



Düzce Üniversitesi Bilim ve Teknoloji Dergisi

Araştırma Makalesi

Yüz Tanıma Sistemleri İçin Geliştirilmiş Veri Artırma Temelli Adaptif Yüz Tanıma Modeli

 Mustafa TAN ^{a,*},  Cem EMEKSİZ ^b

^a Elektrik-Elektronik Mühendisliği Bölümü, Lisansüstü Eğitim Enstitüsü, Tokat Gaziosmanpaşa Üniversitesi, Tokat, TÜRKİYE

^b Elektrik-Elektronik Mühendisliği Bölümü, Mühendislik ve Mimarlık Fakültesi, Tokat Gaziosmanpaşa Üniversitesi, Tokat, TÜRKİYE

* Sorumlu yazarın e-posta adresi: mustan79@gmail.com

DOI:10.29130/dubited.1024670

ÖZ

Hızla gelişen bilgisayar ve grafik ara yüzüne sahip cihaz teknolojileri, yüz tanıma çalışmalarında yeni ufuklar açmışlardır. Özellikle derin öğrenme ağ mimari yapılarından biri olan evrimsel sinir ağları (Convolutional Neural Network-CNN), yüz tanıma çalışmalarında büyük başarılar sağlamaktadır. Bu başarılar da veri setlerinin büyüklüğü önemli rol oynamaktadır. Özellikle kullanılan veri setlerindeki yetersizlik başarı oranlarını etkileyebilmektedir. Bunun önüne geçmek için ise veri tipine göre değişik veri artırma teknikleri uygulanmaktadır. Yapılan bu çalışmada yüz tanımlama problemi için derin öğrenmeye dayalı adaptif bir yüz tanıma modeli (AYTM) geliştirildi. Geliştirilen bu model kontrast sınırlı uyarlanabilir histogram eşitleme (Contrast Limited Adaptive Histogram Equalization-CLAHE), CNN ve çok katmanlı algılayıcı (Multi Layer Perceptron-MLP)'ndan oluşmaktadır. İki farklı veri seti grubu kullanılarak geliştirilen modelin performans değerlendirilmesi yapılmıştır. Özellikle veri artırma işleminin model başarısını ciddi oranda artırdığı gözlemlendi ve veri artırma işleminin derin öğrenme uygulamalarında gerekliliği vurgulanmıştır.

Anahtar Kelimeler: Evrimsel sinir ağları, Veri artırma, Histogram eşitleme

Data Augmentation Based Adaptive Face Recognition Model Developed for Face Recognition Systems

ABSTRACT

The rapidly developing computer and device technologies with graphical interfaces opened new horizons in face recognition studies. Especially Convolutional Neural Networks (CNN), which is one of the deep learning network architecture structures, provides great success in face recognition studies. The size of the datasets plays an important role in these achievements. Especially the inadequacy of the data sets used can affect the success rates. In order to prevent this, different data augmentation techniques are applied according to the data type. In this study, an adaptive face recognition model based on deep learning was developed for the face identification problem. This developed model consists of contrast limited adaptive histogram equalization (CLAHE), CNN and multi-layer perceptron (MLP). Performance evaluation of the model developed was made by using two different data set groups. In particular, it was observed that data augmentation significantly increased the success of the model, and the necessity of data augmentation in deep learning applications was emphasized.

Keywords: Convolutional neural network, Data augmentation, Histogram equalization

I. GİRİŞ

İnsan yüzü, bedenin bir parçası olmakla birlikte, bir kişinin kimliğinin tanınması için en temel birincil bilgileri içermektedir. Özellikle değişik lokasyonlarda insanların kimliğini doğru bir şekilde tanımlamak günümüzde önemli bir araştırma konusu haline gelmiştir [1]. Bu sebeple, yüz tanıma teknolojileri alanında çalışmalar yoğun bir şekilde devam etmektedir. Yüz tanıma teknolojileri temel olarak yüzün algılanması, anahat özelliklerinin çıkartılması ve hizalanması gibi uygulamaları barındırır [2, 3]. Ayrıca yüz doğrulama, yüz izleme ve yüz ifadesi analizleri yüzle ilgili uygulamalara yönelik diğer önemli araştırma konularıdır. Birçok çalışmada gerçek yaşam koşullarında elde edilen yüzlere ait çeşitli özelliklerin, çevresel etkilerden kaynaklı baskılanması araştırmalarda problem oluşturabilmektedir. Dolayısıyla bilgisayarlarla yapılan görüntü analizlerinin ilk dönemlerinde manuel olarak oluşturulan özelliklerden güçlü sınıflandırıcılar elde edilmeye odaklanılmıştır [4].

Genel olarak yüz algılama, genel nesne algılamanın özel bir durumu olarak görülmektedir [5]. Yüz hizalama ise yüz analiz sistemleri için önemli bir ön işleme adımıdır [6]. Yüzün özellik noktalarının konumlandırılması, konumlandırma doğruluğu ve konumlandırma hızı yüz tanıma sistemini doğrudan etkiler. Gerçekte yüz hizalaması farklı pozlar, ifadeler, aydınlatma ve yüz görüntülerindeki kısıtlamalar gibi faktörler nedeniyle son derece zorlaşmaktadır [7]. Amour ve arkadaşları yüz özelliklerini çıkarmada Gabor filtresi ve çift kodlu ikili desene dayalı bir hibrit özellik çıkarımı önermişler ve mükemmel bir performans sergilediğini bildirmişlerdir [8]. Cuimei ve arkadaşları üç sınıflandırıcı ve Haar Cascade algoritmasıyla beraber ten rengi histogram eşleştirme, göz algılama ve ağız algılamaya dayalı ek sınıflandırıcılar kullanarak yeni bir insan yüzü algılama algoritması önermişlerdir [9]. Shu ve arkadaşları yüksek doğruluk gösteren yeni bir yüz algılama yöntemi önermişlerdir [10]. Aynı kişiye ait iki görüntünün zamana, poza, yüz ifadesine ve aydınlatma koşullarına göre önemli ölçüde değişebildiğinden yüz tanıma birçok zorluğa sahiptir. Özellikle ortam aydınlatması, nesne algılama ve yüz tanıma üzerinde etkilidir. Lv ve arkadaşları farklı aydınlatma ortamlarını simüle etmek için 3B yüz modeli yöntemini kullanarak, yüz tanıma modelini farklı aydınlatmalara karşı dayanıklı hale getirmişlerdir [11]. Faraji ve Qi logaritmik fraktal boyuta dayalı yöntem kullanarak yüz görüntüleri için aydınlatmanın etkisini azaltmayı önermişlerdir [12]. Azimifar ve arkadaşları aydınlatma gibi kontrolsüz koşulların etkilerini azaltmak için ön işlemede yoğunluk ve normalleştirme yerine özellik çıkarma kullanan çok yönlü bir yaklaşım sunmuşlardır [13].

Nesne tanımda hızlı ve otomatik tanıma için makine öğrenmesi ve optimizasyon tabanlı geleneksel yöntemler sıklıkla tercih edilmektedir. Fakat bu yöntemlerin en büyük dezavantajı manuel olmaları ve bu nedenle öznetelik çıkarma ve öznetelik seçimi adımlarında bilgi kaybı yaşanmasıdır [14,15]. Geleneksel makine öğrenimi algoritmaları yerine ise veriye dayalı derin öğrenme algoritmaları, girdi verilerinden ayırt edici özellikleri etkin ve doğru bir şekilde otomatik olarak öğrenme yeteneğine sahiptir ve son yıllarda daha çok tercih edilmeye başlanmıştır [16]. Derin öğrenmede, özellikler daha derin katmanlar aracılığıyla ham verilerden otomatik olarak çıkarılır ve bu özelliklerin güçlü bir ayırt etme yeteneği vardır. Son zamanlarda derin öğrenme yöntemleri yüz ve nesne tanıma problemlerinde yaygın olarak tercih edilmektedir. Literatür çalışmaları incelendiğinde elde edilen sonuçların yüksek doğruluk ve tutarlılığa sahip olduğu görülmektedir [17]. Derin Sinir Ağı (Deep Neural Network-DNN), Derin İnanç Ağı (Deep Belief Network-DBN), Evrimsel Sinir Ağı (Convolutional Neural Network-CNN) ve Tekrarlayan Sinir Ağı (Recurrent Neural Network-RNN) gibi çok sayıda farklı derin öğrenme ağ mimarileri mevcuttur [18]. Bilgisayar teknolojisinde gelişen grafik işlemci birimleri (GPU) ile vektör ve matris işlemlerinin daha hızlı yapılması ve derin öğrenmenin sahip olduğu çok katmanlı yapı ile çok hızlı öğrenme olanağı ortaya çıkmıştır [19]. Nesne ve görüntü tanıma alanında yapılan çalışmalarda kullanılan derin öğrenme algoritmaları içerisinde CNN en yaygın olanıdır [20]. Evrimsel Sinir Ağı yapısının yaygın olarak kullanılmasının nedeni yapılan sınıflandırma ve görüntü işleme problemlerinde daha doğru ve tutarlı sonuçlar vermesidir [21].

