandırılması, tıbbi görüntü analizi ve görüntüde sahteciliğin tespiti gibi birçok alanda kullanılmaktadır [7-9]. Biyometride kullanılan derin öğrenme tabanlı yüz tanıma görevleri, genellikle iki adımdan oluşan geleneksel işlem hattı izlenerek gerçekleştirilmektedir. Birinci aşamada görüntülerdeki yüzler tespit edilir, ikinci aşamada ise tespit edilen yüzlerin farklı ağ mimarileri kullanılarak tanınması sağlanır [10-13]. Araştırmacılar, yüz tanıma sistemlerinin performans ve doğruluğunu geliştirmek amacıyla derin öğrenme teknikleri ve CNN modellerinin kullanımına yönelik çalışmalarını sürdürmektedir. Aynı zamanda, yüz algılama ve tanıma tekniklerine ek olarak Raspberry Pi kartlarının erişim kontrolü sağlayan sistemlerde kullanımını da araştırmışlardır [14]. Raspberry Pi geliştirme kartları, çeşitli sensör ve elektronik cihazlarla kolayca entegre edilebilen, düşük maliyetli ve akıllı cep telefonu boyutlarındaki gömülü bilgisayarlar olarak tarif edilebilir. IoT (Internet of Things) sistemler, robotik ve ev otomasyonları gibi çeşitli uygulamalarda yaygın olarak kullanılmaktadır. Raspberry Pi geliştirme kartının erişim kontrolü sağlayan sistemlerde kullanımı, stabil çalışması, düşük maliyeti, esnekliği ve kolay entegrasyonu gibi çeşitli avantajlar sunmaktadır [15]. Ayrıca, Raspberry Pi geliştirme kartı, diğer cihazlar arası iletişimi sağlamak için Wi-Fi ve Bluetooth modülleri ile uzaktan yönetim için bulut tabanlı hizmetler ve çevresel gözetim için kullanılan sensörlerin entegrasyonu gibi çeşitli yaklaşımlarla kullanılabilir [16]. Ancak, Raspberry Pi

geliştirme kartının yüz tanıma teknikleriyle kullanımı halen aktif bir araştırma alanıdır. Yüz tanıma teknolojisi, güvenlik sektöründe çoğunlukla erişim kontrol sistemleri, güvenlik izleme ve kimlik doğrulama süreçlerinde yaygın olarak kullanılmaktadır. Özellikle yüksek güvenlik gerektiren alanlarda, geleneksel yöntemlere göre daha güvenilir ve kullanışlı bir çözüm olarak öne çıkmaktadır. Ancak, mevcut uygulamalar genellikle yüksek maliyetli donanımlar ve karmaşık yazılım çözümleri gerektirmektedir. Bu çalışmada önerilen sistem, mevcut çözümlerden farklı olarak, düşük maliyetli bir platform olan Raspberry Pi üzerinde çalışmakta ve derin öğrenme modelleri ile yüksek doğrulukta yüz tanıma işlemleri gerçekleştirmektedir. Bu çalışma, güvenlik alanında kullanılabilir, düşük maliyetli ve verimli bir akıllı kapı kilit sistemi geliştirmeyi amaçlamaktadır. Yüz tanıma teknolojisi ve derin öğrenme modelleriyle donatılmış bu sistem, özellikle yüksek güvenlik gerektiren yerlerde, yetkisiz girişleri önlemeyi hedeflemektedir. Ayrıca, sistemin esnekliği ve kullanım kolaylığı, yeni kişilerin hızlı ve etkili bir şekilde tanınmasını sağlamakta, bu da sistemin pratikliğini ve geniş çapta uygulanabilirliğini artırmaktadır. Bu sebeple çalışma hem akademik literatürde hem de endüstriyel uygulamalarda önemli bir katkı sağlamayı amaçlamaktadır. Çalışmada, derin öğrenme modelleri ve yüz tanıma teknolojisini kullanarak güvenli bir kapı kilit mekanizması tasarımı amaçlanmıştır. Raspberry Pi 4 B modeli kullanılarak, verimli ve etkin bir performans sağlanması hedeflenmiştir. Önerilen sistemde Keras VGG-Face kütüphanesi ile ResNet50 ve VGG16 tabanlı derin öğrenme modelleri kullanılarak model verisi eğitilmiş ve eğitilen model üzerinde yüz tespiti gerçekleştirilmiştir. Tespit edilen yüz bölgesinden HAAR-Cascade ve HOG öznitelikleri elde edilerek resimler sınıflandırılmış ve yüz tanıma işlemi gerçekleştirilmiştir. Bunun yanı sıra, yeni kişilerin sisteme eklenmesi durumunda tüm ağı yeniden eğitmek yerine, FaceNet modelini kullanan önceden eğitilmiş InceptionResNetV1 yapısı FaceNet-PyTorch kütüphanesi üzerinden uygulanmış, eklenen kişinin verisi gömmeler (embedding) şeklinde saklanmış ve tanıma işlemi MTCNN algoritması yardımı ile bu gömmeler üzerinden gerçekleştirilmiştir. Böylece sistem, yeni kişileri daha verimli ve hızlı bir şekilde tanıyabilmiştir. Çalışma, sırasıyla derin öğrenme modellerini, kullanılan materyalleri, yazılım ve donanım bileşenlerini, bulguların karşılaştırmasını, değerlendirmesini ve gelecekteki iyileştirmelere yönelik önerileri kapsamaktadır. Ayrıca makalede, sistemin tasarımı, veri kümesi yöntemi, başarı ölçütleri ve deneysel sonuçlar ayrıntılarıyla ele alınmakta ve önerilen akıllı kapı kilit sistemi tanıtılmaktadır.

1.1 Literatür Taraması

Muthu ve diğ., ResNet50 modelini kullanarak LFW ve CelebA veri setlerinde %98 doğruluk oranına ulaşarak yüz tanıma performansını artırmayı başarmışlardır [17]. Smith ve diğ., VGG16

modelini LFW ve CelebA veri setlerinde test etmiş ve %95 doğruluk oranı elde ederek modelin yüz tanıma görevlerindeki etkinliğini göstermiştir [18]. Jones ve diğ., InceptionResNet modelini LFW ve CelebA veri setlerinde uygulayarak %98.5 doğruluk oranı ile modelin güçlü performansını ortaya koymuştur [19]. Lee ve diğ., HAAR özniteliklerini WIDER FACE ve LFW veri setlerinde kullanarak %88 doğruluk oranına ulaşmış ve bu yöntemin yüz algılama performansını değerlendirmişlerdir [20]. Chen ve diğ., HOG özniteliklerini WIDER FACE ve LFW veri setlerinde test etmiş ve %91 doğruluk oranıyla bu yöntemin yüz tanıma performansını ortaya koymuştur [21]. Wu ve diğ., MTCNN algoritmasını WIDER FACE ve CelebA veri setlerinde kullanarak %94 doğruluk oranı ile çoklu görevli kaskadlı sinir ağı modelinin başarısını göstermişlerdir [22].

Tablo 1'de, çalışmada kullanılan model ve algoritmaların güncel literatürdeki karşılaştırmaları verilmiştir.

Tablo 1. Literatür karşılaştırma tablosu

Model/Algoritma	Veri Seti	Doğruluk (%)	Referans
ResNet50	LFW, CelebA	%98.00	[17] (Muthu ve diğ., 2023)
VGG16	LFW, CelebA	%95.00	[18] (Smith ve diğ., 2022)
InceptionResNet	LFW, CelebA	%98.50	[19] (Jones ve diğ., 2023)
HAAR	WIDER FACE, LFW	%88.00	[20] (Lee ve diğ., 2021)
HOG	WIDER FACE, LFW	%91.00	[21] (Chen ve diğ., 2022)
MTCNN	WIDER FACE, CelebA	%94.00	[22] (Wu ve diğ., 2023)

Tablo 1'deki verilere göre, ResNet50 ve InceptionResNet modelleri, LFW ve CelebA veri setleri üzerinde en yüksek doğruluk oranlarına (%98.00 ve %98.50) ulaşmışlardır. Bu, derin öğrenme tabanlı yöntemlerin bu alandaki etkinliğini vurgulamıştır. VGG16 modeli ise biraz daha düşük bir doğruluk oranı (%95.00) elde etmiştir, ancak yine de oldukça başarılıdır. HAAR ve HOG gibi yöntemler, daha düşük doğruluk oranları (%88.00 ve %91.00) göstermiştir, bu da bu yöntemlerin derin öğrenme tabanlı modellere kıyasla daha sınırlı performans sunduğunu gösterir. MTCNN modeli, WIDER FACE ve CelebA veri setlerinde %94.00 doğruluk oranı ile diğer geleneksel özellik çıkarım yöntemlerine kıyasla üstün bir performans sergilemiştir.

