

Yüz İmgelerinden Göz Bölgelerinin Tespitinde ESA Tabanlı Alternatif Bir Yaklaşım

Kenan DONUK^{1*}, Ali ARI², Davut HANBAY³

¹Bilgisayar Programcılığı Bölümü, Şırnak Üniversitesi, Cizre Meslek Yüksekokulu, Şırnak, Türkiye

²Bilgisayar Mühendisliği Bölümü, İnönü Üniversitesi, Malatya, Türkiye

³Bilgisayar Mühendisliği Bölümü, İnönü Üniversitesi, Malatya, Türkiye

*¹kenandonuk@sirnak.edu.tr, ²ali.ari@inonu.edu.tr, ³davut.hanbay@inonu.edu.tr

(Geliş/Received: 22/06/2021;

Kabul/Accepted: 05/08/2021)

Öz: Artan işlemci hızlarıyla beraber kullanımı yaygınlaşan derin sinir ağları birçok farklı alanda gösterdiği başarılarla etkinliğini ispatlamıştır. Bu çalışmada bir imge seçici yardımıyla yüz üzerinden örnekler toplanarak elde edilen imgelerin derin sinir ağlarının örüntü tanımada etkili bir algoritması olan evrimsel sinir ağı modeline verilmesiyle yüz resimlerinden gözlerin tespiti gerçekleştirilmiştir. Geçmişten günümüze önerilen birçok göz algılama yöntemi mevcuttur. Fakat bu yöntemlerin birçoğunda aydınlatma koşulları, duruş pozisyonları, düşük çözünürlükteki görüntüler, kapalı göz, gözlük, gözlerin algılanmasında önemli bir sorun olmuştur. Önerilen yöntemin evrimsel sinir ağı modeli ile göz tespitinde zorluk çıkaran durumların model tarafından birçok örnek veri ile öğrenilmesiyle üstesinden gelinmiştir. Önerilen yöntemin performansı günümüzde göz tespitinde yaygın olarak kullanılan Viola-Jones algoritmasının XML tabanlı yüz ve göz tanıma uygulaması ile karşılaştırılmıştır. Karşılaştırma sonunda önerilen yöntemin gözlerin algılanmasında zorluk çıkaran yüz resimlerinde daha iyi sonuçlar verdiği görülmüştür. Doğruluk (%98,99), F1-skor (%98,99), Matthews korelasyon katsayısı (%97,99) ve R-kare (%95,98) gibi yaygın kullanılan ölçütler ile önerilen yöntemin başarısı ortaya konmuştur.

Anahtar kelimeler: Derin öğrenme, evrimsel sinir ağları, viola-jones, göz algılama, nesne algılama

An Alternative Approach Based On CNN for Detecting Eye Areas from Facial Images

Abstract: Deep neural networks, which are becoming widespread with increasing processor speeds, have proven their effectiveness with their success in many different areas. In this study, by collecting the samples over a hundred with the help of an image selector, the deep neural networks of the obtained images were given to the convolutional neural network model, which is an effective algorithm in pattern recognition, and the eyes were detected from the face images. There are many eye detection methods proposed from the past to the present. However, in many of these methods, lighting conditions, posture positions, low resolution images, closed eyes, glasses, and eyes have been an important problem in perception. With the convolutional neural network model of the proposed method, the situations that cause difficulties in eye detection have been overcome by learning by the model with many sample data. The performance of the proposed method was compared with the XML-based face and eye recognition application of the Viola-Jones algorithm, which is widely used in eye detection today. The success of the proposed method has been demonstrated by commonly used criteria such as accuracy (98.99%), F1-score (98.99%), Matthews correlation coefficient (97.99%) and R-squared (95.98%).

Key words: Deep learning, convolutional neural networks, viola-jones, eye detection, object detection

1. Giriş

Göz tespiti, sürücü davranış analizi, yüz tanıma, psikolojik analiz, adli bilişim ve bakış takibi gibi birçok alanda sistem girdisi olarak kullanılmaktadır [1-6]. Gözlerin tespit edilmesi yüz tespitinden sonraki ilk aşamadır. Göz tespitinde yapılan çalışmaların çoğunda aydınlatma, çözünürlük, duruş pozisyonları gibi birçok dış faktörün etkisi ortadan kaldırılmaya çalışılmıştır. Bu dış faktörler gözlerin doğru bir şekilde tespit edilmesini zorlaştırmaktadır. Konu ile ilgili literatürde bulunan çalışmalar şu şekilde sıralanabilir. Viola P. vd. xml tabanlı başarılı bir nesne algılama algoritması geliştirmişlerdir. Bu algoritmanın başarılı olmasında Haar-like özellikler, İmge integrali, Adaboost, Basamaklı sınıflandırıcı gibi bileşenler etkili olmuştur. Bu algoritma daha çok yüz ve göz tespitinde kullanılmaktadır [7]. Zhang K. vd. yüz ve yüzdeki dönüm noktalarını tahmin eden üç aşamalı derin evrimsel ağlara sahip basamaklı bir yapı (MTCNN) geliştirmişlerdir. Önerilen yapıda ilk olarak, yüz dönüm noktaları için aday pencereler hızlı bir şekilde P-Net ile üretilir. Bir sonraki aşamada bu aday pencereler bir R-Net aracılığıyla iyileştirilir ve son aşamada O-Net aracılığıyla son sınırlayıcı kutu ve yüz işaretlerinin konumunu üretir. Bu yapı kullanılarak imgedeki yüz, göz, ağız, burun gibi tüm dönüm noktaları algılanır [8]. Kazemi V. vd. HELEN

* Sorumlu yazar: kenandonuk@sirnak.edu.tr. ¹ 0000-0002-7421-5587, ² 0000-0002-5071-6790, ³ 0000-0003-2271-7865