CNN'nin başarısı doğrudan örnek sayısına ve kısıtlanmamış ortam problemlerine bağlıdır. Ancak kısıtlanmamış ortama maruz kalmamış çok sayıda veriye ulaşmak her zaman mümkün olmamaktadır. Bu nedenle araştırmacılar hem kısıtlanmamış ortamın etkilerini azaltmak hem de ağın başarısını

arttırmak için veri setindeki görüntülerin değiştirilmiş versiyonlarını oluşturarak eğitim veri setinin boyutunu yapay olarak genişletirler. Bu amaçla ham görüntülere uygulanan bu tekniklere veri artırma teknikleri denir [15]. Büyük veri seti oluşturmada, verilerin hem sayısını hem de çeşitliliğini artırmak için veri artırma yöntemi etkili bir tekniktir. Görüntüler üzerinde yaygın olarak yapılan veri artırma teknikleri, görüntünün birkaç piksel çevrilmesi, görüntünün yatay olarak çevrilmesi veya renk uzayında değişiklikler yapılması gibi örneklendirilebilir [22]. Sharma ve Naaz Mir, derin öğrenme tabanlı görsel nesne tespitindeki en son gelişmeleri kapsamlı olarak incelemişlerdir. Bunun için literatürdeki son çalışmaları gözden geçirmişler ve derin öğrenme ile görsel nesne tespiti için gelecekteki birkaç yönelimi de sunmuşlardır [23]. Shorten ve Khoshgoftaar, çalışmalarında sınırlı veri sorununun çözümünde eğitim veri setlerinin boyutunu ve kalitesini artıran veri artırmaya odaklanmışlardır. Böylece daha iyi derin öğrenme modelleri oluşturduklarını bildirmişlerdir [24]. Çağlı ve arkadaşları, çalışmalarında CNN'nin performansını önemli ölçüde artırmak için veri artırma tekniği ile donatmayı önermişlerdir. Evrimsel sinir ağı yaklaşımının veri artırmayla birleştirildiğinde etkili bir yöntem sunduğunu bildirmişlerdir [25].

Bu çalışmada doğru ve güvenilir bir şekilde yüz tanıma işlemi gerçekleştirebilmek için adaptif bir yüz tanıma modeli geliştirilmiştir. Geliştirilen bu modelin uygulanmasında birtakım aşamalar gerçekleştirilmektedir. Bu aşamaların ilki veri setinin seçimidir. Metodoloji kısmında ayrıntıları verilen Yalefaces ve Data_org yüz veri setleri tercih edilmiştir. Her iki veri seti de kısıtlanmamış ortam görüntülerine sahiptir. İkinci aşama veri setlerine ilişkin bilgilerin ve veri özelliklerinin çıkarılmasıdır. Bu bölüm ayrıca çalışmanın en fazla odaklandığı bölümdür. Kısıtlanmamış ortam etkileri ve yetersiz veri seti, yapılan yüz tanıma işleminin doğruluğunu büyük ölçüde etkiler ve değiştirir. Bu nedenle çalışmanın bu bölümünde yüz tanıma işlemindeki tutarlılığın ve doğruluğun artırılması için CLAHE'nin kullanıldığı veri ön işleme yaklaşımı sunulmuştur. CLAHE ile veri setleri 4 kat artırılmıştır. Son olarak sınıflandırma aşamasında ise CNN-MLP kombinasyonunda optimizasyon için adaptif gradyan algoritması (Adagrad) kullanılmıştır. Çalışmanın en temel yenilikçi yönleri ve sağladığı katkılar aşağıda detaylı bir şekilde özetlenmektedir:

(1) Bu çalışmada AYTİM isimli etkili bir yüz tanıma modeli geliştirildi, (2) eğitim aşamasında veri artırma tekniği (CLAHE) ile görüntü çözünürlüğü değiştirildi ve veri sayısı 4 kat artırıldı. Böylece veri artırımının yüz tanıma sonuçlarına etkisi incelenmiştir, (3) orijinal görüntüden ve artırılmış özelliklerden öz nitelikleri çıkarmak için bir CNN model ile Adagrad optimizasyon tabanlı gelişmiş bir MLP sınıflandırıcısı kullanıldı, (4) önerilen AYTİM'de veri artırımının doğruluk ve güvenilirliği yükselttiği tespit edilmiştir.

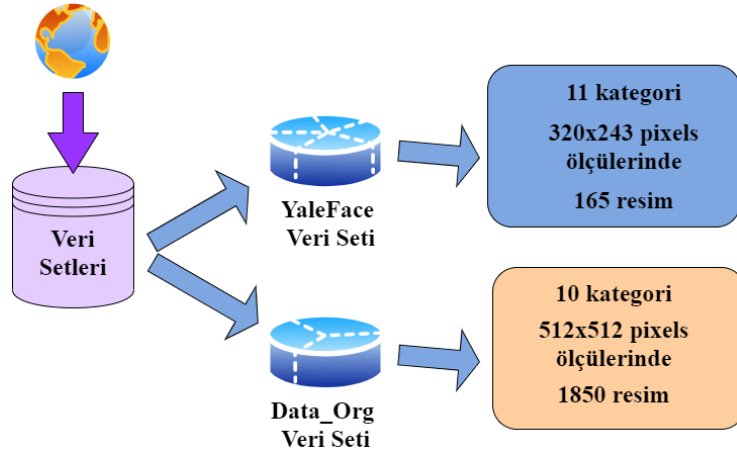
Çalışmanın geri kalanı ise şu şekilde yapılandırılmıştır: Bölüm 2, veri kümelerini açıklar ve önerilen modelin çerçevesi tartışılır. Deneysel sonuçlar Bölüm 3'te sunulurken, Bölüm 4'te çalışmada elde edilen sonuçlar değerlendirilip tartışılmıştır.

II. METODOLOJİ

Çalışmada önerilen yenilikçi adaptif sınıflandırma modeline ait bölümler ve bu bölümlerde kullanılan bazı metod ve yaklaşımlar aşağıda detaylı bir şekilde sunulmaktadır.

A. VERİ SETİ

Yapılan çalışmada kullanılan veri setlerinin genel tanımlaması Şekil 1'de gösterilmektedir. Veri tabanı kullanılarak iki farklı veri seti oluşturulmuş ve önerilen modelin performans karşılaştırılması da bu veri setlerine göre gerçekleştirilmiştir.



Şekil 1. Veri setlerinin genel özellikleri

Birinci veri seti Yalefaces yüz veri setidir. Yalefaces yüz veri seti, çalışmalar için daha uygun bir veri kümesidir çünkü tanıma problemi daha zordur. Veri seti, her biri 320x243 piksel boyutunda 11 gri tonlamalı görsele sahiptir ve 15 kişiden (14 erkek, 1 kadın) 165 GIF görüntüsünü içerir. Değişik ışık koşullarında (merkezi ışık, soldan ışık, sağdan ışık), yüz ifadelerinde (mutlu, normal, üzgün, uykulu, şaşırılmış, göz kırpmış) ve gözlüklü, gözlüksüz gibi farklı durumlardan oluşmaktadır. İkinci veri seti Data-org ise New York sokaklarındaki insanlardan, on ortak ifadenin önerdiği duyguları, gerçekleştirmelerini isteyerek toplanmıştır. Bu veri kümesi insanların gerçek hayatlarında sık kullandıkları yüz ifadelerini içerir. Bu veri seti, 10 kategoride ve 185 farklı kişinin renkli görüntülerini içerir. Veri seti, her biri 512x512 piksel boyutunda ve kişilerin değişik yüz ifadelerinde (mutlu, normal, üzgün, uykulu, şaşırılmış, göz kırpmış vb.) oluşmaktadır. Şekil 2 ve 3'te veri setlerinden örnekler verilmiştir. Yalefaces orijinal veri setinin %80'ni yani 132 görüntü eğitim verisi, geri kalan 33 görüntü ise test veri seti olarak kullanılmıştır. Data_org orijinal veri setinin %80'ni yani 1480 görüntü eğitim verisi, geri kalan 370 görüntü ise test veri seti olarak kullanılmıştır. Histogram eşitleme yöntemiyle 4 kat artırımı yapılan diğer çalışmada yalefaces eğitim veri setinin %80'i olan 528 görüntü ve Data_org veri setinin %80'i olan 5920 görüntü eğitim verisi olarak kullanılmıştır.