2. MATERYAL VE METOT

Bu çalışmada, farklı eğitim parametreleri ile uygulanan CNN modelleri, çeşitli öznitelik çıkarım algoritmalarıyla birlikte kullanılarak karşılaştırılmıştır. Bu kapsamda, Keras VGG-Face ve FaceNet-PyTorch kütüphaneleri ile ResNet50, VGG16 ve önceden eğitilmiş

InceptionResNetV1 derin öğrenme modelleri kullanılmıştır. Ayrıca, ResNet50 ve VGG16 modellerinin performansını artırmak amacıyla Adam, Adagrad, RMSProp ve SGD gibi optimizasyon algoritmaları yardımcı parametreler olarak uygulanmıştır. Yüz tanıma algoritmaları olarak ise HOG (Histogram of Oriented Gradients) özellik tanımlayıcı, HAAR-Cascade sınıflandırıcı ve MTCNN (Multi-task Cascaded Convolutional Networks) yüz algılama algoritmaları tercih edilmiştir. HOG, HAAR ve MTCNN algoritmalarının ürettiği benzerlik oranlarının tespitiyle birlikte, Kosinüs Benzerliği (Cosine Similarity) yöntemi kullanılarak daha sağlıklı bir güven skoru elde edilmesi hedeflenmiştir. Bu modüller aşağıda kısaca açıklanmıştır:

2.1 VGG-Face

VGG-Face, Oxford Üniversitesi'nden Visual Geometry Group (VGG) tarafından geliştirilmiş bir yüz tanıma modelidir [6]. Modelin temel amacı, farklı yüzleri doğru bir şekilde ayırt etmek ve tanımdır. VGG-Face'in mimarisi, derin konvolüsyonel sinir ağları (CNN) kullanarak yüzlerin özelliklerini çıkarır. Model, çok sayıda konvolüsyonel ve maksimum havuzlama katmanına sahip olup, her bir yüzü yüksek seviyeli özelliklerle temsil eder.

VGG-Face'in temel avantajlarından biri, derinlik ve öğrenme kapasitesinin yüksek olmasıdır. Bu, modelin daha karmaşık yüz özelliklerini öğrenmesini ve daha doğru tanıma sonuçları üretmesini sağlar. Keras kütüphanesinde uygulanan bu model, içinde ResNet50, SeNet50 ve VGG16 olmak üzere üç temel derin öğrenme sinir ağı modelini içermektedir. Çalışmada, bu kütüphaneye ait ResNet50 ve VGG16 modelleri ile kullanılan veri kümesi üzerinde eğitim ve test işlemleri gerçekleştirilmiştir.

2.2 FaceNet

FaceNet, Google tarafından geliştirilmiş bir yüz tanıma modelidir ve temel amacı, doğrudan bir kişinin yüzünü bir gömme vektörü olarak temsil etmektir [23]. Bu vektör, yüz özelliklerinin anlamlı bir özetini sunar; benzer yüzler arasındaki mesafeyi küçültürken, farklı yüzler arasındaki mesafeyi büyütür. Böylece yüz tanıma işlemi, sınıflandırma yerine mesafe ölçümü ile gerçekleştirilir.

FaceNet modeli, bir "triplet loss" fonksiyonu ile eğitilir. Bu fonksiyon, bir anchor (referans yüz), bir positive (benzer yüz) ve bir negative (farklı yüz) örnekten oluşan üçlüleri kullanır. Amaç, anchor ve positive arasındaki mesafeyi minimize ederken, anchor ve negative arasındaki mesafeyi maksimize etmektir.

Çalışmada, FaceNet, transfer öğrenme yöntemi ile kullanılarak önceden eğitilmiş model ağırlıkları üzerine küçük veri kümeleriyle ince ayar yapılmıştır. Bu yaklaşım, özellikle düşük donanım kaynaklarına sahip sistemlerde etkili bir performans sergilemiştir. TensorFlow veya PyTorch gibi kütüphanelerle birlikte uygulanabilen bu model, yüz tanıma problemlerinde sıkça tercih edilmektedir. FaceNet, genellikle InceptionResNetV1 modeli ve MTCNN (Multi-task Cascaded Convolutional Networks) gibi yüz algılama algoritması ile birlikte kullanılarak, yüzlerin doğru bir şekilde tespit edilmesini ve ardından tanınmasını sağlar. Bu kombinasyon, özellikle gerçek zamanlı uygulamalar ve düşük kaynaklı sistemlerde etkili sonuçlar elde edilmesine olanak tanır.

2.3 Kullanılan Derin Öğrenme Modelleri

2.3.1. ResNet50

He vd. tarafından geliştirilen ResNet, görüntü tanıma üzerine geniş ve yenilikçi bir evrişimli sinir ağıdır [5]. ResNet, 2015 yılında ImageNet Büyük Ölçekli Görsel Tanıma Yarışması'nı %3,57'lik bir hata oranıyla kazanmış olup, ağ bozulma sorunlarını ve küçük ya da ani değişimleri gidermekte etkilidir. Bu modelin seçilme nedeni, özellikle derin öğrenme görevlerinde sunduğu yüksek doğruluk ve düşük hata oranıdır. Avantajı, residual bağlantılar sayesinde daha derin ağların eğitilebilmesi ve daha iyi genelleme yeteneği sunmasıdır. Dezavantajı, eğitim sürecinin uzun ve hesaplama açısından yoğun olmasıdır. Uygulama sürecinde, modelin 50 katmanlı yapısı ile görüntüdeki özellikler hiyerarşik bir şekilde çıkarılır ve sınıflandırma yapılır. Karşılaşılan zorluklardan biri, modelin eğitimi sırasında aşırı öğrenme (overfitting) riskidir.

ResNet modelinin farklı varyantları bulunmaktadır: ResNet18, ResNet50 ve ResNet101. Bu çalışmada kullanılan ResNet50 versiyonu, adından da anlaşılacağı üzere 50 katmandan oluşmaktadır.

2.3.2. VGG16

2014 yılında Oxford Üniversitesi'nden Karen Simonyan vd. tarafından önerilen VGG 16, 2014 ILSVRC yarışmasında birinci ve ikinci sırayı almıştır [6]. VGG16, görüntü sınıflandırmada yüksek doğruluğa sahip, önceden eğitilmiş (pre-trained) ana modeli baz almaktadır. Görüntülerden zengin özellikler çıkararak sınıflandırmada yüksek doğruluk sağlamaktadır. VGG16, önemli donanım kaynakları ve eğitim için zaman gerektiren derin ve karmaşık bir modele sahiptir. Bu modelin seçilme nedeni, derin yapıların sınıflandırma doğruluğunu artırmasıdır. Avantajı, zengin

özellik çıkarma kapasitesi ve geniş bir veri kümesi ile eğitildiğinde yüksek performans göstermesidir. Dezavantajı, çok büyük bir model olması nedeniyle eğitim süresinin uzun ve hesaplama kaynakları açısından pahalı olmasıdır. Uygulama sürecinde, girdi görüntüleri katmanlar boyunca işlenir ve nihayetinde sınıflandırılır. Karşılaşılan zorluklar arasında, modelin eğitimi için gerekli büyük veri setlerinin elde edilmesi ve işlenmesi sayılabilir.

2.3.3. InceptionResNetV1

InceptionResNetV1, Google tarafından geliştirilen bir modeldir ve Inception ve Residual bağlantılarını birleştirir [24]. Bu model, derin öğrenme görevlerinde yüksek performans ve verimlilik sunar. InceptionResNetV1, özellikle yüz tanıma gibi görevlerde etkili sonuçlar sağlayabilir. Bu modelin seçilme nedeni, daha hızlı öğrenme ve daha derin ağlarda bile hızlı konverjans sağlamasıdır. Avantajı, hem hız hem de doğruluk açısından dengeli bir performans sunmasıdır. Dezavantajı, uygulama sürecinde diğer modellere kıyasla daha karmaşık yapıların yönetilmesi gerekliliğidir. Uygulama sürecinde, modelin residual bağlantıları ve inception modülleri, geniş ve derin özellik çıkarımı sağlar. Karşılaşılan zorluklardan biri, modelin karmaşıklığı nedeniyle yapılandırma ve ince ayarların zorlayıcı olabilmesidir. Model hem Inception yapılarının hem de residual bağlantıların avantajlarını bir arada sunarak, derin ağlarda daha etkili öğrenme ve daha hızlı konverjans sağlar.

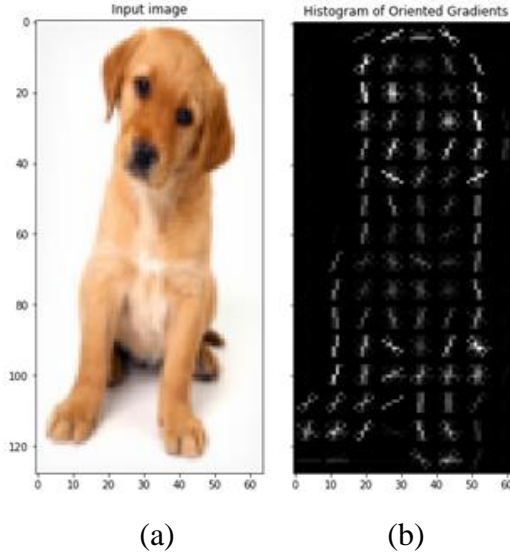
Bu modellerin her biri, derin öğrenme ve görüntü tanıma alanında farklı avantajlar ve performans özellikleri sunar. Çalışmada kullanılan bu modeller, yüz tanıma gibi görevlerde geniş bir yelpazede karşılaştırılmış ve analiz edilmiştir.