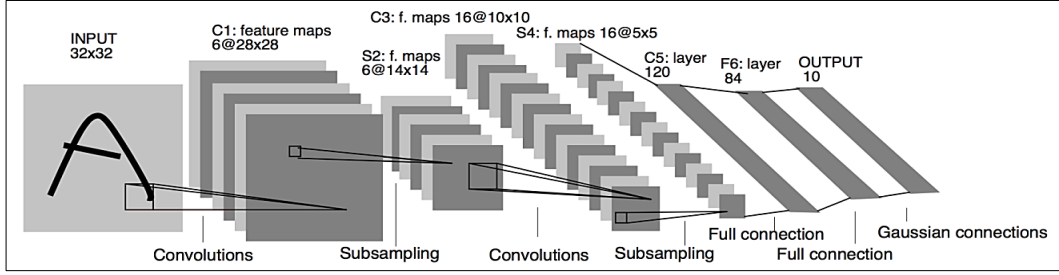
veri kümesi ile eğitilen, yüzün dönüm noktası konumlarını (yüz, göz, burun, ağız) tahmin etmek ve yüksek kaliteli tahminlerle süper gerçek zamanlı performans elde etmek için regresyon ağaçları topluluğunu kullanmışlardır. Çalışmalarında gradyan arttırmaya dayalı bir öğrenim yöntemi uygulamışlardır [9]. Soetedjo A. yaptığı çalışmada gözün tespiti için 3 adımdan oluşan bir yöntem kullanmıştır. Viola-Jones algoritmasıyla elde edilen yüz bölgesine ilk olarak dikey gradyan uygulanarak kaba bir göz bölgesi elde edilir, ikinci adımda gözdeki beyaz sklerayı çıkarmak için beyaz renk eşiği uygulanır ve son olarak gözün elips şekline uyacak bir elips oturtma yöntemiyle göz bölgesi tespit edilmiştir. Bu çalışmada gözün sınırı %91 oranında başarı ile tespit edilmiştir [10]. Majumder G. vd. yaptıkları çalışmada gri tonlamalı görüntüler üzerinde kontrast artırma uygulayarak aydınlatma etkisinin ortadan kaldırılmasıyla elde ettikleri görüntüleri Hızlı Köşe Dedektörü adı verilen bir özellik algılama tekniği uygulamışlardır. Önerilen yöntemin doğruluğu, rastgele seçilen görüntüler üzerindeki manuel olarak belirlenmiş göz konumları ile önerilen yöntemin tespit ettiği göz konumları karşılaştırılarak ortaya konmuştur [11]. Knapik M. vd. sürücü yorgunluğunun izlenmesi amacıyla aydınlatma koşullarından kaynaklı zorlukların üstesinden gelmek için göz tespitinde termal görüntüler kullanmışlardır. Termal görüntülerin kontrastını arttırmak için yeni sanal yüksek dinamik aralık prosedürü ile bir ön işleme önermiştir. Önerilen yöntem hızı ve doğruluğu YOLOv3 derin sinir ağı tabanlı nesne dedektörü ile karşılaştırılarak ortaya konmuştur [12]. Gou C. vd. yaptıkları çalışmada önce göz algılamayı daha sonra algılanan gözün durumunu tahmin eden kademeli bir regresyon çerçevesi sunmuşlardır. Eğitiminde gerçek ve sentetik görüntülerin kullanıldığı kademeli regresyon çerçevesi hem göz algılama hem de göz durumu (açık-kapalı) tahmini için son teknoloji metotlara kıyasla üstün performans göstermiştir [13]. Yu M. vd. gri yoğunluk Varyans Filtresine (VF) ve Destek Vektör Makinelerine (DVM) dayalı bir hibrit göz algılama yöntemi önermişlerdir. VF ile göz olmayan bölgeler kaldırılırken, eğitilmiş DVM aracılığıyla da iki göz bölgesi kolayca belirlenmiştir [14]. Ghazali K. H. vd. yaptıkları çalışmada yeni bir göz algılama ve izleme algoritması geliştirmişlerdir. Göz algılaması için arka plan ve ten rengi gibi kısıtlamalar olmaksızın gözün yerini etkili ve doğru bir şekilde tespit edebilen eğitilmiş göz filtresi kullanmışlardır. Göz tespit performansını Gabor filtresi ve OpenCV Cascade sınıflandırıcı ile karşılaştırmıştır. 11,5 ms'lik çalışma zamanı performansıyla daha iyi sonuç elde etmiştir [15]. Turan B. vd. yaptıkları çalışmada dizüstü bilgisayar kullanıcılarının göz bölgesi özelliklerini gerçek dünya koşulları altında ortaya çıkarmaya çalışmışlardır. Gözlerin tespiti için bilgisayarın dahili kamerası kullanılarak 107 kullanıcı görüntüsü kullanılarak bir Yapay Sinir Ağı (YSA) eğitilmiştir. Eğitilen YSA ile kullanıcıların göz bölgeleri yüksek bir doğrulukla tespit edilmiştir [16].

Yapılan bu çalışmada yüz görüntülerinden gözlerin tespit edilebilmesi için Evrişimsel Sinir Ağı (ESA) tabanlı bir yöntem önerilmektedir. Önerilen yöntemin ESA modelinde eğitim ve test aşamasında toplamda 260 bin göz ve göz olmayan veri ile çalışılmıştır. Bu veriler "Large-scale CelebFaces Attributes (CelebA)" ve "Closed Eyes In The Wild (CEW)" veri setlerindeki yüz resimlerinden elde edilmiştir [17,18]. Bu veri setleri yeterli miktarda sağa ve sola dönük yüz, düşük çözünürlüklü yüz, farklı aydınlatmalar altındaki yüz ve gözlüklü yüzler içerdiğinden tercih edilmiştir. Verilerin %80'i eğitim aşamasında geri kalanı ise test aşamasında kullanılmıştır. Önerilen yöntem test aşamasında %97,97 gibi bir başarıya sahiptir. Önerilen yöntem Viola-Jones algoritmasının XML tabanlı yüz ve göz tespit uygulaması ile karşılaştırılmıştır. Karşılaştırma sonunda önerilen yöntem, göz tespiti süresi açısından Viola-Jones algoritmasının gerisinde kalırken göz tespitinin doğruluğu açısından yaklaşık %13 gibi daha iyi başarıyı sergilemiştir.