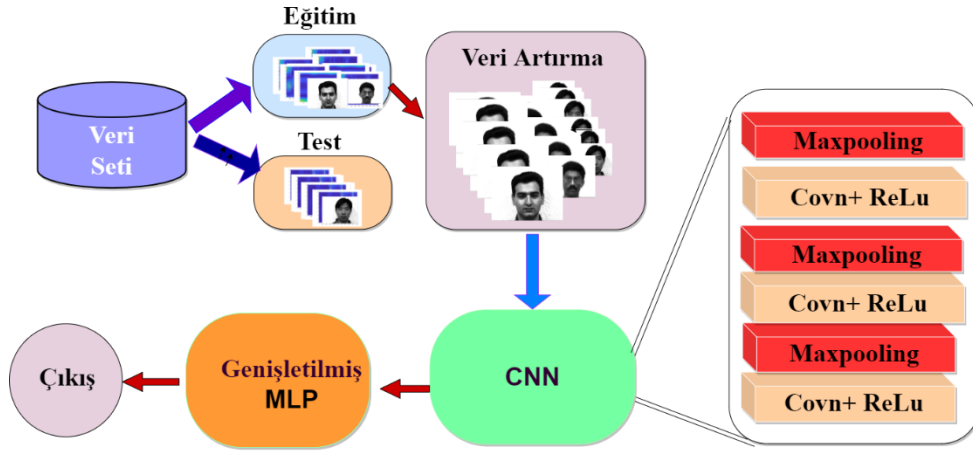
Şekil 2. Yalefaces veri setinden örnekler



Şekil 3. Data_org veri seti örnekleri

B. ADAPTİF YÜZ TANIMA MODELİ

Önerilen AYTМ modelinin genel akış şeması Şekil 4'te gösterilmektedir. Önerilen sınıflandırma modelinde üç modül vardır: (1) CLAHE yöntemi ile veri artırma yaparak görüntü sayısını genişletmek, (2) CNN model ile veriye ait özellikleri çıkarmak (Özellik oluşturma) ve (3) Orijinal ve artırılmış verilerden CNN ile çıkarılan öznelikler geliştirilmiş MLP sınıflandırıcısına (Derin Sınıflandırıcı) eklenerek daha sonra sonuçlar elde etmektir. Modelde yer alan metodlar aşağıda sunulan alt başlıklar altında detaylı bir şekilde açıklanmaktadır.



Şekil 4. Adaptif yüz tanıma modeli

B. 1. Kontrast Sınırlı Adaptif Histogram Eşitleme (CLAHE)

Derin öğrenme uygulamalarının görüntü verileri ile yapılan çalışmalarında en iyi sınıflandırma eğitim verilerinin miktarını yüksek tutarak gerçekleştirilmektedir. Böylece hata oranları azaltılarak daha yüksek performans değerlerine ulaşılmaktadır [26]. Böylece düşük performans riski azaltılmış ve kullanılan derin öğrenme modeli için daha fazla uyum sağlanmış olur. Veri artırma işleminde genel olarak veri setinde yer alan verilere birtakım yöntemler uygulayarak ilave veriler oluşturulmaktadır [27]. Burada esas olan yeni oluşturulan verilerin etiketlerinin başlangıçtaki yapıdan farklı olmamasıdır. Böylece model ezberleme (overfit) eğiliminden uzaklaşmaktadır [28]. Bu nedenle bu çalışmada veri setlerindeki resimler üzerinde CLAHE tekniği kullanılarak veri artırma işlemi gerçekleştirilmiştir. Görüntü verilerinin kontrastlarında yapılan değişiklikler ile eğitim veri seti 4 kat

artırılmıştır. Böylece CNN ağ yapısında veri artırımının tahminleme performansını da yükselttiği gözlenmiştir.

Görüntü iyileştirmede adaptif histogram eşitleme sıkça kullanılan bir yöntemdir. Bu yöntem global kontras değerlerini kullanmaktadır. Bunu sonucu olarak görüntüde farklı bölgelerde çok karanlık bölgeler oluşabilmektedir. Bu sorunu çözebilmek adına alternatif bir yöntem olan CLAHE kullanılmaktadır [29]. CLAHE özellikle kullanıcı tarafından tanımlanan klip limiti değeri ile ne kadar kontrast artırılması gerekliliğini de belirler [30]. Ayrıca piksel dağılımları genellikle Rayleigh dağılımının özellikleri sergileyecek şekilde eşlenir. CLAHE'nin uygulamasında her bir bölgeye ait histogramlar ayrı ayrı hesaplanır. Daha sonra kullanıcı tarafından tanımlanan klip limit değerine göre her bir histogram yeniden düzenlenir. Son işlem olarak kümülatif dağılım fonksiyonları kontrast sınırlı histogramlar için hesaplanır. Bu adımlar şu şekilde detaylandırılmaktadır;

Adım 1: İlk olarak görüntü $M \times N$ şeklinde bölümlendirilir. Her bölgeye ait histogramı hesaplanır. Ortalama pixel sayısı [31]:

$$N_{avg} = \frac{(N_X \times N_Y)}{N_{gr}} \quad (1)$$

Denklemden yer alan N_X ve N_Y sırasıyla bölgenin x ve y boyutlarındaki piksel sayısıdır. Gri seviyelerin miktarı ise N_{gr} ile ifade edilir.

Adım 2: CLAHE Denklem 2 kullanılarak hesaplanır.

$$N_{CL} = N_{clip} \times N_{avg} \quad (2)$$

N_{CL} ve N_{clip} sırasıyla kontrast limiti ve kontrast limitin [0,1] aralığında normalize edilmiş halidir. Burada N_{CL} sınır değerdir ve bu değer üstüne çıkılırsa pikseller kırılır. Bu durumda eğer toplam piksel sayısı $N_{\Sigma clip}$ ile ifade edilirse her gri seviyeye dağıtılacak ortalama kalan pikseller Denklem 3 olarak sunulur.

$$N_{avggr} = \frac{N_{\Sigma clip}}{N_{gr}} \quad (3)$$

Histogram kırılma işleminde aşağıda verilen kural kullanılır:

$$\text{If } H(j) > N_{CL} \text{ then } H_{clip}(j) = N_{CL}$$

$$\text{Else if } (H(j) + N_{avggr}) \gg N_{CL} \text{ then } H_{clip}(j) = N_{CL} \quad (4)$$

$$\text{Else } H_{clip}(j) = H(j) + N_{CL} \quad (5)$$

j^{th} gri seviyede yer alan kırılmış ve orijinal histogram değeri $H_{clip}(j)$ ve $H(j)$ ile ifade edilir. Kalan piksellerin dağılımı ise tüm piksellerin dağıtımını bitene kadar devam eder.

Adım 3: Kırılmış histogram olarak tanımlanan $P_{input}(j)$ hem transferin gerçekleştirilmesinde hem de kümülatif olasılık hesaplamasında kullanılır. Buna göre çıkış olasılık yoğunluğu Denklem 6 kullanılarak aşağıdaki gibi hesaplanır:

$$p(y(i)) = \frac{1}{\alpha^2} (y(i) - y_{min}) e^{(-\frac{1}{2\alpha^2}(y(i)-y_{min})^2)} \quad \text{for } y(i) \geq y_{min} \quad (6)$$

α Rayleigh dağılımının ölçek parametresidir. y_{min} piksel değerinin alt sınırınıdır. $y(i)$ ifadesi Rayleigh ileri dönüşümü olarak tanımlanır ve aşağıdaki gibi hesaplanır:

$$y(i) = y_{min} + \sqrt{2\alpha^2 \ln\left(\frac{1}{1-p_{input}(i)}\right)} \quad (7)$$

Denklem 7 doğrusal kontrast kullanılarak aşağıdaki gibi yeniden şekillendirilebilir.

$$y(i) = \frac{x(i) - x_{min}}{x_{max} - x_{min}} \quad (8)$$

Transfer fonksiyonunun maksimum ve minimum değerleri x_{max} ve x_{min} ile tanımlanır. $x(i)$ ise transfer fonksiyonundan girdi değeri olarak değerlendirilir.

B. 2. Özellik Çıkarma

Makine öğreniminde özellikler, bir algoritmanın girdi olarak aldığı değişkenlerdir. Özellikler, girdiden elde edilen ham değerlerin belirli parçaları veya bu verilerden elde edilen değerler olabilir [32]. İleri beslemeli bir derin öğrenme modeli olarak CNN, çeşitli bilgisayarla görme ve örüntü tanıma görevlerinde bir dönüm noktası haline gelmiştir. CNN yapısında öznelik çıkarma konvolüsyon katmanlarından oluşmaktadır. Özellik çıkarma aşaması üç tür katman içerir: konvolüsyon katmanı, aktivasyon katmanı ve havuz katmanı [33]. Birden fazla özellik çıkarma aşamasını dönüşümlü olarak istifleyerek derin bir CNN ağı oluşturulabilir. Özellik çıkarma aşamaları, düşük boyutlu öznelik çıkarmayı amaçlar. Bunun içinde yüksek boyutlu verileri kullanır.