2.4 Kullanılan Öznitelikler

2.4.1. HOG

İlk olarak Dalal ve Triggs tarafından önerilen HOG (Histogram of Oriented Gradients) algoritması, özellikle insan tespiti gibi görevlerde başarılı sonuçlar üretmesiyle bilinir [25]. Bu algoritma, görüntüdeki kenarları ve gradyanları analiz ederek nesnelere tanır. Bu algoritmanın tercih edilme sebebi, özellikle basit geometrik şekillerin tanınmasında yüksek başarı göstermesidir. Avantajları arasında, kenar tespiti ve ışık değişimlerine karşı dayanıklılık bulunur. Ancak dezavantajı, karmaşık ve çoklu nesnelere bulunduğu sahnelerde performansın düşebilmesidir. Uygulama sürecinde, girdi görüntülerindeki gradyanlar hesaplanarak, nesne kenarları ve diğer önemli özellikler belirlenir.

Şekil 1’de örnek olarak girdi görüntüsündeki gradyanlar hesaplanıp, piksel renk değerlerindeki keskin değişiklikler belirlenmiş, bu değişiklikler sonucunda nesne kenarları ve diğer önemli özellikler vurgulanmıştır.



Şekil 1. (a) Girdi görüntüsü, (b) HOG özellik tanımlayıcısı

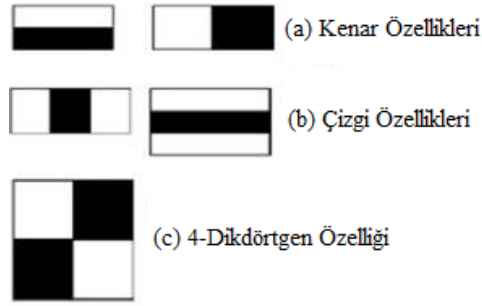
2.4.2. HAAR

Viola ve Jones tarafından öne sürülmüş piksel tarama bazlı HAAR özellik kademeli sınıflandırıcılar, nesnelerin tespiti için uygulanmış sınıflandırıcılarda yüksek performans üretmektedir [26]. Shamrat vd., değişik ortamlardaki farklı kişilerin yüzlerini tespit eden integral görüntüler sayesinde herhangi bir boyutun HAAR benzeri karakteristiklerini sabit zamanda hesaplayan bir sistem önermişlerdir [27].

HAAR sınıflandırıcısı, yüz tanımda kullanılan bir makine öğrenme yöntemidir. Yüzler ve yüz ifadeleri de dahil olmak üzere fotoğraflardaki öğeleri algılamak için kademeli bir piksel tarama işlemi kullanır. Algoritma, görüntülerden özellikler çıkarmakta ve bunları farklı ortamlardaki yüzleri algılayan bütünleşik görüntüleri kullanarak sabit zamanda hesaplamaktadır. Bu algoritmanın seçilme sebebi, hızlı işlem kapasitesi ve düşük donanım gereksinimidir. Avantajı, düşük hesaplama gücü gerektiren ortamlarda bile hızlı sonuç vermesidir. Ancak dezavantajı, düşük aydınlatma koşullarında veya karmaşık arka planlarda başarı oranının düşebilmesidir. Uygulama sürecinde, girdi görüntüsündeki piksel tarama işlemi ile nesnelere tanınır ve HAAR özellikleri yardımıyla sınıflandırılır.

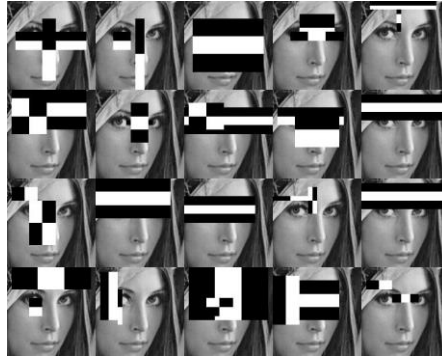
Şekil 2’de verilen HAAR özellikleri, genellikle farklı biçimlerde dikdörtgen şeklindeki bölgeye yoğunlaşan bir tür görüntü özelliğini temsil eder. Özellik vektörleri, görüntü üzerinde belirli

bölgelerdeki parlaklık farklarını ölçer. Bu özelliklerin kullanımı, nesne tanıma algoritmalarının hızlı ve etkili çalışmasını sağlamak için tasarlanmıştır.



Şekil 2. HAAR özellikleri; (a) kenar, (b) çizgi ve (c) 4-dikdörtgen özellikleri

Şekil 3'te ise örnek olarak HAAR özelliklerinin insan yüzündeki farklı bölgelere uygulanması verilmiştir. HAAR özellikleri, bir görüntüyü adım adım taramak üzere oluşturulmuş bir dizi sınıflandırıcı (classifier) kullanılarak uygulanır. Bu sınıflandırıcılar genellikle Cascade (Kademeli) adı verilen bir yapıda düzenlenir. Her bir aşamada, bir bölgeyi bir nesne olarak tanımlama olasılığını artıran veya azaltan sınıflandırıcılar bulunur.

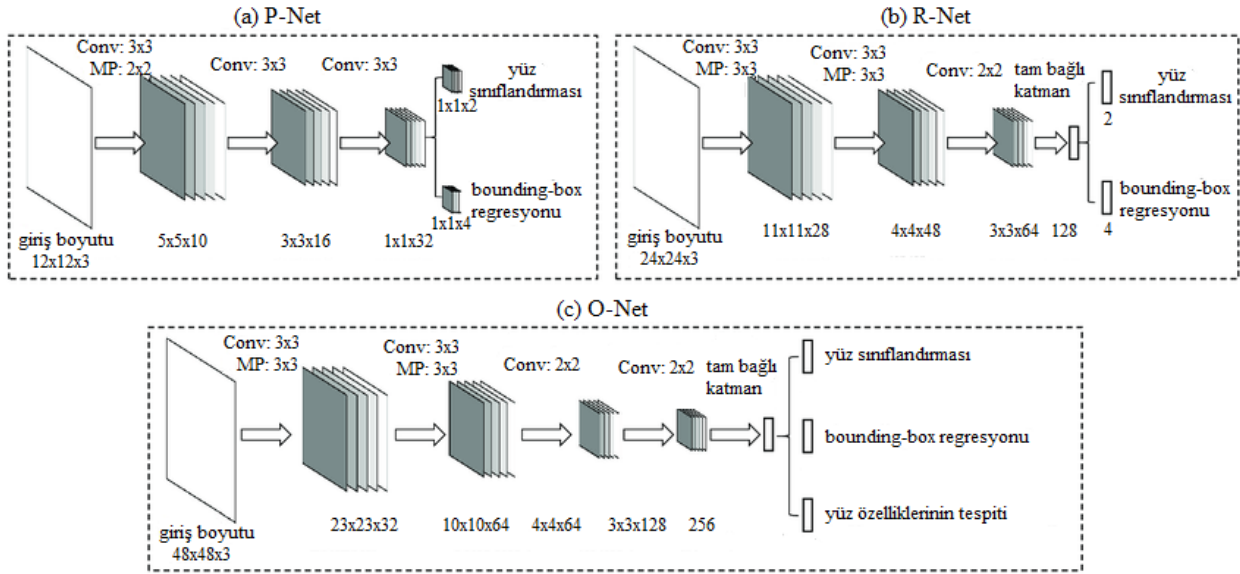


Şekil 3. HAAR özelliklerinin insan yüzüne uygulanmış biçimleri

2.4.3. MTCNN

MTCNN, yüz algılama ve yüz özelliklerinin tespiti için kullanılan bir modeldir [28]. Bu model, üç aşamalı bir yapıya sahiptir ve her aşamada farklı görevleri gerçekleştirir. P-Net (Proposal Network), ilk aşamada, potansiyel yüz bölgelerini ve bu bölgelerin olasılıklarını tahmin eder. R-Net (Refinement Network), P-Net'ten gelen aday bölgeleri daha da iyileştirir ve doğruluğu artırır. O-Net (Output Network), Son aşamada, yüz bölgelerinin kesin sınırlarını belirler ve yüz özelliklerini tespit eder. MTCNN, çeşitli yüz tanıma uygulamalarında yaygın olarak kullanılmaktadır. Seçilme nedeni, hem hızlı hem de hassas yüz tespiti yapabilmesi ve ön eğitilmiş derin öğrenme modelleri ile uyumlu çalışabilmesidir. Avantajı, karmaşık ve zorlu koşullarda bile yüksek başarı oranı sunmasıdır. Dezavantajı, yüksek işlem gücü ve kaynak gerektirebilmesidir.

Uygulama sürecinde, MTCNN modeli P-Net, R-Net ve O-Net olmak üzere üç farklı ağ kullanarak yüzlerin yerlerini ve özelliklerini belirler. Şekil 4'te MTCNN modelinin bileşenleri verilmiştir.