2. Evrişimsel Sinir Ağları (ESA)

Derin sinir ağları veya derin öğrenme yapay sinir ağlarında katmanların derinleşmesiyle birlikte ortaya çıkmıştır. Örnekte tanımda etkili bir makine öğrenimi olan derin öğrenmenin Evrişimsel Sinir ağları, Derin Otomatik Kodlayıcılar, Uzun Kısa Süreli Bellek Ağları, Derin Boltzmann Makineleri ve Tekrarlayan Sinir Ağları gibi farklı algoritmaları mevcuttur [19-23]. Bu algortimalardan biri olan ve günümüzde de nesne tanıma başta olmak üzere birçok farklı alanda etkinliğini ispatlayan ESA, ilk olarak hayvanların görsel korteksinden esinlenilerek ortaya çıkmıştır [24]. İlk ESA mimarisi 1988 yılında LeCun tarafından geliştirilen LeNet ağıdır [19]. 1998 yılında LeCun ve arkadaşları tarafından geliştirilen LeNet-5 mimarisi ESA'nın gelişiminde bir milat olmuştur. Bu mimari ile MINST veri tabanındaki el yazması rakamlar yüksek bir doğrulukta tahmin edilmiştir [25]. LeNet-5 ağı Şekil 1'de gösterilmiştir. LeNet ile başlayan ESA tabanlı mimarilerin gelişimi 2012 yılında ImageNet ILSVRC-2012 yarışmasında AlexNet, ILSVRC-2014 yarışmasında GoogLeNet ve VGG(16) mimarileriyle büyük ivme kazanmıştır. ILSVRC-2015 ImageNet yarışmasında Microsoft ResNet mimarisi görüntü sınıflandırmada %3,6'lık bir hata oranıyla insan doğruluk hatasını geride bırakmıştır [26]. ESA yapısı geleneksel yapay sinir ağlarından farklı olarak evrişimsel katman ve havuzlama katmanına sahiptir. ESA

yapısında genel olarak ilk evrişim katmanlarında giriş görüntüsünden konumsal bilgiler çıkarılırken sonraki evrişim katmanlarında daha çok anlamsal bilgiler çıkarılmaktadır [27].



Şekil 1. LeNet-5 ESA tabanlı mimari [25].

3. Materyal ve Metod

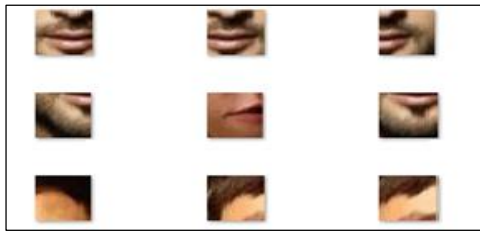
3.1. Veri seti

Önerilen yöntemde tasarlanan ESA modelinin eğitimi için veri setinin büyük bir kısmı 200K dan fazla ünlü resmi içeren “Large-scale CelebFaces Attributes (CelebA)” veri setinden karşılandı [17]. Veri setindeki yüz bin adet resimden göz bölgeleri dlib yüz işaretçisi kullanılarak alınmıştır. Dlib kütüphanesi, Kazemi ve Sullivan'ın 2014 yılındaki “Regresyon Ağaçları Topluluğu ile Bir Milisaniyelik Yüz Hizalama” makalesinin bir uygulaması olan ve iBUG 300-W veri kümesi ile eğitilen yüz ve yüz işaret detektörüdür [9,28,29]. Elde edilen yaklaşık iki yüz bine yakın 30x30 piksellik sağ ve sol göz görüntülerinden işlemci kaynakları göz önünde bulundurularak 150 bin adet kullanılmıştır. Bu görüntülerin içinde gözlüklü, gözlüksüz, farklı duruş pozisyonu, farklı parlaklık ve çözünürlüğe sahip göz görüntüleri mevcuttur. Yapılan kırpma işlemine ait örnek görüntüler Şekil 2’de verilmiştir.

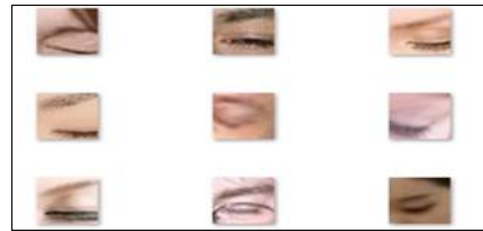


Şekil 2. Kırpılmış göz görüntüleri.

Yine aynı veri setinden yüzde göz olmayan bölgeler 30x30 piksellik 100 bin adet görüntü toplanmıştır. Şekil 3’te göz olmayan kırpılmış görüntüler verilmiştir. Bunun yanında “Closed Eyes In The Wild (CEW)” veri setinde yer alan 1192 resimden 30x30 piksellik sağ ve sol olmak üzere yaklaşık 1600 kapalı göz görüntüsü elde edilmiştir [18]. Bu veri setindeki bazı yüzler dlib yüz dedektörü tarafından tespit edilemediğinden bu yüzlere ait kapalı gözler eğitime dahil edilememiştir. Şekil 4’te kapalı gözler için görsel verilmiştir.