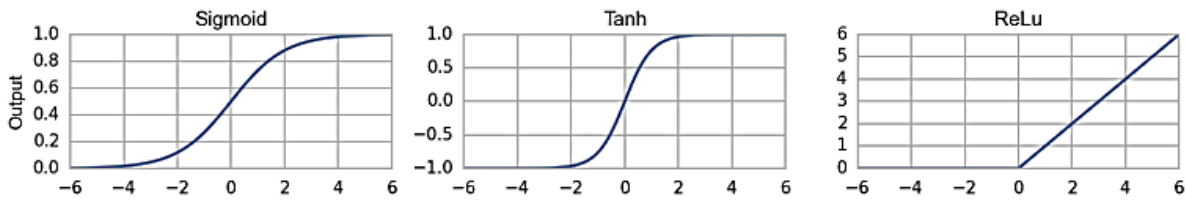
Çalışma alanına göre konvolüsyon katmanının sıralı bağlanma derinliği değişmektedir. Sınıflandırma katmanı ise düşük boyutlu öznelikler ile kategorilerin eşleşmesini sağlar ve genellikle tam bağlı bir yapay sinir ağı mimarisine sahiptir [34]. Denklem 9'da gösterilen konvolüsyon işleminde görüntünün bir maske yardımıyla filtrelenmesi sağlanır. Böylece keskinleştirilmiş ve yumuşatılmış görüntüler, açıl kenarlar, dikey kenarlar ve yatay kenarlar gibi orijinal görüntülerle aynı boyuta sahip bu görüntülerin türevleri elde edilmektedir [35].

$$X_{ij}^l = \sum_a^n \sum_b^n W_{ab} y_{(i+a)(j+b)}^{l-1} \quad (9)$$

Konvolüsyon katmanında konvolüsyondan sonra aktivasyon aşaması bulunmaktadır. Konvolüsyon işleminden sonra, sigmoid işlevi, tanh işlevi ve doğrultulmuş doğrusal birim (ReLU) [36] gibi doğrusal olmayan aktivasyon işlevi genellikle karmaşık özellikleri ifade etmek için kullanılır. Sigmoid ve tanh işlevleriyle karşılaştırıldığında, ReLU yalnızca öğrenmeyi hızlandırmakla kalmayıp aynı zamanda gradyan kaybolması sorununu da hafiflettiği için CNN'de yaygın olarak kullanılmaktadır. CNN'da denklem 10'da gösterilen ReLU fonksiyonu tercih edilmektedir. Bu fonksiyon sayesinde özellikle şekillerdeki negatif değerlerin elenmesi sağlanmaktadır.

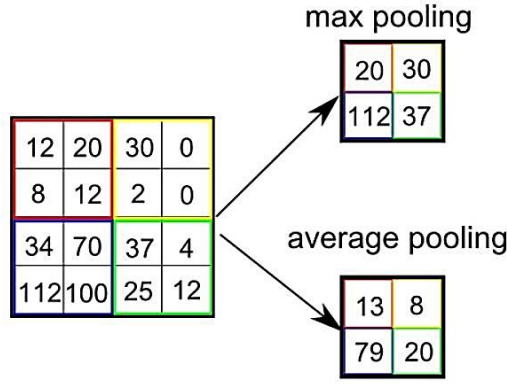
$$y = \max(0; x) \quad (10)$$

Ayrıca Şekil 5'te bazı aktivasyon fonksiyon grafikleri gösterilmiştir.



Şekil 5. Aktivasyon fonksiyon grafikleri

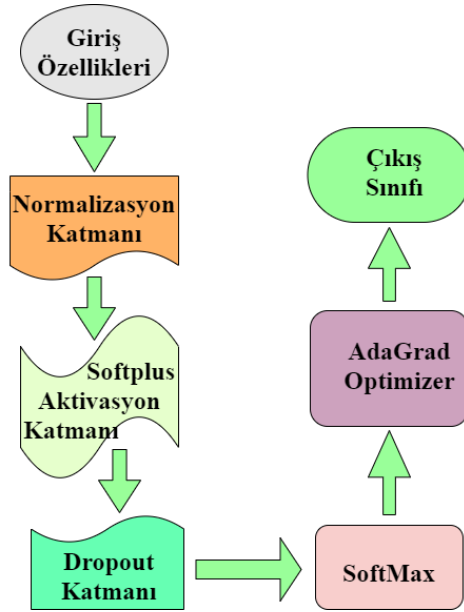
Konvolüsyon katmanındaki son aşamada havuzlama (pooling) işlemi fazlalığı önlemeyi ve Konvolüsyondan sonra özelliklerin daha sağlam olmasını sağlamayı amaçlar. Ortak havuzlama işlemleri, maksimum havuzlama ve ortalama havuzlamayı içerirken, aralarındaki fark havuzlama bölgesindeki maksimum veya ortalama değer alınır alınmadığıdır [37]. Şekil 6'da pooling örneği verilmiştir. Öznelikler çıkarıldıktan sonra elde edilen düşük boyutlu imgeler sınıflandırma katmanına aktarılır. Sinir ağının çıkışında hesaplanan hata değeri geri yayılım yöntemi kullanarak konvolüsyon filtre katsayılarının ve tam bağlı yapının katman ağırlıklarının güncellenmesinde kullanılır [35].



Şekil 6. Havuzlama katmanı

B. 3. MLP Kullanarak Sınıflandırma

Çok katmanlı algılayıcının [38] performansının artırılmasında Adagrad optimizier [39] ekleyerek geliştirilmiş MLP kullanılmıştır. Sınıflandırıcı, Şekil 7'de gösterildiği gibi üç katmandan oluşur.



Şekil 7. MLP katmanları

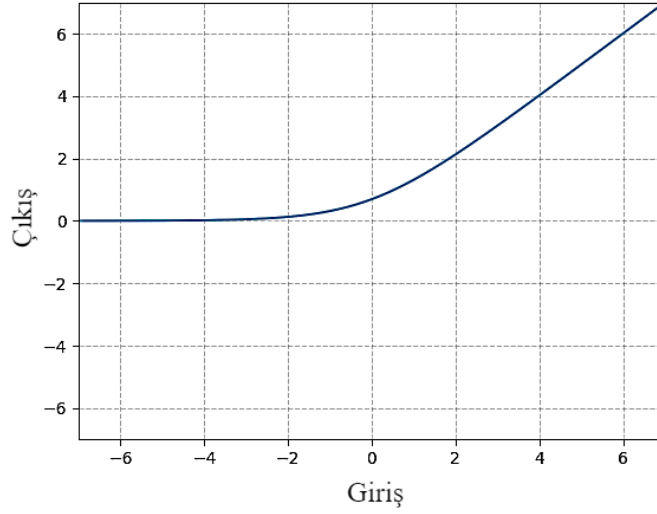
İlk katman verinin tüm görüntü özelliklerini normalleştirir. Bu normalleştirme girdi verilerinin benzer şekilde dağıtılmasını sağlamak için önemli bir adımdır. Bu, ağ eğitimi sırasında daha hızlı bir yakınsama sağlar. Normalizasyon fonksiyonu Denklem 11'de gösterilmektedir:

$$x'_i = \frac{(x_i - \bar{x})}{x_{var}} = \frac{x_i - \frac{1}{n} \sum_{j=1}^n x_j}{\sqrt{\frac{1}{n} \sum_{j=1}^n (x_j - \bar{x})^2}} \quad (11)$$

İkinci katmanda her bir düğümün sonucunu 0 ile $+\infty$ arasında eşleştirmek için Softplus aktivasyon fonksiyonu kullanılır. Bu katmanın sonucunda fonksiyon, sürekli 0 ile $+\infty$ arasında ölçekleyen, aşağıdaki Denklem 12 kullanılarak hesaplanır.

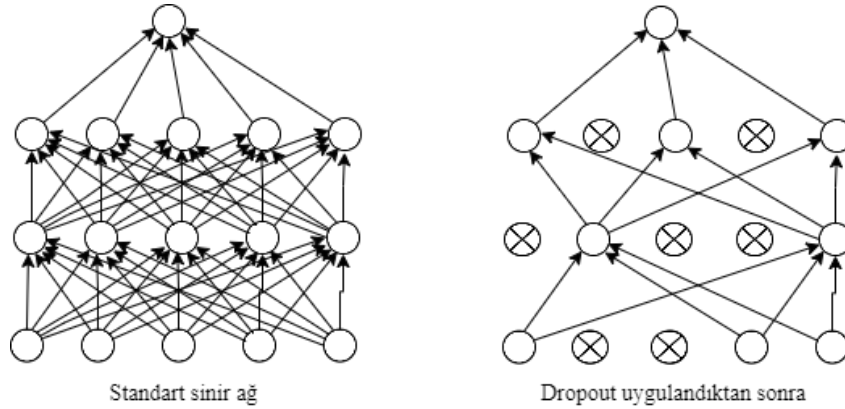
$$f(x) = \frac{1}{\beta} \log(1 + e^{\beta \cdot x}) \quad (12)$$

Softplus (Şekil 8), ReLU fonksiyonunun pürüzsüz bir yakınsamasıdır ve fonksiyonu sürekli pozitif çıktı vermeye zorlar. Eğer β yeterince büyürse, fonksiyon ReLU gibi olmaya başlar.