Şekil 4. MTCNN aşamaları

2.5 Kullanılan Optimizasyon Parametreleri

Önerilen sistemde, Keras VGG-Face kütüphanesi ile eğitim dahilinde, makine öğrenimi ve derin öğrenme alanlarında yaygın kullanılan Adam, Adagrad, RMSProp ve SGD optimizasyon algoritmaları, model eğitiminde yardımcı parametre olarak kullanılmış olup, sonuçta bu 4 farklı parametrenin ürettiği performanslar kıyaslanmıştır. Bu algoritmalar, modelin eğitim sürecinde parametrelerini güncellemek ve kayıp fonksiyonunu minimize etmek için kullanılmaktadır. Bu parametrelerin kısaca özellikleri aşağıda verilmiştir;

2.5.1. Adagrad

Adagrad, özellikle seyrek veri kümelerinde etkili olup, adaptif öğrenme hızları sağlar. Ancak, derin öğrenme modellerinde, eğitim ilerledikçe güncellemelerin yavaşlamasına yol açabilir. Bu, öğrenme hızının azalmasına neden olur ve modelin performansını olumsuz etkileyebilir [29].

2.5.2. Adam

Adam ise, hem gradyanların hem de gradyan karelerinin hareketli ortalamalarını kullanarak parametre güncellemeleri yapar. Bu yöntem, genellikle hızlı ve etkili sonuçlar verir. Ancak, modelin büyüklüğüne ve karmaşıklığına bağlı olarak farklı öğrenme oranları gerektirebilir. Adam'ın dinamik yapısı, çeşitli veri kümeleri üzerinde istikrarlı performans sunmasını sağlar [30].

2.5.3. RMSProp

RMSProp, Adagrad'ın öğrenme hızındaki düşüş sorununu hafifletmek için geliştirilmiştir. Gradyan karelerinin ağırlıklı hareketli ortalamasını kullanarak, öğrenme hızını sabit tutar ve genellikle daha iyi bir denge sağlar. Ancak, ek hiper parametre ayarları gerektirebilir, bu da modelin hassas ayarını zorlaştırabilir [31].

2.5.4. SGD

SGD (Stokastik Gradyan İnişi), basit bir gradyan iniş algoritması olarak her bir eğitim örneği için gradyan hesaplayarak parametreleri günceller. SGD, yerel minimumlara takılma eğilimindedir ve hiper parametre ayarı gerektirir. Bu algoritma, genellikle dalgalı gradyanlar nedeniyle istikrarsız sonuçlar üretebilir [32].

2.6 Kullanılan Yardımcı Algoritmalar

2.6.1. Kosinüs Benzerliği

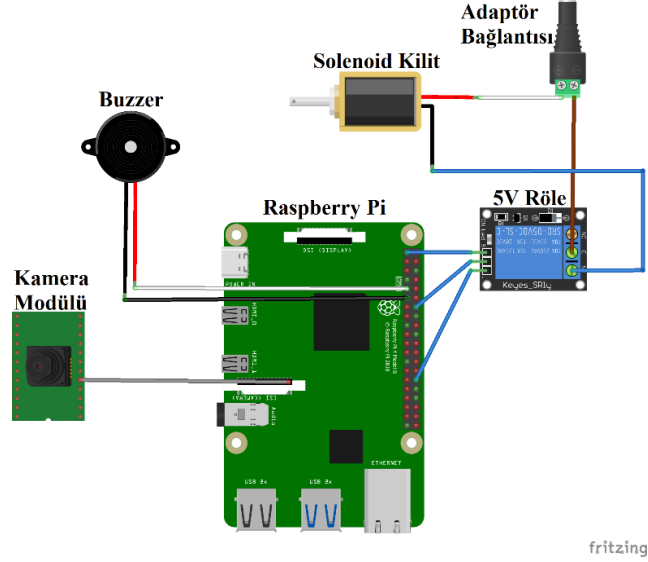
Kosinüs benzerliği, özellikle metin sınıflandırma ve bilgi erişimi gibi alanlarda yaygın olarak kullanılan bir yöntemdir [33]. İki vektör arasındaki benzerlik derecesini ölçmek için kullanılır ve 0 ile 1 arasında değerler alır. Kosinüs benzerliğinin 0 olması, iki vektörün tamamen farklı olduğunu, 1 olması ise tamamen aynı olduklarını gösterir. Bu çalışmada, 224x224 boyutlarında yüz görüntülerinden elde edilen öznitelikler, bir boyutlu dizilere dönüştürülerek kosinüs benzerliği yöntemi ile analiz edilmiştir. Elde edilen kosinüs benzerliği sonucu, güven puanı olarak değerlendirilmiş ve sonuçlara eşik değeri uygulanarak, tahmin edilen değer 0 veya 1 olarak yuvarlanmıştır. Bu yöntem, yüz tanıma doğruluğunu artırmak ve yanlış pozitifleri azaltmak amacıyla kullanılmıştır. Denklem 1'de verilen Kosinüs benzerliği formülünden elde edilen sonuca eşik değeri uygulanarak tahmin edilen sonuç 0 ya da 1 olarak belirlenmiştir.

$$S_C(A, B) = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} \quad (1)$$

2.7 Projede Kullanılan Donanım Kaynakları

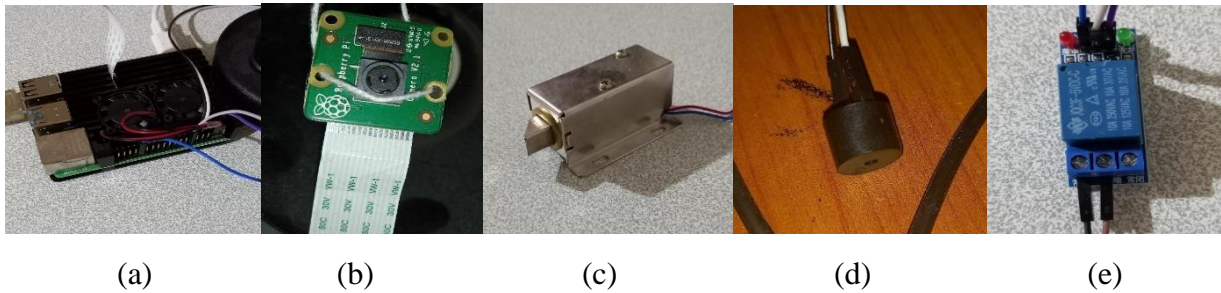
Önerilen sistem, çeşitli donanım bileşenlerinden oluşmaktadır. Bu donanımlar, yüz tanıma sistemi için gerekli işlevleri yerine getirmek amacıyla kullanılmıştır. Sistem, bir adet geliştirme

kartı, bir adet röle, bir adet buzzer, ve bir solenoid kapı kilidinden oluşmaktadır. Bu bölümde, kullanılan donanımlar hakkında temel bilgiler verilmiştir. Şekil 5, önerilen sistemin Fritzing platformunda çizilmiş devre şemasını göstermektedir.



Şekil 5. Çalışmanın tasarımı

Bu şemada, programlanabilir kart üzerinde yer alan 12. pin'in 5 Volt'luk bir röleyi, 17. pin'in ise bir buzzer'ı tetiklediği ve diğer iki pinin topraklama için ayrıldığı göstermektedir. Kullanılan donanım, çeşitli I/O pinlerine sahip, Raspberry Pi 4 Model B 8 GB versiyonudur. Çalışmada kullanılan donanım bileşenlerinin fotoğrafları Şekil 6'da verilmiştir.



Şekil 6. Donanım bileşenleri: (a) Raspberry pi 4 model b, (b) Pi kamera V2, (c) Solenoid kapı kilidi, (d) buzzer, (e) 5V röle modülü

Şekil 6(a) Raspberry Pi 4 Model B:

- Bu cihaz, bir bilgisayar için gerekli olan işlemci, bellek ve giriş-çıkış gibi tüm birimleri tek bir kartta toplayan mini bir bilgisayardır. Raspberry Pi 4 Model B'nin spesifikasyonları şu şekildedir:
 - Dört çekirdekli 1.5 GHz ARM Cortex-A72 CPU

- 64-bit SoC Broadcom 2711 entegre
- 8GB LPDDR4 RAM
- 4kp60 HEVC video destekli VideoCore VI Grafik Ünitesi
- 5V-3A USB-C güç girişi
- Gigabit Ethernet portu
- 2 adet mikro HDMI portu, 4K video desteği (1 × 4K@60Hz veya 2 × 4K@30Hz)
- 2 adet USB 3.0 ve 2 × USB 2.0 portları
- Bluetooth 5.0 BLE
- Boyutlar: 68.63 x 94.09 x 26.63 mm

Şekil 6(b) Raspberry Pi Camera V2:

- Raspberry Pi tarafından üretilen ve Raspberry Pi kartıyla uyumlu olan bu kamera modülü, kısa ribbon kablo ile direkt olarak kartın CSI konektörüne bağlanabilmektedir. Çalışmada, gerçek zamanlı video aktarımı bu kamera modülü ile gerçekleştirilmektedir.

Şekil 6(c) Solenoid Kapı Kilidi:

- Elektromanyetik bir mekanizma görevi gören bu modül, başlangıçta kapalı durumda olan kilidi açıp kapatır. Yüz tanıma işlemi sonucunda kişi tanınırsa kilit açılır, tanınmazsa kapalı kalır.