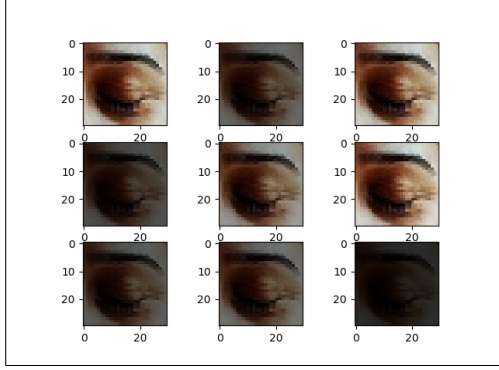


Şekil 3. Göz olmayan görüntüler.

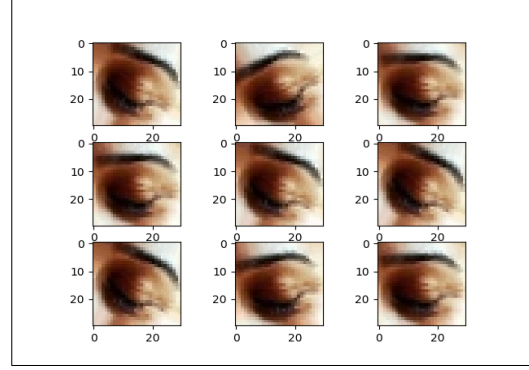


Şekil 4. Kapalı göz görüntüleri.

Kapalı göz görüntülerinin eğitim aşamasında model tarafından daha iyi öğrenilmesi için yaklaşık 1600 kapalı göz görüntüsü üzerine rastgele parlaklık ve rotasyon ile veri artırma yöntemi uygulanarak kapalı göz görüntüsü sayısı 10 bine çıkarılmıştır. Veri artırımına ait görsel Şekil 5 ve Şekil 6’de verilmiştir. Kapalı gözlerde dahil olmak üzere toplamda 260 bin adet eğitim verisinde, açık-kapalı göz görüntüleri için “1”, göz olmayan görüntüler için “0” etiketi kullanılarak ESA modeli ile eğitilmiştir.



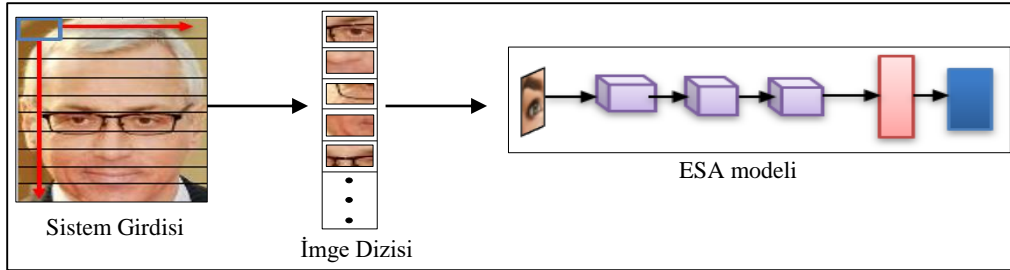
Şekil 5. Rastgele parlaklık.



Şekil 6. Rastgele rotasyon.

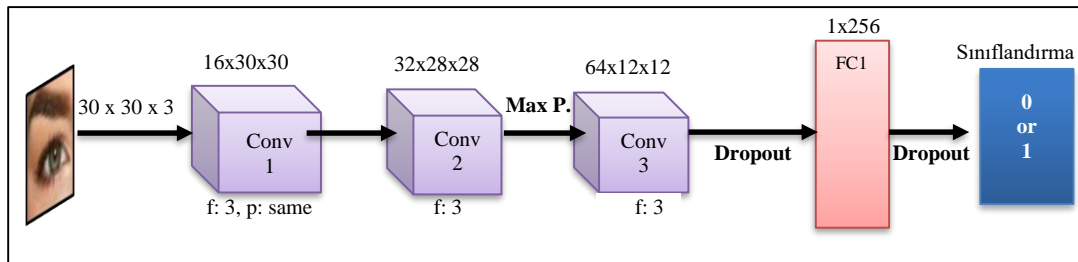
3.2. Önerilen ESA tabanlı sistem akış şeması

Önerilen yöntemde göz tespitinin yapılabilmesi için girdi olarak verilen yüz resmi üzerinde 30x30’luk uzamsal boyuttaki kare imge seçici belirli bir oranda soldan sağa ve yukarıdan aşağıya kaydırılarak resimden örnekler toplamaktadır. Bir dizi değişkeniyle toplanan 30x30’luk imgeler sınıflandırılmak üzere eğitilen ESA’ya verilir. ESA bu imgeleri “göz” veya “göz değil” olarak sınıflandırmaktadır. Önerilen yöntemle ait görsel Şekil 7’de verilmiştir.



Şekil 7. Önerilen ESA tabanlı yöntem.

Önerilen yöntemle ait ESA modelinde 30x30x3 piksel ve kanala sahip göz görüntüsüne 2 evrişim katmanından geçtikten sonra max pooling (maksimum havuzlama) uygulanmıştır. Daha sonra tekrar evrişim katmanına girdikten sonra dropout (atma) uygulanmıştır. Sonra düzleştirilip tam bağlantı katmanına iletilmiştir. 256 nöronluk tam bağlantı katmanından sonra tekrar dropout uygulanarak sınıflandırma katmanına iletilmiştir. Son katmanda ikili sınıflandırma için softmax, diğer tüm katmanlarda ReLU aktivasyon fonksiyonu kullanılmıştır. Şekil 8’de tasarlanan ESA modeli verilmiştir.