Şekil 8. Softplus fonksiyon grafiği

Üçüncü katmanda Dropout katmanı ile başlar. Dropout, bir sinir ağının gizli nöronlar arasındaki birkaç bağlantıyı rastgele kaldırarak ve yaygın olarak kullanılan bir düzenleme yöntemi. Bazı nöronları kaldırmak, eğitim sırasında aşırı öğrenmeyi (overfitting) engellemek için kullanılmaktadır [40]. Ayrıca, Dropout'un bir sınıflandırıcının bir girdi örneğindeki veya bir gizli özellik haritasındaki en ayırt edici bölgeye odaklanmasını engelleyebileceği [41] ve dolayısıyla genellemeyi iyileştirebileceği de gösterilmiştir. Şekil 9'da Dropout yöntemi gösterilmiştir.



Şekil 9. Dropout yöntemi

Dropout işlemini, girdi reel sayıları olarak bir V vektörü alan ve onu üstel girdi sayısı ile orantılı V olasılıklarından oluşan bir olasılık dağılımına normalleştirilen softMax fonksiyonunun kullanımı takip eder. Softmax işlevi, etkinliği nedeniyle sınıflandırma görevi için en popüler seçeneklerden biri haline gelmiştir. Matematiksel olarak Softmax işlemi şu şekilde tanımlanabilir [42]:

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}} \quad (13)$$

burada $\sigma(z)_i$, z 'nin i kategorisine ait olma olasılığını temsil eder.

Konvolüsyonel ağları oluşturduktan sonra ağı eğitimi için gerekli optimum parametreleri elde edilmesinde genellikle gradyan iniş yöntemleri kullanılır. Bu yöntemler arasında toplu gradyan azalma, stokastik gradyan azalma ve mini toplu gradyan azalma bulunur. Toplu gradyan azalma, konvolüsyonel sinir ağlarını eğitmek için yaygın olarak kullanılmaktadır [43]. Sinir ağına tüm parametrelerin güncellenmesinde aynı öğrenme oranının kullanımı, özelliklerin farklı frekansları varsa hepsini aynı ölçüde güncellenmeye bilmektedir. Bu zorluğa yanıt olarak, Adagrad, Adadelata [44], Kök Ortalama Kare prop ve Uyarlamalı Moment Tahmini [45] dahil olmak üzere uyarlanabilir optimizasyon algoritmaları ailesi geliştirilmiştir. Özellikle Adagrad, öğrenme hızını parametrelere göre uyarlamada yani sık görülen özelliklerle ilgili parametrelerde daha küçük güncellemeler yapması ama seyrek olanlar için daha büyük güncellemeler yapması sebebiyle kullanımı önerilmektedir [46]. Çalışmada son olarak, gradyan tabanlı optimizasyon için Adagrad algoritması kullanılmıştır.

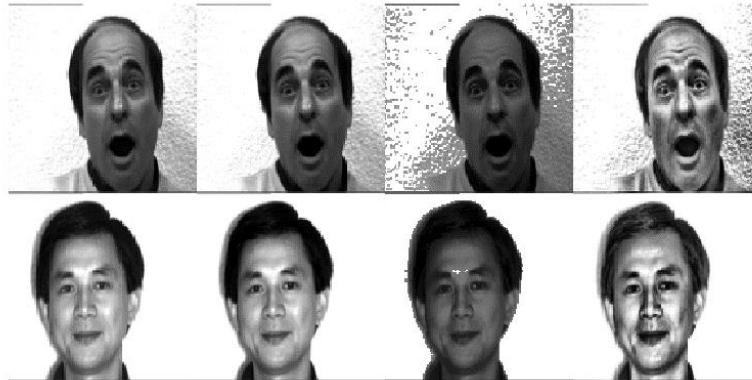
III. BULGULAR

Bu çalışmada 2 farklı veri seti ile 2 vaka çalışması gerçekleştirilmiştir. İlk vaka çalışmasında veri setleri üzerinde standart CNN yapısı kullanılarak sınıflandırma yapılmıştır. İkinci vaka çalışmasında ise önerilen AYTMM kullanılarak veri setlerinde sınıflandırma yapılmıştır. Yapılan sınıflandırma neticesinde elde edilen performanslar değerlendirilmiştir. Kullanılan veri setlerine ait özellikler Tablo 1’de verilmiştir.

Tablo 1. Veri seti özellikleri

	Yalefaces	Data_org
Veri Kümesi Sayısı	165	1850
Veri Kümesi Sınıf Sayısı	11	10
Renk Uzayı	Binary	RGB
ESA Benzerliği	Az	Az

CNN yapısının performansının artırılması için eğitim verileri setinde kullanılan görüntü verileri CLAHE tekniği ile 4 kat artırılmıştır. Veri artırma öncesi ve sonrası görüntülerin gösterimi Şekil 10’da verilmiştir.



Şekil 10. Veri artırma örnekleri (Sol baştakiler orijinal görüntülerdir)

Uygulamada kullanılan CHALE yöntemiyle veri kümeleri farklı sayıda eğitim verisi sayılarına ulaşmıştır. Ayrıca yalefaces veri seti 64x64, data_org veri seti ise 32x32 çözünürlüğe dönüştürülmüştür. CNN’nda öğrenme katsayısı olarak 0.001 kullanılmıştır.

DeneySEL sonuçlardaki doğruluk oranı, eğitim verisi dışından kalan test verilerinin sisteme uygulanarak elde edilen doğru tahminlerin yüzde oranıdır. Veri artırımı yöntemiyle eğitilen sistemler yüksek tahmin başarımıyla derin öğrenme yöntemlerinde ne kadar etkili sonuçlar ortaya koyduğu önemli bir gelişme olarak not edilmiştir. Veri artırımını gerçekleştirmeden önce yapılan sınıflandırma sonuçları ile veri artırımı yaptıktan sonra gerçekleştirilen sınıflandırma sonuçlarındaki hesaplanan doğruluk oranlarında belirgin farklılıkların olduğu Tablo 2’de görülmektedir.

Tablo 2. Veri artırımı sonrası sınıflandırma sonuçları

Eğitim Sonuçları				
Veri Kümeleri	Aşamalar	Veri boyutu	Doğruluk (%)	Yürütülme Zamanı(dk)
Yalefaces	Orjinal veri	165	%27.3	77
	Artırılmış veri	660	%97.0	223
Data_org	Orjinal veri	1850	%23.6	167
	Artırılmış veri	7800	%91.6	345

Tablo 2’de görüldüğü gibi önerilen AYTm yaklaşımı, veri setlerine ait eğitim veri sayısının artması, sınıflandırma doğruluk oranının da ciddi artış meydana getirmektedir. Her ne kadar kullanılan veri setleri yüz resimlerinden oluşsa da düşük veri setine sahip Yalefaces ve Data_org veri setleri üzerinde önerilen modelin, veri kümesinin özelliklerini çok iyi öğrendiğini göstermektedir. Çalışmada elde edilen Yalefaces veri setinin orijinal veriye ait karışıklık matrisi Şekil 11’de gösterilmiştir. Şekil üzerinde kategoriler tek tek ele alındığında en başarılı tanıma sahip 3 kategorinin %66.7 ile sağdan aydınlatma, soldan aydınlatma ve gözlüklü olduğu görülmektedir. Ayrıca 2 görüntü için %100 olmak üzere ve diğer 5 kategoride ise kendi kategorileri haricinde tanıma gerçekleştiği görülmektedir. Test verisinin başarısı %27.3 olmuştur.