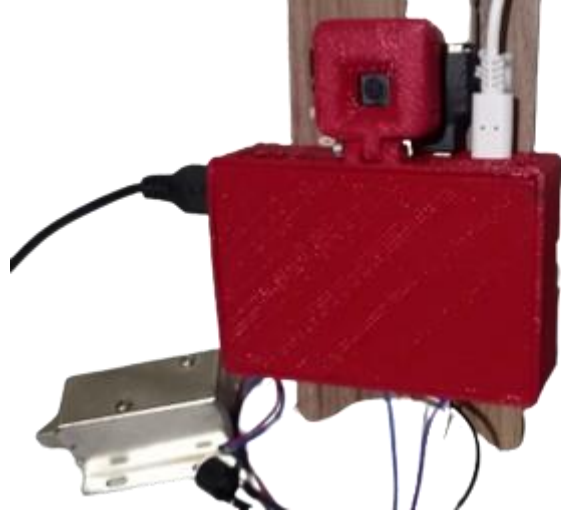
Şekil 6(d) Buzzer:

- Yüz tanıma işlemi sonucunda, kişi tanınırsa kısa bip sesi, tanınmazsa uzun bip sesi üretmektedir.

Şekil 6(e) 5V Röle Modülü:

- Bu modül, bağlandığı GPIO pinini kontrol ederek solenoid kilidi açma ve kapama işlemlerini gerçekleştirir.

Şekil 7’de Raspberry Pi’in kapalı kutu sistemi görülmektedir. Bu sistem, donanım bileşenlerini koruma amacıyla tasarlanmıştır ve tüm işlevlerin düzgün çalışmasını sağlamak için gerekli bağlantılar içerir.

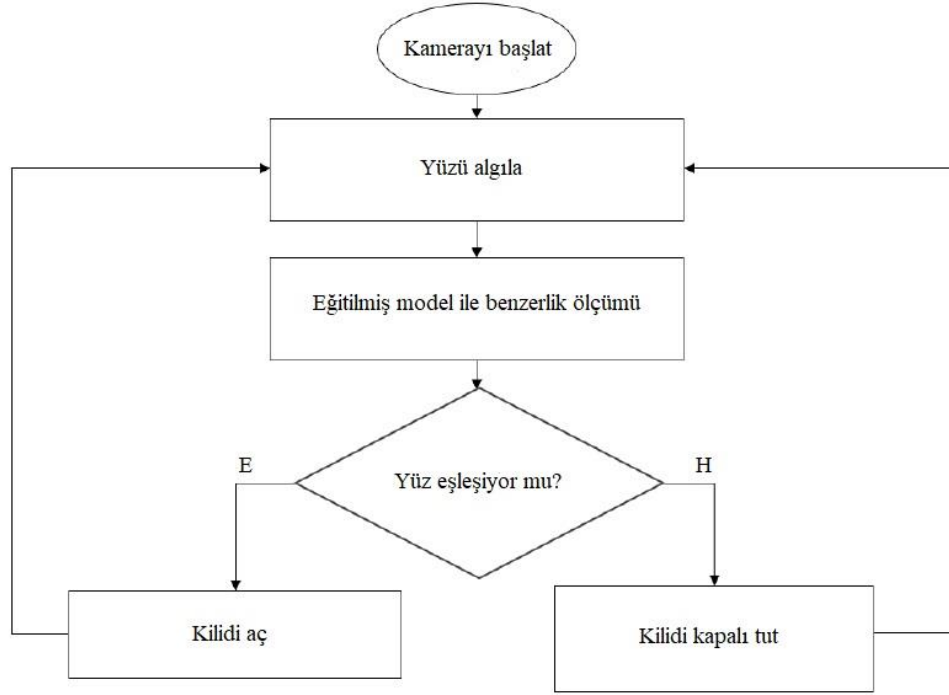


Şekil 7. Raspberry pi kapalı kutu sistemi

Raspberry Pi kartı çalıştırıldığında, 64-bit Raspberry Pi OS (eski adıyla Raspbian) işletim sistemi açılmaktadır. Bu işletim sistemi üzerinde, Visual Studio Code ortamında geliştirilmiş iki farklı Python kod dosyası bulunmaktadır.

Birinci kod dosyasında, Keras VGG-Face kütüphanesi kullanılarak ResNet50 ve VGG16 modelleri ile eğitim gerçekleştirilmiştir. Model eğitimi sırasında Adam, Adagrad, RMSProp ve SGD gibi optimizasyon parametreleri kullanılmıştır. Eğitim sürecinin sonucunda, her biri .h5 formatında olmak üzere toplam 8 adet model dosyası üretilmiştir. Bu modellerle, HOG ve HAAR öznitelikleri kullanılarak yüz tanıma işlemleri gerçekleştirilmiştir.

İkinci kod dosyasında, FaceNet-PyTorch kütüphanesinden ön eğitilmiş InceptionResNetV1 modeline odaklanılmıştır. FaceNet modeli, yüz tanıma işlemi için genellikle MTCNN yüz algılama algoritması ile birlikte kullanılır. Model, yeni kişilerin yüz verilerini mevcut veri setine ekleyerek modelin doğruluğunu artırır. Model, yüzleri temsil eden gömme vektörleri çıkararak, bu vektörleri kullanarak kişileri tanırlar ve karşılaştırırlar. Yeni yüz verileri, mevcut eğitim veri setine eklenerek modelin doğruluğunu artırmak için kullanılabilir. Şekil 8’de çalışmanın iş akış şeması verilmiştir.



Şekil 8. Çalışmanın iş akış diyagramı

İlk olarak, kamera açılarak kişinin yüzü algılanır. Algılanan yüz, eğitilmiş model kullanılarak tanınır. Kosinüs benzerliği yöntemiyle tahmin edilen sonuçlar değerlendirilir. Tahmin edilen yüzün, tanıma başarısı bir yüzdelik başarı skoru ile ifade edilir. Eğer yüz tanınırsa, kilit açılır; aksi takdirde kilit kapalı kalır. Sistem, kullanıcı programı kapatana kadar yüz algılama işlemini sürekli olarak gerçekleştirir.

2.8 Veri Seti

Bu çalışmada, Stirling Üniversitesi'ndeki Psikolojik Görüntü Koleksiyonu'ndan (PICS) "Aberdeen" veri kümesi kullanılmıştır [34]. Bu veri kümesi, 90 kişinin 687 renkli yüz görüntüsünü içermektedir. Eğitim sürecinde, 18'den az örneğe sahip olan görüntüler veri kümesinden çıkarılmıştır. Sonuç olarak, 18 ve daha fazla resim içeren kişilerden seçilen 18'er örnek ile toplam $18 \times 23 = 414$ yüz resmi içeren bir veri kümesi oluşturulmuştur. Şekil 9'da çalışmada kullanılan veri setinden örnek yüz resimleri verilmiştir.



Şekil 9. Veri setinden örnek yüz resimleri

Keras VGG-Face ile model eğitimi sırasında veri kümesinin %75'i eğitim, %25'i test için kullanılmış olup 310 adet yüz verisi eğitim için, 104 adet yüz verisi ise test için ayrılmıştır. Eğitim ve test aşamalarında RGB giriş görüntüleri varsayılan 224x224 piksel boyutlarında kullanılmıştır. Batch size ise modelin kapasitesine göre 16 olarak belirlenmiştir. Daha düşük bir batch size, eğitim aşamalarının yavaşlamasına sebep olurken, daha yüksek bir batch size ise Raspberry Pi'nin donanım kaynaklarının yetersiz kalmasına ve bellek uyarısına sebep olabilmektedir.

Keras VGG-Face ile doğrudan ağı yeniden eğitmeden yeni kişiler eklemek genellikle mümkün değildir. Bu senaryoyu önerilen sisteme uygulamak amacıyla, FaceNet-PyTorch kütüphanesinde bulunan ön eğitilmiş InceptionResNetV1 modeli kullanılmıştır. FaceNet, yüzlerin gömme vektörlerini oluşturarak, bu vektörler üzerinden yeni kişileri sisteme entegre etmeye olanak sağlar. Bu yöntem, mevcut modelin yeniden eğitilmesine gerek kalmadan yeni yüzleri hızlı bir şekilde sisteme eklemeyi sağlar. Bu yaklaşım, modelin performansını artırırken sistem kaynaklarının daha etkin bir şekilde yönetilmesini sağlar ve büyük veri setleri ile sınırlı donanım kaynaklarıyla çalışırken avantaj sağlar.

2.9 Değerlendirme Metriği

Çalışmanın performansını değerlendirmek için kullanılan metrik, her bir yöntem kombinasyonu için Doğruluk (Accuracy) değerlerini sağlamaktadır. Doğruluk değeri, aşağıdaki Denklem 2 ile hesaplanmaktadır:

$$\frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

Bu denklemde TP (True Positive) doğru tahmin edilmiş pozitif değerler, TN (True Negative) doğru tahmin edilmiş negatif değerler, FP (False Positive) yanlış tahmin edilmiş pozitif değerler, FN (False Negative) yanlış tahmin edilmiş negatif değerler olarak temsil edilmektedir.