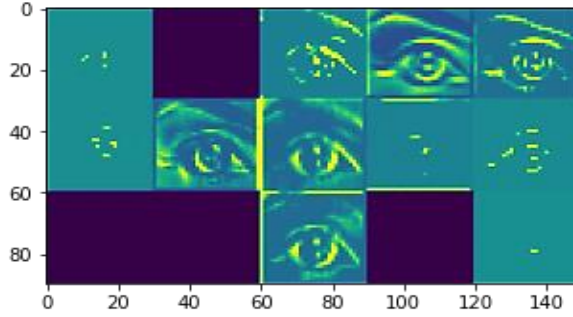


Şekil 8. Tasarlanan ESA modeli.

Şekil 8'de verilen Conv1, Conv2, Conv3 evrişim katmanları 30x30x3 uzamsal boyuta sahip matris formundaki giriş görüntüsüne sırasıyla uygulanmaktadır. 3x3x3 uzamsal boyutta ve 0, 1, -1'den oluşan filtreler giriş matrisi üzerinden adım adım kaydırılarak evrişim işlemi gerçekleştirilmektedir. Evrişim işlemi sonunda filtrenin uzamsal boyutuna bağlı olarak çıktının uzamsal boyutu değişmektedir. Evrişim işleminden sonra elde edilen çıktılar bir sonraki evrişim katmanı için girdi olacaktır. Görüntüye uygulanan evrişim işleminin formülü Denklem 1'de verilmiştir [30].

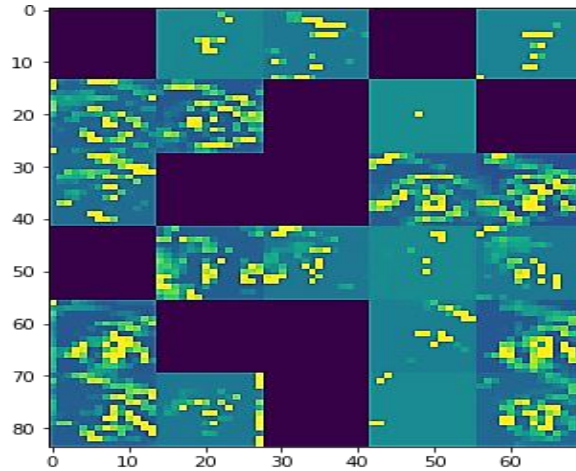
$$F_c = \text{conv}(i, j) = (I \otimes K)(i, j) = \sum_m \sum_n I(m, n)k(i - m, j - n) \quad (1)$$

Denklem 1'de verilen F_c evrişim çıktısını, \otimes işlevi evrişim işlemi, m ve n ise görüntü uzamsal boyutunu ifade etmektedir. I elemanı girdi görüntüsünü K elemanı ise filtre (çekirdek)'yi ifade etmektedir. K filtresi I görüntüsü üzerinde soldan sağa ve yukarıdan aşağıya doğru kaydırılarak girdi matrisinin her bir elemanı çekirdeğin her bir elemanı ile çarpılır. RGB görüntü matrislerinde ise evrişim işlemi her bir katman için yapılır ve evrişim çıktısı oluşturmak için birleştirilir [30]. Tasarlanan ESA modelinin ilk evrişim katmanı çıktısı Şekil 9'da verilmiştir.



Şekil 9. Conv1 katmanı evrişim çıktıları.

Ağda uygulanan hiper parametrelerden Max pooling, giriş verilerine (2x2)'lik bir filtre uygulanmasıyla verideki maksimum değerler dikkate alınarak uzamsal bir alt boyut elde edilir. Böylelikle hesaplama kolaylığı ve önemli özelliklerin çıkarılmasında rol oynar [31]. Tasarlanan modelin Max pooling layer (Maksimum havuzlama katmanı) çıktıları Şekil 10'da veri görselleştirme yöntemiyle verilmiştir. Bir diğer parametre olan Dropout, katmanlarda yer alan özellik dedektörlerinin (nöron) belirlenen oranda rastgele atılmasıyla gerçekleşir. Sinir ağının eğitimi sırasında, ağın veri setini aşırı öğrenme (overfitting)'sinin önüne geçmek ve yeni veriler için ağın genelleme gücünü arttırmak için kullanılır [32,33].



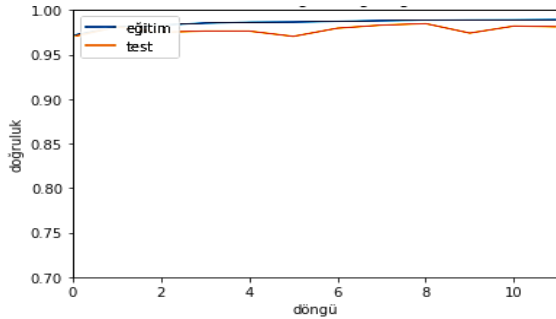
Şekil 10. Maksimum havuzlama katmanı çıktısı.

3.3. Sistem parametreleri

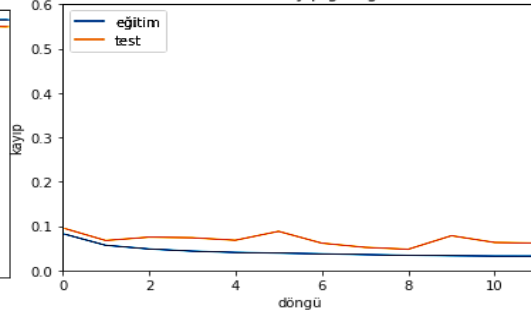
Açık-kapalı göz, yüzde göz olmayan imgeler ile toplamda 260 bin veri üzerinde eğitim ve test işlemleri gerçekleştirilmiştir. Eğitim aşaması %99 ve test aşaması %97,97 doğruluk oranıyla sonuçlanmıştır. Bu sonuçlar ağın doğru bir şekilde eğitildiğini göstermektedir. Tasarlanan ESA modeli verileri Tablo 1’de verilmiştir. Tablodan görüleceği üzere eğitim aşamasında 260K verinin %80’i eğitimde geriye kalan veride test aşamasında kullanılmıştır. Eğitim aşaması 15 döngüde gerçekleştirilmiştir. Her döngüde veri seti 32 imgelik toplu parçalar halinde eğitilmiştir. Kayıp fonksiyonu olarak categorical crossentropy ve ağın optimizasyonu için ise “Adam” optimizasyon algoritması kullanılmıştır. Şekil 11 ve 12’de sırasıyla model doğruluk ve model kayıp grafiği verilmiştir.