Çıkış Sınıfı	Şaşırın	0 0.0%	0 0.0%	1 3.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100%
	Göz kırpan	1 3.0%	0 0.0%	0 0.0%	0 0.0%	1 3.0%	0 0.0%	1 3.0%	1 3.0%	0 0.0%	0 0.0%	1 3.0%	100%
	Merkez aydınlatma	0 0.0%	0 0.0%	1 3.0%	1 3.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	50.0%
	Gözlüklü	1 3.0%	0 0.0%	1 3.0%	2 6.1%	0 0.0%	0 0.0%	0 0.0%	1 3.0%	0 0.0%	0 0.0%	0 0.0%	40.0%
	Mutlu	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN%
	Soldan aydınlatma	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 6.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100%
	Gözlüksüz	0 0.0%	1 3.0%	0 0.0%	0 0.0%	0 0.0%	1 3.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100%
	Normal	0 0.0%	1 3.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100%
	Sağdan aydınlatma	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 6.1%	0 0.0%	0 0.0%	0 0.0%	100%
	Üzgün	0 0.0%	1 3.0%	0 0.0%	0 0.0%	2 6.1%	0 0.0%	2 6.1%	1 3.0%	1 3.0%	1 3.0%	1 3.0%	11.1%
	Uyuyan	1 3.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 6.1%	1 3.0%	25.0%
		0.0%	0.0%	33.3%	66.7%	0.0%	66.7%	0.0%	0.0%	66.7%	33.3%	33.3%	27.3%
		100%	100%	66.7%	33.3%	100%	33.3%	100%	100%	33.3%	66.7%	66.7%	72.7%
	Şaşırın	Göz kırpan	Merkez aydınlatma	Gözlüklü	Normal	Soldan aydınlatma	Gözlüksüz	Normal	Sağdan aydınlatma	Üzgün	Uyuyan		
	Giriş Sınıfı												

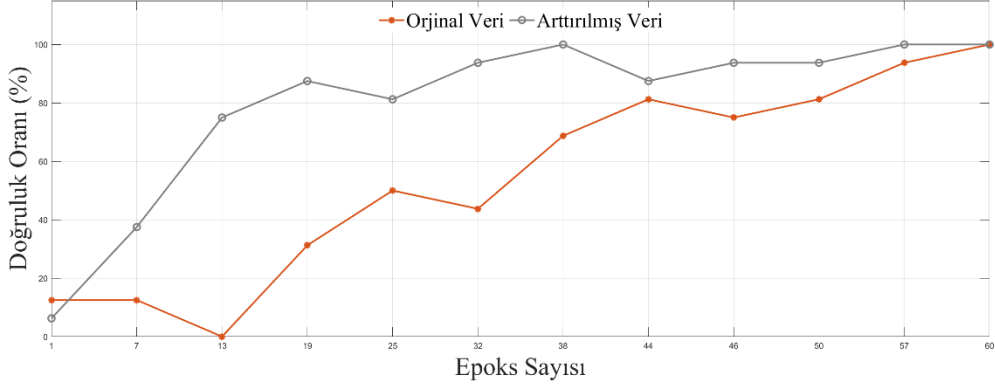
Şekil 11. Yalefaces veri seti ile birinci çalışma sonucu oluşan karışıklık matrisi

Yine aynı veri setinin veri artırımı ile elde edilen set ile gerçekleştirilen ikinci çalışmada ki karışıklık matrisi Şekil 12’de gösterilmiştir. Şekil incelendiğinde 10 kategorinin %100 tanıma gerçekleştirdiği ve sadece normal kategorisinin %66.7’de kaldığı görülmektedir. Ayrıca sadece bir kategoride “gözlüksüz” hatalı tanıma olduğu görülmektedir.

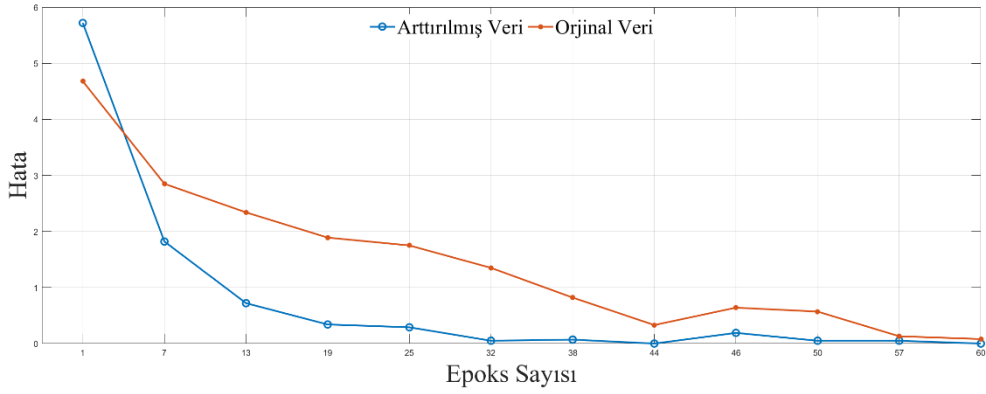
Çıkış Sınıfı	Şaşırın	3 9.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100%
	Göz kırpan	0 0.0%	3 9.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100%
	Merkez aydınlatma	0 0.0%	0 0.0%	3 9.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100%
	Gözlüklü	0 0.0%	0 0.0%	0 0.0%	3 9.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100%
	Mutlu	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 9.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100%
	Soldan aydınlatma	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 9.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100%
	Gözlüksüz	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 9.1%	1 3.0%	0 0.0%	0 0.0%	0 0.0%	75.0%
	Normal	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 6.1%	0 0.0%	0 0.0%	0 0.0%	100%
	Sağdan aydınlatma	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 9.1%	0 0.0%	0 0.0%	100%
	Üzgün	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 9.1%	0 0.0%	100%
	Uyuyan	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 9.1%	100%
		100%	100%	100%	100%	100%	100%	100%	66.7%	100%	100%	100%	97.0%
		0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	33.3%	0.0%	0.0%	0.0%	3.0%
	Şaşırın	Göz kırpan	Merkez aydınlatma	Gözlüklü	Normal	Soldan aydınlatma	Gözlüksüz	Normal	Sağdan aydınlatma	Üzgün	Uyuyan		
	Giriş Sınıfı												

Şekil 12. Artırılmış Yalefaces veri seti ile ikinci çalışma sonucu oluşan karışıklık matrisi

Şekil 13. a ve b’de, Yalefaces veri seti kullanılarak yapılan analizlerde önerilen modelde veri artırımı ve artırısız performans sonuçlarına ait doğruluk ve hata değerlerinin değişim grafiği gösterilmektedir. Her iki değişim grafiğinde incelendiğinde 44’üncü epok’da modelin yüksek performansa ulaştığı ve sonrasında elde edilen bu değer civarında değişimlerin devam ettiği görülmektedir. Böylece arttırılmış verilerde modelin daha yüksek performansa ulaştığını söylemek mümkündür.



(a)



(b)

Şekil 13. Yalefaces veri seti çalışmalarına ait (a) Doğruluk (b) Hata grafiği

Çalışmada diğer veri setinin orjinal ve veri artırımı yapılarak uygulanan çalışmalardan elde edilen karışıklık matrisi sırasıyla Şekil 14 ve Şekil 15’te gösterilmiştir. Şekil 14 incelendiğinde her kategoride tanıma olmasına rağmen %42.1 ile “öpücük atan” kategorisi en iyi tanıma başarısına sahiptir. En az karıştırılan kategori ise “Hayal kırıklığı” olmuştur. Şekil 15 incelendiğinde ise 9 kategorinin %89’un üstünde olduğu ve bunlardan 3 tanesinin ise %94’ten yüksek olduğu görülmüştür. Dolayısıyla her kategoride tanıma olmasına rağmen %94.9 ile sırtmak kategorisi en iyi tanıma başarısına sahiptir. En az karıştırılan kategori ise %97.2 ile “öpücük atan” kategorisi olmuştur.

Çıkış Sınıfı	Giriş Sınıfı										
	Kızgın	Ağlayan	Hayal kırıklığı	Sırtan	Öptüçük atan	Üzgün	Gülümseyen	Şaşırın	Konuşan	Göz kırpan	
Kızgın	7 1.8%	4 1.0%	5 1.3%	1 0.3%	5 1.3%	8 2.1%	3 0.8%	5 1.3%	5 1.3%	3 0.8%	15.2% 84.8%
Ağlayan	6 1.6%	13 3.4%	8 2.1%	6 1.6%	1 0.3%	10 2.6%	2 0.5%	5 1.3%	4 1.0%	7 1.8%	21.0% 79.0%
Hayal kırıklığı	3 0.8%	3 0.8%	4 1.0%	1 0.3%	5 1.3%	4 1.0%	3 0.8%	3 0.8%	2 0.5%	3 0.8%	12.9% 87.1%
Sırtan	3 0.8%	3 0.8%	0 0.0%	15 3.9%	0 0.0%	1 0.3%	10 2.6%	2 0.5%	1 0.3%	7 1.8%	35.7% 64.3%
Öptüçük atan	3 0.8%	4 1.0%	10 2.6%	1 0.3%	16 4.2%	3 0.8%	1 0.3%	4 1.0%	6 1.6%	1 0.3%	32.7% 67.3%
Üzgün	2 0.5%	3 0.8%	3 0.8%	0 0.0%	4 1.0%	7 1.8%	2 0.5%	2 0.5%	5 1.3%	4 1.0%	21.9% 78.1%
Gülümseyen	6 1.6%	3 0.8%	1 0.3%	12 3.1%	1 0.3%	1 0.3%	11 2.9%	2 0.5%	4 1.0%	4 1.0%	24.4% 75.6%
Şaşırın	4 1.0%	3 0.8%	4 1.0%	0 0.0%	4 1.0%	1 0.3%	4 1.0%	12 3.1%	4 1.0%	5 1.3%	29.3% 70.7%
Konuşan	1 0.3%	1 0.3%	1 0.3%	0 0.0%	1 0.3%	3 0.8%	1 0.3%	2 0.5%	3 0.8%	2 0.5%	20.0% 80.0%
Göz kırpan	2 0.5%	1 0.3%	2 0.5%	3 0.8%	1 0.3%	1 0.3%	2 0.5%	1 0.3%	4 1.0%	2 0.5%	10.5% 89.5%
	18.9% 81.1%	34.2% 65.8%	10.5% 89.5%	38.5% 61.5%	42.1% 57.9%	17.9% 82.1%	28.2% 71.8%	31.6% 68.4%	7.9% 92.1%	5.3% 94.7%	23.6% 76.4%