3. BULGULAR

Bu bölümde, ResNet50 ve VGG16 modellerinin, Adagrad, Adam, RMSProp, ve SGD optimizasyon algoritmalarının, HOG ve HAAR özneliklerinin kullanıldığı kombinasyonların performansları ile InceptionResNetV1 ve MTCNN yüz tanıma algoritmalarının değerlendirilmesi yapılmıştır. Elde edilen doğruluk değerleri, çeşitli eşik seviyeleri için (%75, %80, %85, %90, %92, %94, %96 ve %98) Tablo 2, 3, 4, 5, 6 ve 7'de özetlenmiştir.

Tablo 2. HOG öz niteliklerinin kullanılması ile %75, %80 ve %85 eşik seviyeleri altında başarı performansını gösteren sonuç tablosu

Kullanılan CNN Modeli	Optimizasyon Algoritması	%75 Eşik Seviyesindeki Sonuçlar	%80 Eşik Seviyesindeki Sonuçlar	%85 Eşik Seviyesindeki Sonuçlar
ResNet50	Adam	%100	%100	%100
ResNet50	Adagrad	%100	%87	%35
ResNet50	RMSProp	%100	%100	%100
ResNet50	SGD	%100	%74	%43
VGG16	Adam	%100	%100	%100
VGG16	Adagrad	%100	%96	%65
VGG16	RMSProp	%100	%100	%100
VGG16	SGD	%100	%65	%30

Tablo 2 incelendiğinde ResNet50 modeli ile Adam ve RMSProp optimizasyon algoritmaları, tüm eşik seviyelerinde %100 başarı göstermektedir. Adagrad ve SGD algoritmaları, %85 eşik seviyesinde belirgin bir başarı düşüşü göstermektedir. VGG16 modeli ile benzer sonuçlar elde edilmiştir; Adam ve RMSProp %100 başarı sağlarken, Adagrad ve SGD'de %85 eşik seviyesinde başarı kayıpları görülmektedir.

Tablo 3. HOG öz niteliklerinin kullanılması ile %90, %92, %94, %96 ve %98 eşik seviyeleri altında başarı performansını gösteren sonuç tablosu

Kullanılan CNN Modeli	Optimizasyon Algoritması	%90 Eşik Seviyesindeki Sonuçlar	%92 Eşik Seviyesindeki Sonuçlar	%94 Eşik Seviyesindeki Sonuçlar	%96 Eşik Seviyesindeki Sonuçlar	%98 Eşik Seviyesindeki Sonuçlar
ResNet50	Adam	%100	%96	%91	%83	%30
ResNet50	Adagrad	%4	X	X	X	X
ResNet50	RMSProp	%100	%91	%70	%57	%9
ResNet50	SGD	X	X	X	X	X
VGG16	Adam	%100	%100	%83	%43	%4
VGG16	Adagrad	%22	%9	X	X	X
VGG16	RMSProp	%91	%65	%30	%4	X
VGG16	SGD	%4	%4	X	X	X

X: Başarısız

Tablo 3 incelendiğinde ResNet50 modeli ile Adam ve RMSProp algoritmaları, %90-98 eşik seviyeleri arasında kademeli bir başarı düşüşü göstermektedir. Adagrad ve SGD algoritmaları, %90 üzerindeki eşik seviyelerinde başarısız olmuştur. VGG16 ile Adam ve RMSProp algoritmaları da benzer şekilde başarı kayıpları göstermektedir.

Tablo 4. HAAR özneteliklerinin kullanılması ile %75, %80 ve %85 eşik seviyeleri altında başarı performansını gösteren sonuç tablosu

Kullanılan CNN Modeli	Optimizasyon Algoritması	%75 Eşik Seviyesindeki Sonuçlar	%80 Eşik Seviyesindeki Sonuçlar	%85 Eşik Seviyesindeki Sonuçlar
ResNet50	Adam	% 100	% 100	% 100
ResNet50	Adagrad	% 100	% 70	% 17
ResNet50	RMSProp	% 100	% 100	% 100
ResNet50	SGD	% 100	% 65	% 17
VGG16	Adam	% 100	% 100	% 100
VGG16	Adagrad	% 100	% 96	% 61
VGG16	RMSProp	% 100	% 100	% 100
VGG16	SGD	% 100	% 52	% 13

Tablo 4 incelendiğinde ResNet50 ve VGG16 modelleri ile Adam ve RMSProp optimizasyon algoritmaları, %75 ve %80 eşik seviyelerinde %100 başarı sağlamaktadır. Adagrad ve SGD algoritmaları, %85 eşik seviyesinde belirgin bir düşüş göstermektedir.

Tablo 5. HAAR özneteliklerinin kullanılması ile %90, %92, %94, %96 ve %98 eşik seviyeleri altında başarı performansını gösteren sonuç tablosu

Kullanılan CNN Modeli	Optimizasyon Algoritması	%90 Eşik Seviyesindeki Sonuçlar	%92 Eşik Seviyesindeki Sonuçlar	%94 Eşik Seviyesindeki Sonuçlar	%96 Eşik Seviyesindeki Sonuçlar	%98 Eşik Seviyesindeki Sonuçlar
ResNet50	Adam	% 100	% 96	% 87	% 70	% 30
ResNet50	Adagrad	% 4	X	X	X	X
ResNet50	RMSProp	% 96	% 78	% 78	% 39	X
ResNet50	SGD	X	X	X	X	X
VGG16	Adam	% 100	% 96	% 78	% 26	X
VGG16	AdaGrad	% 4	% 4	X	X	X
VGG16	RMSProp	% 78	% 52	% 13	X	X
VGG16	SGD	X	X	X	X	X

X: Başarısız

Tablo 5 incelendiğinde ResNet50 ve VGG16 modelleri, %90 üzerindeki eşik seviyelerinde başarı kayıpları yaşamaktadır. Adam ve RMSProp algoritmaları, %96 ve %98 eşik seviyelerinde başarısızlığa yakın sonuçlar göstermektedir.

Keras VGG-Face kütüphanesi ile birlikte kullanılan VGG16 ve ResNet50 modellerinin ürettiği performans verileri incelendiğinde, ResNet50 modeli, Adam optimizasyon algoritması ile HOG özniteliğinin kullanıldığı kombinasyon doğruluk ortalaması bakımından en yüksek performansı sergilemiştir.

Tablo 6 ve 7, MTCNN algoritması ile birlikte kullanılan ön eğitimli InceptionResNetV1 modelinin performansını göstermektedir. Bu modelin başarı oranları, çeşitli eşik seviyelerinde incelenmiştir:

Tablo 6. MTCNN algoritması kullanılması ile %75, %80 ve %85 eşik seviyeleri altında başarı performansını gösteren sonuç tablosu

Kullanılan CNN Modeli	%75 Eşik Seviyesindeki Sonuçlar	%80 Eşik Seviyesindeki Sonuçlar	%85 Eşik Seviyesindeki Sonuçlar
InceptionResnetV1	%100	%100	%99