Tablo 1. Tasarlanan ESA model verileri.

Veri seti (Data set)	260.000
Eğitim verisi (Training data)	%80
Test verisi (Test data)	%20
Giriş verileri (Input data)	Açık + Kapalı göz imgeleri (150.000 + 10.000) Yüzde göz olmayan imgeler (100.000)
Kayıp fonksiyonu (Loss fonk.)	Categorical crossentropy
Optimizasyon (optimization)	Adam
Parti boyutu (Batch size)	32
Eğitim döngü sayısı (Epoch)	15
Doğruluk (Eğitim)% (Training accuracy)	99.00
Doğruluk (Test)% (Test accuracy)	97.97



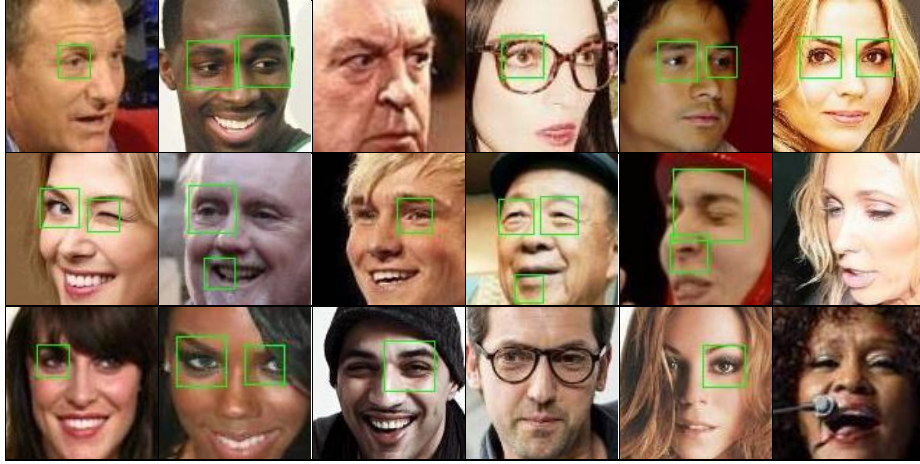
Şekil 11. Model doğruluk grafiği.



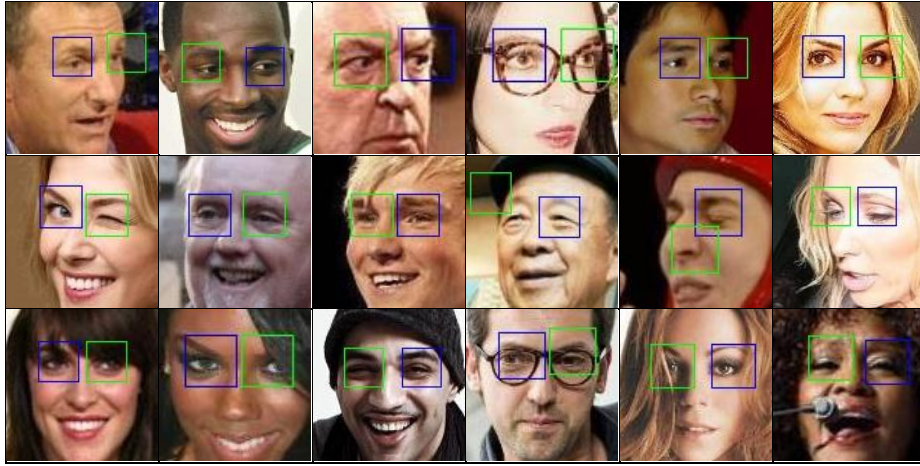
Şekil 12. Model kayıp grafiği.

4. Sonuçlar ve Tartışmalar

Önerilen yöntemde girdi olarak verilen resimden toplanan imgeler ESA modeline verildikten sonra model tarafından yapılan sınıflandırma sonucunda imgeler softmax aktivasyon fonksiyonu sayesinde 0 ile 1 arasında tahmin değerlerine sahip olmaktadır. 1’e yakın değer alan imgeler “göz”, 0’a yakın değer alan imgeler ise “göz değil” olarak kabul edilmiştir. Göz olarak kabul edilen imgeler ikiden fazla olabilmektedir. Çünkü yüz üzerinde imge toplama işlemi gerçekleştirilirken kayma miktarına bağlı olarak göz bölgesinden birden fazla örnek alınmaktadır. Dolayısıyla en yüksek tahmin skoruna sahip imgeler göz olarak kabul edilmektedir. Şekil 13 ve Şekil 14’te sırasıyla Viola Jones algoritması ve önerilen ESA tabanlı sistemin CelebA veri setinde yer alan ve eğitimde kullanılmayan resimlere uygulanması sonucu elde edilen sonuçların bir kısmı verilmiştir. Resimlerin seçiminde sağa, sola ve öne bakan resimlerin kullanılmasına dikkat edilmiştir. Viola Jones algoritmasının uygulanmasında kullanılan opencv kütüphanesinin detectMultiScale fonksiyonundaki göz tespit parametrelerinden scaleFactor=1,2 ve minNeighbors=3 alınmıştır [34].