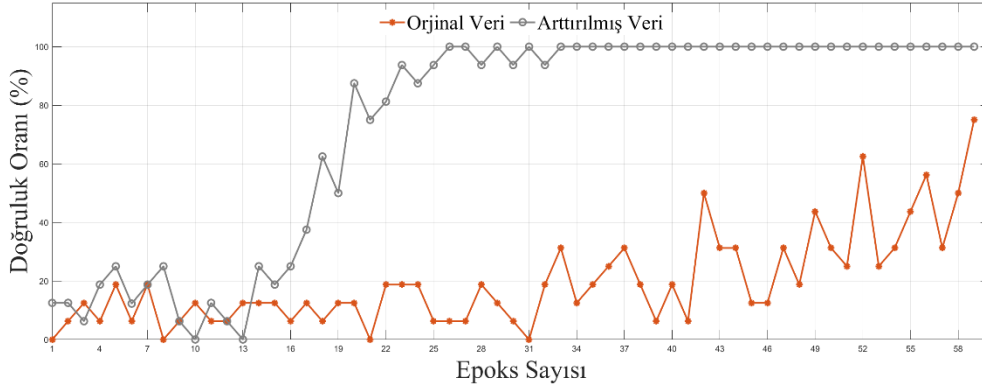
Şekil 14. Data_org veri seti ile birinci çalışma sonucu oluşan karışıklık matrisi

Çıkış Sınıfı	Giriş Sınıfı										
	Kızgın	Ağlayan	Hayal kırıklığı	Sırtan	Öptüçük atan	Üzgün	Gülümseyen	Şaşırın	Konuşan	Göz kırpan	
Kızgın	32 8.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.3%	0 0.0%	0 0.0%	97.0% 3.0%
Ağlayan	0 0.0%	36 9.4%	1 0.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.3%	0 0.0%	2 0.5%	90.0% 10.0%
Hayal kırıklığı	0 0.0%	0 0.0%	34 8.9%	0 0.0%	1 0.3%	2 0.5%	2 0.5%	0 0.0%	1 0.3%	1 0.3%	82.9% 17.1%
Sırtan	1 0.3%	0 0.0%	0 0.0%	37 9.7%	0 0.0%	0 0.0%	1 0.3%	0 0.0%	0 0.0%	0 0.0%	94.9% 5.1%
Öptüçük atan	0 0.0%	0 0.0%	0 0.0%	1 0.3%	35 9.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	97.2% 2.8%
Üzgün	1 0.3%	1 0.3%	2 0.5%	0 0.0%	1 0.3%	36 9.4%	0 0.0%	0 0.0%	1 0.3%	1 0.3%	83.7% 16.3%
Gülümseyen	1 0.3%	1 0.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	35 9.2%	0 0.0%	1 0.3%	0 0.0%	92.1% 7.9%
Şaşırın	1 0.3%	0 0.0%	0 0.0%	0 0.0%	1 0.3%	1 0.3%	0 0.0%	36 9.4%	0 0.0%	0 0.0%	92.3% 7.7%
Konuşan	1 0.3%	0 0.0%	1 0.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	35 9.2%	0 0.0%	94.6% 5.4%
Göz kırpan	0 0.0%	0 0.0%	0 0.0%	1 0.3%	0 0.0%	0 0.0%	1 0.3%	0 0.0%	0 0.0%	34 8.9%	94.4% 5.6%
	86.5% 13.5%	94.7% 5.3%	89.5% 10.5%	94.9% 5.1%	92.1% 7.9%	92.3% 7.7%	89.7% 10.3%	94.7% 5.3%	92.1% 7.9%	89.5% 10.5%	91.6% 8.4%

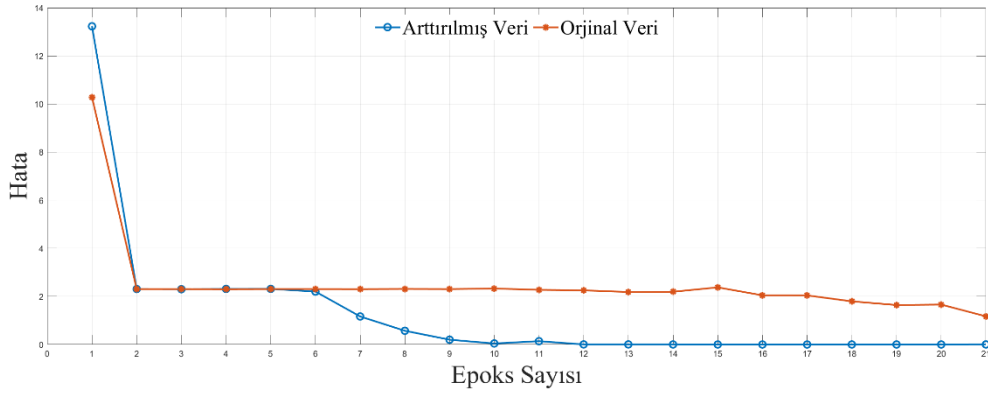
Şekil 15. Artırılmış Data_org veri seti ile ikinci çalışma sonucu oluşan karışıklık matrisi

Şekil 16. a ve b’de, Data_org veriseti kullanılarak yapılan analizlerde önerilen modelde veri artırılmış ve artırılmırsız performans sonuçlarına ait doğruluk ve hata değerlerinin değişim grafiği gösterilmektedir. Her iki değişim grafiğinde incelendiğinde, orijinal verinin en yüksek doğruluk değeri %75 olarak elde edilmiştir. Buna karşılık artırılmış veride ise yaklaşık olarak 32’inci epok’da en

yüksek performansa ulaşmıştır. Çalışılan bu veri grubunda da artırılmış veri seti ile daha yüksek performansa ulaşıldığını söylemek mümkündür.



(a)



(b)

Şekil 16. Data_org veri seti çalışmalarına ait (a) Doğruluk (b) Hata grafiği

IV. SONUÇLAR

Spesifik olarak, görüntüler üzerinden yüz tanıma konusunda geleneksel yöntemlere değinilmekle birlikte, esasen derin öğrenme mimarilerinin bu konudaki katkıları araştırılmıştır. Özellikle bu çalışmada, yüz tanıma çalışmalarında veri artırma işleminin önemi vurgulanmış ve yeni bir AYTMM geliştirilmiştir. Önerilen modelde özellikle konvolüsyonel sinir ağı modeli bilgisayarlı görüntü alanında başarılı sonuçlar vermesi nedeniyle tercih edilmiştir. Ayrıca kullanılan veri setleri ve yapılan çalışmaların başarımları tablolar halinde sunulmuştur. Tablo 2’de görüldüğü gibi önerilen yaklaşım ile veri setlerine ait eğitim verilerinin artırılması, sınıflandırmada doğruluk oranının da ciddi bir artış meydana getirmiştir.

Yalefaces veri setinin orijinal ve artırılmış veri karışıklık matrisleri tek tek ele alındığında orijinal veride en başarılı tanıma oranı %56.7 olurken artırılmış veride ise bir kategori hariç tüm kategorilerde %100 olarak gerçekleşmiştir. Artırılmış verinin, orijinal veriye göre eğitimde daha yüksek performansa ulaştığını ve hata oranlarında ise orijinal verinin işlendiği sınıflandırmalarda daha yüksek çıktığı görülmüştür. Çalışmada diğer veri setinden elde edilen karışıklık matrisinde ise en iyi tanıma oranı %42.1 olmuş, diğer yandan artırılmış veride ise %94.9 olarak gerçekleşmiştir.