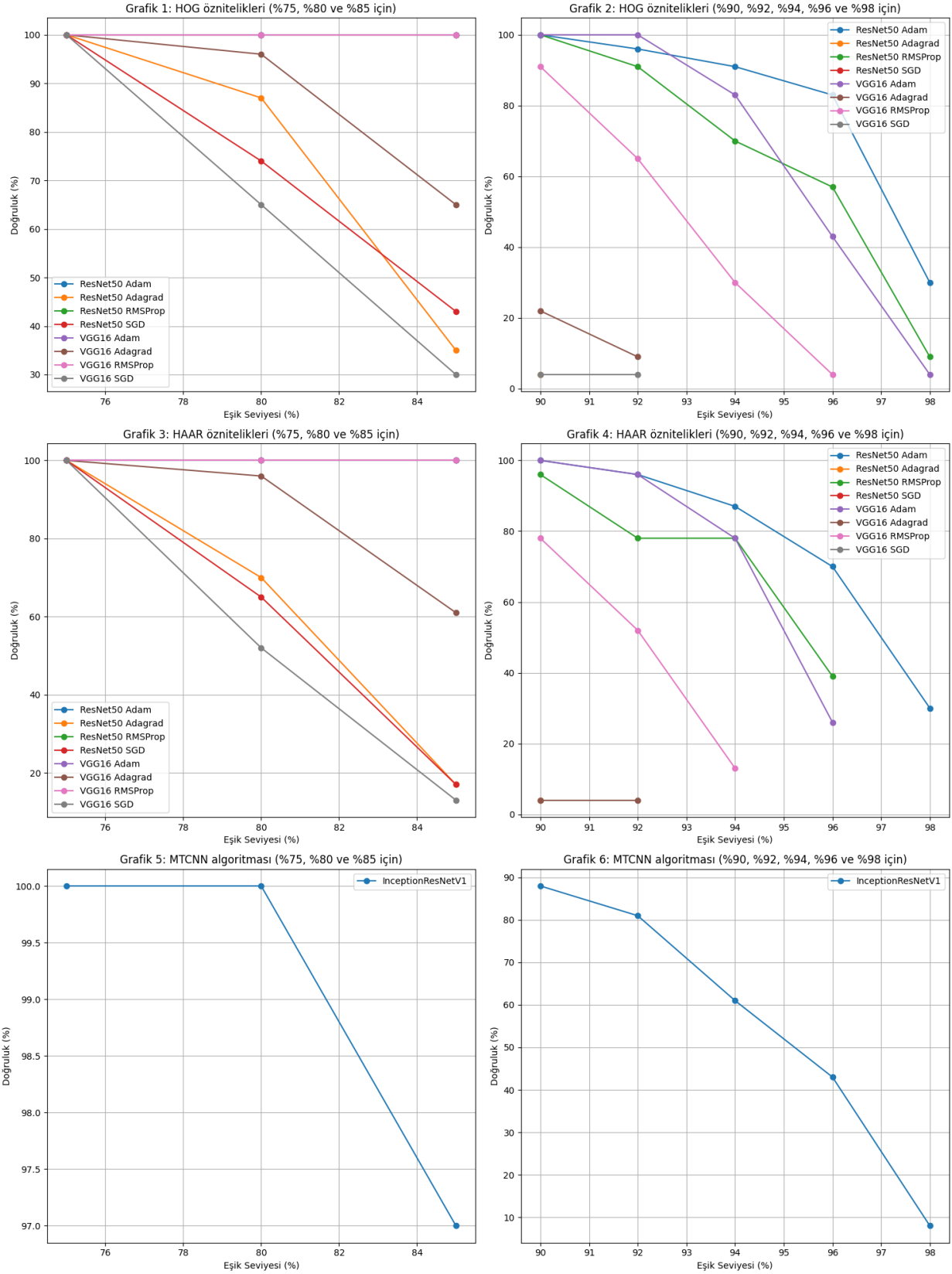
Tablo 6 incelendiğinde, InceptionResNetV1 modelinin %75 ve %80 eşik seviyelerinde %100 başarı sağladığı, ancak %85 eşik seviyesine gelindiğinde başarı oranının %99'a düştüğü görülmüştür.

Tablo 7. MTCNN algoritması kullanılması ile %90, %92, %94, %96 ve %98 eşik seviyeleri altında başarı performansını gösteren sonuç tablosu

Kullanılan CNN Modeli	%90 Eşik Seviyesindeki Sonuçlar	%92 Eşik Seviyesindeki Sonuçlar	%94 Eşik Seviyesindeki Sonuçlar	%96 Eşik Seviyesindeki Sonuçlar	%98 Eşik Seviyesindeki Sonuçlar
InceptionResnetV1	%91	%81	%66	%42	%10

Tablo 7 incelendiğinde InceptionResnetV1 modeli, %90 ve üzeri eşik seviyelerinde başarı kaybı yaşar. Özellikle %98 eşik seviyesinde başarı %10'a kadar düşmektedir.

Şekil 10'da karşılaştırılan kombinasyonlara ait sonuç doğruluk değerini gösteren grafikler verilmiştir.



Şekil 10. Karşılaştırılan model-optimizasyon çiftlerinin doğruluk değerleri

4. DENEYSEL SONUÇLAR VE TARTIŞMA

Bu çalışmada, yüz tanıma uygulamalarında en yüksek performans çıktısını Adam optimizasyon algoritması ve ResNet50 modeli ile HOG özniteliklerinin kullanıldığı kombinasyon sağlamıştır. Çalışma kapsamında elde edilen sonuçlar, farklı eşik seviyelerinde (%75, %80, %85, %90, %92, %94, %96 ve %98) yapılan testlerle farklı model, parametre ve yüz tanıma algoritması kombinasyonlarının performanslarını göstermektedir.

4.1 Başarı Faktörleri

Adam ve RMSProp optimizasyon algoritması, yüksek performans göstermiş ve doğru tanıma oranlarında %100'e kadar başarı sağlamıştır. Özellikle Adam, çeşitli eşik seviyelerinde istikrarlı ve yüksek doğruluk oranları sunarak bu çalışmada en etkili parametre olarak öne çıkmıştır.

Adagrad ve SGD algoritmaları ise genel olarak düşük performans göstermiştir. Adagrad, özellikle yüksek eşik seviyelerinde başarısız sonuçlar vermiştir. SGD (Stokastik Gradient Descent) ise genellikle düşük doğruluk oranlarına sahip olmuş ve yüz tanıma çalışmalarında kullanılması önerilmemektedir.

4.2 Model Performansı

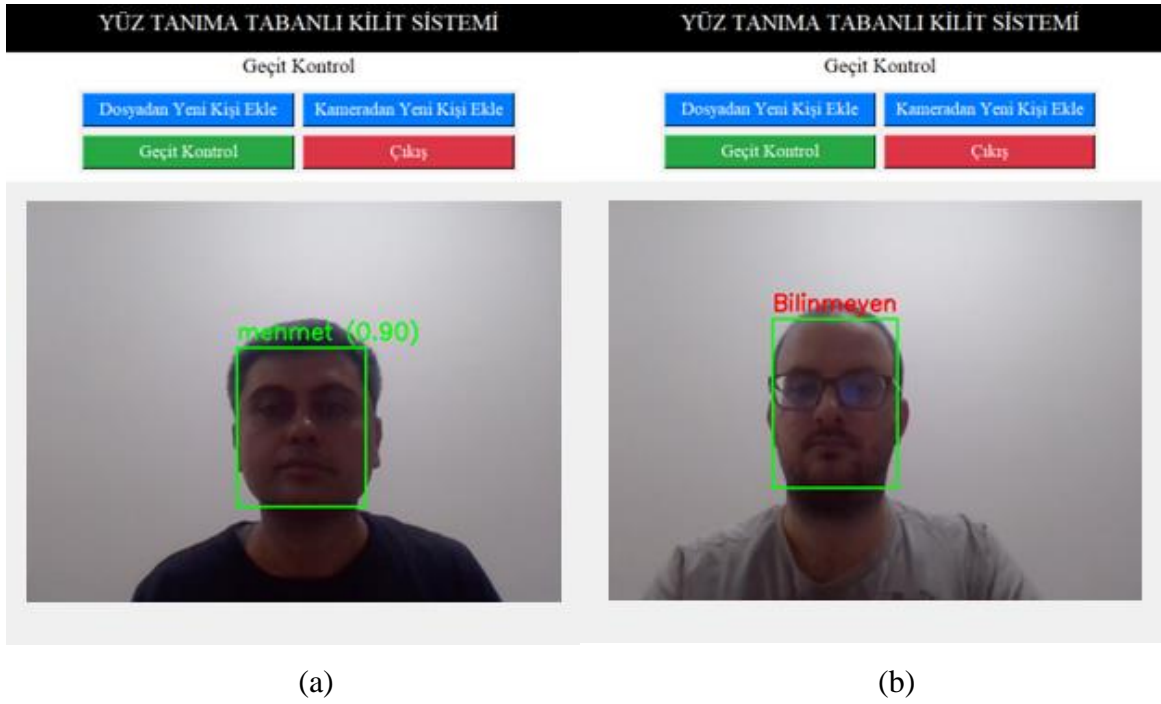
Çalışmada kullanılan farklı optimizasyon algoritmalarının aynı model üzerinde farklı performanslar sergilediği gözlemlenmiştir. Bu durum, optimizasyon algoritmasının modelin başarısını önemli ölçüde etkilediğini göstermektedir. Özellikle Adam ve RMSProp algoritmalarının yüksek başarı sağlaması, bu algoritmaların model performansını olumlu yönde etkilediğini ortaya koymaktadır.

HAAR ve HOG öznitelikleri ile yapılan karşılaştırmalarda, HOG öznitelikleri ile elde edilen sonuçlar genellikle daha yüksek doğruluk oranlarına sahip olmuştur. HAAR özniteliklerinin bazı eşik seviyelerinde düşük performans göstermesi, yüz tanıma süreçlerinde HOG'un daha uygun bir seçenek olabileceğini göstermektedir.

InceptionResNetV1 modeli, özellikle düşük eşik seviyelerinde yüksek performans göstermiş, ancak yüksek eşik seviyelerinde belirgin bir başarısızlık sergilemiştir. Bu sonuçlar, modelin kullanım senaryolarına göre dikkatli bir şekilde seçilmesi gerektiğini ortaya koymaktadır. MTCNN algoritması ile birlikte kullanıldığında, modelin düşük eşik seviyelerinde (%75-%85) mükemmel yakın performans gösterdiği, ancak eşik seviyesi arttıkça doğruluğun hızla düştüğü

gözlemlenmiştir. Bu durum, MTCNN'in yüksek eşik seviyelerinde daha az etkili olduğunu ve ince ayar gerektirdiğini göstermektedir.

Şekil 11(a)'da tanınan yüz karşısında kim olduğu ve tahmin değeri paylaşılırken kapı kilidi açılmış, Şekil 11(b)'de ise tanınmayan bir yüz örneği olduğu için “Bilinmeyen Kişi” etiketi döndürülerek kilit açılmamıştır.