Şekil 13. Viola-Jones algoritması uygulanan resimler.



Şekil 14. Önerilen yöntemin uygulandığı resimler.

Önerilen yöntemin Viola-Jones algoritması ile karşılaştırılması için eğitim aşamasında kullanılmayan 600 resim seçilmiştir. Seçilen resimlerde farklı duruş pozisyonları ve gözlüklü resimlerin de mevcut olmasına dikkat edilmiştir. İki yöntemin kıyaslamasında 600 resim üzerinde tespit edilen sonuçlar Tablo 2’de gösterilmiştir. Tablo 2’de, önerilen yöntemin 600 resimde de en az bir göz tespiti yaptığı görülmüştür. Önerilen yöntem iki göz tespitinde %98,2’lik bir başarı sağlamıştır. Viola-Jones algoritmasında bu oran %84,5’tir.

Tablo 2. 600 resim üzerinde karşılaştırma.

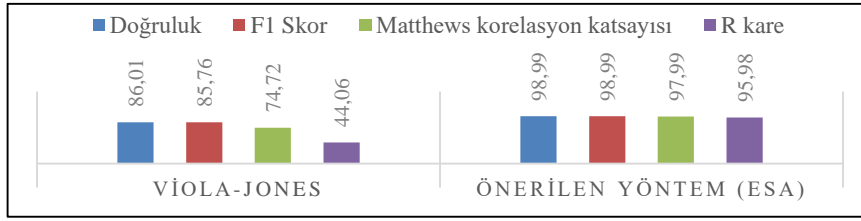
Viola-Jones ve Önerilen yöntem				
Algoritmalar	Tek Göz	İki Göz	Hiçbiri	Toplam resim
Viola-Jones	61	507	32	600
Önerilen yöntem	11	589	0	600

Önerilen yöntem doğruluk testlerinden de başarıyla geçmiştir. Doğruluk testlerinin yapılabilmesi için eğitimde kullanılmamış 10 bin göz görüntüsü ve 10 bin yüzde göz olmayan görüntü olmak üzere toplamda 20 bin görüntü Doğruluk, F1 skor, Matthews korelasyon ve R kare ölçütlerine tabi tutulmuştur. Bu ölçütlerin

uygulanmasında python sklearn kütüphanesi fonksiyonları kullanılmıştır [35]. Sonuçlar Tablo 3'te verilmiştir. Ayrıca Tablo 3 sonuçları grafiksel olarak ta Şekil 15'te gösterilmiştir.

Tablo 3. Doğruluk ölçütleri.

Algoritmalar	Doğruluk (Accuracy) %	Kesinlik (Precision) %	Çağrışım (Recall) %	F1 Skor (F1 Score) %	Matthews korelasyon katsayısı (MCC) %	R kare (R squared) %
Viola-Jones Algoritması	86.01	99.00	72.70	85.76	74.72	44.06
Önerilen yöntem	98.99	98.69	99.30	98.99	97.99	95.98



Şekil 15. Viola-Jones ve Önerilen yöntem doğruluk karşılaştırması.

Önerilen yöntem sınırlandırılmış yüz görüntüleri üzerinde 30x30 boyutlarında göz araması yapmaktadır. Dolayısıyla bu durum Viola-Jones algoritmasına göre dezavantajlı görünmektedir. Çünkü Viola-Jones algoritması farklı boyutlardaki göz görüntülerini de bulabilmektedir. Önerilen yöntemin farklı boyutlardaki yüz görüntülerinde de göz tespiti yapılabilmesi için girdi olarak verilecek yüz görüntülerinin uzamsal boyutlarının görüntü işleme teknikleri ile azaltılması veya artırılması ile yine 30x30'luk bir göz arama tespiti gerçekleştirilerek bu sorunun üstesinden gelinbilir.

5. Öneriler

Yapılan bu çalışmada ESA'nın örüntü tanımadaki gücünden faydalanılarak bir göz tespit yöntemi önerilmiştir. Önerilen yöntemin göz tespitindeki başarısı Viola-Jones algoritması ile karşılaştırılarak yapılmıştır. Önerilen yöntem Viola-Jones algoritması ile karşılaştırıldığında önerilen yöntemin daha iyi sonuçlar verdiği görülmektedir. Fakat gözlerin algılanmasında çalışma zamanı performansı Viola-Jones algoritması ile karşılaştırıldığında önerilen yöntemin göz algılama süresi, resim üzerinde imgelerin toplanması sırasında kaybedilen süreden dolayı daha uzundur. Resim üzerindeki imgeler toplanırken belirlenen kayma miktarının artırılması ile bu süre düşürülebilmektedir. Sürenin düşürülmesine karşılık doğrulukta bir miktarda düşmektedir. Geliştirilecek bir algortima ile yüzün spesifik özelliklerinden faydalanarak yüzün tamamı üzerinde yapılacak imge toplama işlemi yerine gözlerin bulunduğu lokasyonda imge toplama işlemi yapılarak önerilen yöntemin hızı büyük ölçüde artırılabilir. Ayrıca önerilen yöntemin daha hızlı işlemcilerde kullanılmasıyla gerçek zamanlı uygulamalarda kullanılması uygundur.

Teşekkür

Bu çalışma İnönü Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimi (BAP) tarafından FDK-2020-2110 kodlu proje ile desteklenmiştir.

Kaynaklar

- [1] Noble AM, Miles M, Perez MA, Guo F, Klauer SG. Evaluating driver eye glance behavior and secondary task engagement while using driving automation systems. Accident Analysis and Prevention 2021; 151.
- [2] Rakhmatulin I, Duchowski AT. Deep neural networks for low-cost eye tracking. Procedia Computer Science 2020; 176: 685–694.