Şekil 11. Önerilen Sistem sonucu: (a) Veri kümesinde bulunan kişi, (b) Veri kümesinde bulunmayan kişi

4.3 Bulguların Özeti ve Performans Analizi

- **ResNet50 ve VGG16 Modelleri:** Adam ve RMSProp algoritmaları ile HOG özniteliklerinin kullanılması yüksek doğruluk oranları sağlamıştır. Özellikle Adam, çeşitli eşik seviyelerinde istikrarlı ve yüksek doğruluk oranları sunarak bu çalışmada en etkili parametre olarak öne çıkmıştır. Bunun yanında, RMSProp da çoğu durumda başarılı sonuçlar elde etmiştir, ancak bazı eşik seviyelerinde Adam algoritmasının gerisinde kalmıştır. HOG öznitelikleriyle elde edilen sonuçlar, genellikle HAAR özniteliklerine kıyasla daha iyi bir doğruluk sağlamış, bu da HOG'un yüz tanıma ve öznitelik çıkarma açısından daha etkili olduğunu göstermektedir.
- **InceptionResNetV1 Modeli:** MTCNN algoritması ile birlikte kullanılan bu model yüksek başarı oranları sağlarken, %90 ve üzeri eşik seviyelerinde performans kayıpları yaşanmıştır. Özellikle %96 ve %98 eşik seviyelerinde doğruluk oranları önemli ölçüde düşmüş, bu da modelin daha yüksek hassasiyet gerektiren durumlarda performansının

düştüğünü göstermektedir. Ancak, daha düşük eşik seviyelerinde %100'e yakın doğruluk oranları sunarak bu seviyelerde güçlü bir performans sergilemiştir. Bu sonuç, modelin daha esnek ve daha genel kararlar vermesi gereken durumlarda daha etkili olduğunu, ancak yüksek doğruluk gerektiren özel durumlarda başarısının azaldığını göstermektedir.

Çalışmanın bulguları, yüz tanıma sistemlerinin kapı güvenlik sistemlerinde etkin bir şekilde kullanılabileceğini göstermektedir. Özellikle, yüksek doğruluk oranlarına sahip kombinasyonların kullanılması, güvenlik seviyesini artırabilir. Adam algoritması ve HOG öznelikleri gibi yüksek performans gösteren bileşenlerin kullanılması, çeşitli güvenlik koşullarına uyum sağlayan adaptif sistemlerin geliştirilmesine olanak tanır.

5. ÖNERİLER

Bu çalışmada kapı güvenlik sisteminin derin öğrenme tabanlı bir gerçekleştirimi tasarlanmış olup Keras VGG-Face tabanlı ResNet50 ve VGG16 derin öğrenme modelleri ve PyTorch FaceNet tabanlı ön eğitilmiş InceptionResNetV1 modeli arasında bir karşılaştırma gerçekleştirilmiştir. Elde edilen sonuçlar neticesinde derin öğrenme modellerinin güvenli kapı tasarımı işleminde başarılı sonuçlar verdiği gözlemlenmiştir.

Sonraki çalışmalarda sistem performansının iyileştirilmesi, bunun yanı sıra düşük aydınlatma koşullarında da başarılı olabileceği yöntemlerin elde edilebilmesi için önerilen sistemin geliştirilmesi harici bunun yanında farklı bantlardan yararlanması hedeflenmektedir.

6. TEŞEKKÜR

Erciyes Üniversitesi Bilimsel Araştırma Projeleri Koordinatörlüğü'ne FYL-2022-12418 kodlu projemiz kapsamında sağladıkları destek için teşekkür ederiz.

Kaynakça

- [1] Phin, P., Abbas, H., & Kamaruddin, N. (2020). Physical security problems in local governments: A survey. *Journal of Environmental Treatment Techniques*, 8(2), 679-686.
- [2] Motwani, Y., Seth, S., Dixit, D., Bagubali, A., & Rajesh, R. (2021). Multifactor door locking systems: A review. *Materials Today: Proceedings*, 46(17), pp.7973-7979.
- [3] Thomson, G. (2005). Facial recognition. *Encyclopedia*.
- [4] Ahonen, T., Hadid, A., & Pietikainen, M. (2004). Face recognition with local binary patterns. In *Proceedings of the European Conference on Computer Vision*, pp. 469-481.

- [5] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* ,pp. 770–778.
- [6] Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. In Y. Bengio & Y. LeCun (Eds.), *Proceedings of ICLR*.
- [7] Hasan, R. I., Yusuf, S. M., & Alzubaidi, L. (2020). Review of the state of the art of deep learning for plant diseases: A broad analysis and discussion. *Plants*, 9(10), 1302.
- [8] Ker, J., Wang, L., Rao, J., & Lim, T. (2017). Deep learning applications in medical image analysis. *IEEE Access*, 6, 9375–9389.
- [9] Elmaci, M., Toprak, A. N., & Aslantas, V. (2023). Detection of background forgery using a two-stream convolutional neural network architecture. *Multimedia Tools and Applications*, 1-28.
- [10] Lai, S.-C., Kong, M., Lam, K.-M., & Li, D. (2019). High-resolution face recognition via deep pore-feature matching. In *Proceedings of the IEEE International Conference on Image Processing* , pp. 3477–3481.
- [11] Gilani, S. Z., & Mian, A. (2018). Learning from millions of 3D scans for large-scale 3D face recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 1896–1905.
- [12] Kim, K., Yang, Z., Masi, I., Nevatia, R., & Medioni, G. (2018). Face and body association for video-based face recognition. In *Proceedings of the IEEE Winter Conference on Applications of Computer Vision* , pp. 39–48.
- [13] Zheng, Y., Pal, D. K., & Savvides, M. (2018). Ring loss: Convex feature normalization for face recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5089–5097.
- [14] Nagendra Kumar, J., Kumar, G., Kurikelly, R., & Sri Charan, D. (2023). Face recognition and Raspberry Pi powered smart door unlocking system. *E3S Web of Conferences*, 391, 01087.
- [15] Huang, D., Shan, C., Ardabilian, M., & Chen, L. (2011). Local binary patterns and its variants for face recognition: A comprehensive review. *International Journal of Computer Vision*, 91(1), 68-90.
- [16] Razzaque, M., Milani, A., & Iqbal, S. A. R. (2021). A comprehensive review of the Raspberry Pi: The affordable computing device for education and IoT. *Journal of Computing Sciences in Colleges*, 36(5), 50-59.
- [17] Muthu, K., Jayanthi, S., Rajesh, P., & Sharma, K. (2023). ResNet50 for face recognition: Recent advances and performance analysis. *Journal of Computer Vision*.

- [18] Smith, J., Williams, P., Zhang, Y., & Kumar, R. (2022). VGG16 in face detection and recognition: A comparative study. *Proceedings of the International Conference on Pattern Recognition*.
- [19] Jones, A., Patel, D., Kim, H., & Garcia, M. (2023). InceptionResNet50 for facial recognition: A comparative evaluation. *Computer Vision and Image Understanding*.
- [20] Lee, S., Nguyen, T., Patel, A., & Choi, B. (2021). Haar cascades for real-time face detection: An evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- [21] Chen, L., Zhao, W., Li, K., & Zhang, Q. (2022). Histogram of oriented gradients for face detection: Performance and improvements. *Journal of Artificial Intelligence Research*.
- [22] Wu, Y., Wang, X., Liu, Z., & Huang, L. (2023). MTCNN for face detection and alignment: A review and performance comparison. *IEEE Transactions on Neural Networks and Learning Systems*.
- [23] Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- [24] Szegedy, S., Vanhoucke, V., & Ioffe, S. (2016). Rethinking the Inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- [25] Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. In *Proceedings of Computer Vision and Pattern Recognition (Vol. 1, pp. 886–893)*.
- [26] Viola, P., & Jones, M. (2001). Robust real-time face detection. In *Proceedings of the Eighth IEEE International Conference on Computer Vision (ICCV 2001)*.
- [27] Shamrat, F. M. J., Majumder, A., Antu, P. R., Barmon, S. K., Nowrin, I., & Ranjan, R. (2021). Human face recognition applying Haar cascade classifier. In *Proceedings of the International Conference on Pervasive Computing and Social Networking (Salem, Tamil Nadu, India, March 19-20, 2021)*.
- [28] Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 23(10), 1499-1503.
- [29] Cohn, J. D. (2011). Adaptive subgradient methods for supervised learning. *Journal of Machine Learning Research*, 12, 1-30.
- [30] Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. In *Proceedings of ICLR*.
- [31] Hinton, G. (2012). Lecture 6a: Overview of mini-batch gradient descent and RMSProp. *Coursera: Neural Networks for Machine Learning*.

- [32] Prophet, R. (2016). Stochastic gradient descent: Basics and variants. In *Deep Learning Book* (Chapter 6).
- [33] Han, J., Kamber, M., & Pei, J. (2012). *Data mining: Concepts and techniques*. Morgan Kaufmann.
- [34] Craw, I., & Cameron, P. (1992). Face recognition by computer. In D. Hogg & R. Boyle (Eds.), *Proceedings of BMVC92*, pp. 320-330.