- [3] Alghamdi J, Alharthi R, Alghamdi R, Alsubaie W, Alsubaie R, Alqahtani D, Ramadan RA, Alqarni L, Alshammari R. A Survey on Face Recognition Algorithms. ICCAIS 2020 - 3rd International Conference on Computer Applications and Information Security 2020; 1–5.
- [4] Tarnowski P, Kołodziej M, Majkowski A, Rak RJ. Eye-Tracking Analysis for Emotion Recognition. Computational Intelligence and Neuroscience 2020; 1–13.
- [5] Thiyaneswaran B, Padma S. Iris Recognition using Left and Right Iris Feature of the Human Eye for Biometric Security System. International Journal of Computer Applications 2012; 50(12): 37–41.
- [6] Turan B, Eskikurt HI, Can MS. An Application Based on Artificial Neural Network for Determining Viewpoint Coordinates on a Screen. Elektronika ir Elektrotechnika 2016; 22(2): 86–91.
- [7] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2001; 1: 511–518.
- [8] Zhang K, Zhang Z, Li Z, Qiao Y. Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks. IEEE Signal Processing Letters 2016; 23(10): 1499–1503.
- [9] Kazemi V, Sullivan J. One millisecond face alignment with an ensemble of regression trees. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2014.
- [10] Soetedjo A. Eye Detection Based-on Color and Shape Features. International Journal of Advanced Computer Science and Applications 2012; 3(5).
- [11] Majumder G, Bhowmik MK, Bhattacharjee D. Automatic Eye Detection Using Fast Corner Detector of North East Indian (NEI) Face Images. Procedia Technology 2013; 10:646–653.
- [12] Knapik M, Cyganek B. Fast eyes detection in thermal images. Multimedia Tools and Applications 2020; 80(3):3601–3621.
- [13] Gou C, Wu Y, Wang K, Wang K, Wang FY, Ji Q. A joint cascaded framework for simultaneous eye detection and eye state estimation. Pattern Recognition 2017; 67:23–31.
- [14] Yu M, Lin Y, Wang X. An efficient hybrid eye detection method. Turkish Journal of Electrical Engineering and Computer Sciences 2016; 24(3):1586–1603.
- [15] Ghazali KH, Jadin MS, Jie M, Xiao R. Novel automatic eye detection and tracking algorithm. Optics and Lasers in Engineering 2015; 67: 49–56.
- [16] Turan B, Eskikurt HI. Iris and eye corner detection by processing internal webcam images. Turkish Journal of Electrical Engineering and Computer Sciences 2016; 24(4): 3036–3048.
- [17] Liu Z, Luo P, Wang X, Tang X. Deep learning face attributes in the wild. In ICCV 2015; 3730–3738.
- [18] Song F, Tan X, Liu X, Chen S. Eyes closeness detection from still images with multi-scale histograms of principal oriented gradients. Pattern Recognition 2014; 47(9): 2825–2838.
- [19] Cun YL, Guyon I, Jackel LD, Henderson D, Boser B, Howard RE, Denker JS, Hubbard W, Graf HP. Handwritten Digit Recognition: Applications of Neural Network Chips and Automatic Learning. IEEE Communications Magazine 1989; 27(11): 41–46.
- [20] Vincent P, Larochelle H, Bengio Y, Manzagol PA. Extracting and composing robust features with denoising autoencoders. Proceedings of the 25th International Conference on Machine Learning 2008; 1096–1103.
- [21] Hochreiter S, Schmidhuber J. Long Short-Term Memory. Neural Computation 1997; 9(8): 1735–1780.
- [22] Salakhutdinov R, Hinton G. Deep Boltzmann Machines. 2009;5(2).
- [23] Martens J, Ca IU. Learning Recurrent Neural Networks with Hessian-Free Optimization Ilya Sutskever. Proceedings of the 28th International Conference on Machine Learning (ICML) 2011.
- [24] Hubel DH, Wiesel TN. Receptive fields and functional architecture of monkey striate cortex. The Journal of Physiology 1968; 195(1): 215–243.
- [25] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proceedings of the IEEE 1998; 86(11): 2278–2323.
- [26] Ajit A, Acharya K, Samanta A. A Review of Convolutional Neural Networks. International Conference on Emerging Trends in Information Technology and Engineering (Ic-ETITE) 2020; 1-5.
- [27] Li W, Liu K, Yan L, Cheng F, Lv YQ, Zhang LZ. FRD-CNN: Object detection based on small-scale convolutional neural networks and feature reuse. Scientific Reports 2019; 9(1).
- [28] dlib C++ Library. <http://dlib.net/>, (23.09.2020).
- [29] i'bug-resources-Facial point annotations. <https://ibug.doc.ic.ac.uk/resources/facial-point-annotations/>, (23.09.2020).
- [30] Sahu M, Dash R. A survey on deep learning: Convolution neural network (cnn). Smart Innovation, Systems and Technologies 2021; 153: 317–325.
- [31] Rachapudi V, Lavanya Devi G. Improved convolutional neural network based histopathological image classification. Evolutionary Intelligence 2020; 1–7.
- [32] Hinton GE, Srivastava N, Krizhevsky A, Sutskever I, Salakhutdinov RR. Improving neural networks by preventing co-adaptation of feature detectors, 2012.
- [33] Hahn S, Choi H. Understanding dropout as an optimization trick. Neurocomputing 2020; 398: 64–70.
- [34] Home-OpenCV. <https://opencv.org/>, (30.12.2020).
- [35] Pedregosa Fabianpedregosa F, Michel V, Grisel Oliviergrisel O, Blondel M, Prettenhofer P, Weiss R, Vanderplas J, Cournapeau D, ve diğerleri. Scikit-learn: Machine learning in Python. In Journal of Machine Learning Research 2011; 12(85): 2825-2830.