Sonuç olarak, iki farklı veri seti başarımları karşılaştırıldığında veri artırımının sınıflandırmada ki doğruluk oranının yükselmesine neden olduğunu göstermiştir. Makine öğrenmesi ve derin öğrenme ile veri artırmaya dayalı yüz tanıma çalışmaları yapmak isteyen araştırmacılar için de sunulan çalışmanın yol gösterici olması hedeflenmiştir.

V. KAYNAKLAR

- [1] Y. Q. Li, D. T. Lin, and Z. W. Yeh, "Improving Deep Learning for Face Verification Using Color Histogram Equalization Data Augmentation," in *Proceedings of the 5th World Congress on Electrical Engineering and Computer Systems and Sciences*. 2019, Paper No. MVML 103, Prague, Czech Republic.
- [2] Y. Zhou, D. Liu, and T. Huang, "Survey of face detection on low-quality images," *13th IEEE international conference on automatic face & gesture recognition (FG 2018)*, 2018, pp. 769-773.
- [3] V. Kazemi, and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1867-1874.
- [4] J. Zhang, X. Wu, J. Zhu, and S. C. H. Hoi, "Feature agglomeration networks for single stage face detection," *Neurocomputing*, vol. 380, pp. 180-189, 2020.
- [5] C. Ren, N. An, J. Wang, L. Li, B. Hu, and D. Shang, "Optimal parameters selection for BP neural network based on particle swarm optimization: A case study of wind speed forecasting," *Knowledge-based systems*, vol. 56, pp. 226-239, 2014.
- [6] P. Viola, and M.J. Jones, "Robust real-time face detection," *International journal of computer vision*, vol. 57, no. 2, pp. 137-154, 2004.
- [7] K. Cui, H. Cai, Y. Zhang, and H. Chen, "A face alignment method based on SURF features," *10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 2017, pp. 1-6.
- [8] B. Ammour, T. Bouden, L. Boubchir, and S. Biad, "Face identification using local and global features," *40th International Conference on Telecommunications and Signal Processing (TSP)*, 2017, pp. 784-788.
- [9] L. Cuimei, Q. Zhiliang, J. Nan, and W. Jianhua, "Human face detection algorithm via Haar cascade classifier combined with three additional classifiers," *13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI)*, 2017, pp. 483-487.
- [10] H. Shu, D. Chen, Y. Li, and S. Wang, "A highly accurate facial region network for unconstrained face detection," *IEEE international conference on image processing (ICIP)*, 2017, pp. 665-669.
- [11] J. J. Lv, X. H. Shao, J. S. Huang, X. D. Zhou, and X. Zhou, "Data augmentation for face recognition," *Neurocomputing*, vol. 230, pp. 184-196, 2017.
- [12] M.R. Faraji, and X. Qi, "Face recognition under varying illuminations using logarithmic fractal dimension-based complete eight local directional patterns," *Neurocomputing*, vol. 199, pp. 16-30, 2016.
- [13] F. Shahali, A. Nazemi, and Z. Azimifar, "Single sample face identification utilizing sparse

discriminative multi manifold embedding,” *Artificial Intelligence and Signal Processing Conference (AISP)*, 2017, pp. 1-5.

[14] G. Bazoukis, S. Stavrakis, J. Zhou, S. C. Bollepalli, G. Tse, Q. Zhang, J. P. Singh and A. A. Armoundas, “Machine learning versus conventional clinical methods in guiding management of heart failure patients—a systematic review,” *Heart failure reviews*, vol. 26, no. 1, pp. 23-34, 2021.

[15] M. F. Aslan, K. Sabanci, and A. Durdu, “A CNN-based novel solution for determining the survival status of heart failure patients with clinical record data: numeric to image,” *Biomedical Signal Processing and Control*, vol. 68, p. 102716, 2021.

[16] J. W. Oh, and J. Jeong, “Data augmentation for bearing fault detection with a light weight CNN,” *Procedia Computer Science*, vol. 175, pp. 72-79, 2020.

[17] U.R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, and R. S. Tan, “Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals,” *Applied Intelligence*, vol. 49, no. 1, pp. 16-27, 2019.

[18] M. A. Kızrak, ve B. Bolat, “Derin öğrenme ile kalabalık analizi üzerine detaylı bir araştırma,” *Bilişim Teknolojileri Dergisi*, c. 11, s. 3, ss. 263-286, 2018.

[19] S. Akkol, A. Akilli, and I. Cemal, “Comparison of artificial neural network and multiple linear regression for prediction of live weight in hair goats,” *Yyu J. Agric. Sci*, vol. 27, pp. 21-29, 2017.

[20] Hinton, G., L. Deng, D. Yu, G. E. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, and B. Kingsbury, “Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups,” *IEEE Signal processing magazine*, vol. 29, no. 6, pp. 82-97, 2012.

[21] G. Li, L. Liu, X. Wang, X. Dong, P. Zhao, and X. Feng, “Auto-tuning neural network quantization framework for collaborative inference between the cloud and edge,” *International Conference on Artificial Neural Networks*, 2018, pp. 402-411.

[22] E. D. Cubuk, B. Zoph, D. Mane, V. Vasudevan, and Q. V. Le, “Autoaugment: Learning augmentation strategies from data,” *the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 113-123.

[23] V. Sharma, and R.N. Mir, “A comprehensive and systematic look up into deep learning based object detection techniques: A review,” *Computer Science Review*, vol. 38, p. 100301, 2020.

[24] C. Shorten, and T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” *Journal of Big Data*, vol. 6, no. 1, pp. 1-48, 2019.

[25] E. Cagli, C. Dumas, and E. Prouff, “Convolutional neural networks with data augmentation against jitter-based countermeasures,” *International Conference on Cryptographic Hardware and Embedded Systems*, 2017, pp. 45-68.

[26] L. Perez, and J. Wang, “The effectiveness of data augmentation in image classification using deep learning,” *arXiv preprint*, arXiv:1712.04621, 2017.

[27] B. McFee, E. J. Humphrey, and J. P. Bello, “A software framework for musical data augmentation,” *Proceedings of the 16th ISMIR Conference*, 2015, pp. 248-254.

[28] J. Salamon, and J. P. Bello, “Deep convolutional neural networks and data augmentation for environmental sound classification,” *IEEE Signal processing letters*, vol. 24, no. 3, pp. 279-283. 2017.

- [29] A. M. Reza, "Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement," *Journal of VLSI signal processing systems for signal, image and video technology*, vol. 38, no. 1, pp. 35-44, 2004.
- [30] M. S. Hitam, W. N. J. H. W. Yussof, E. A. Awalludin, and Z. Bachok, "Mixture contrast limited adaptive histogram equalization for underwater image enhancement," *International conference on computer applications technology (ICCAT)*, 2013, pp. 1-5, Sousse, Tunisia.
- [31] M. Kaur, R.K. Sarkar, and M.K. Dutta, "Investigation on quality enhancement of old and fragile artworks using non-linear filter and histogram equalization techniques," *Optik*, vol. 244, no: 167564, 2021.
- [32] O. A. Shawky, A. Hagag, E. S. A. E. Dahshan, and M. A. Ismail, "Remote sensing image scene classification using CNN-MLP with data augmentation," *Optik*, vol. 221, no. 165356, 2020.
- [33] R. Yan, J. Liao, J. Yang, W. Sun, M. Nong, and F. Li, "Multi-hour and multi-site air quality index forecasting in Beijing using CNN, LSTM, CNN-LSTM, and spatiotemporal clustering," *Expert Systems with Applications*, vol. 169, no. 114513, 2021.
- [34] W. Zhang, C. Li, G. Peng, Y. Chen, and Z. Zhang, "A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load," *Mechanical Systems and Signal Processing*, vol. 100, pp. 439-453, 2018.
- [35] K. Fırlıdak, ve M.F. Talu, "Evrişimsel Sinir Ağlarında Kullanılan Transfer Öğrenme Yaklaşımlarının İncelenmesi," *Computer Science*, vol. 4, no. 2, pp. 88-95. 2019.
- [36] V. Nair, and G.E. Hinton, "Rectified linear units improve restricted boltzmann machines," 27th International Conference on Machine Learning (ICML-10), 2010, Haifa, Israel.
- [37] K. Jarrett, K. Kavukcuoglu, M. A. Ranzato, and Y. LeCun, "What is the best multi-stage architecture for object recognition?," *12th international conference on computer vision*, 2009, pp. 2146-2153.
- [38] N. B. Gaikwad, N. B. Gaikwad, V. Tiwari, A. Keskar, and N. C. Shrivaprakash, "Efficient FPGA implementation of multilayer perceptron for real-time human activity classification," *IEEE Access*, vol. 7, pp. 26696-26706, 2019.
- [39] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," *Journal of machine learning research*, vol. 12, no. 7, 2011.
- [40] P. Flach, "Machine learning: the art and science of algorithms that make sense of data," 2012: *Cambridge University Press*.
- [41] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggenberger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Human brain mapping*, vol. 38, no. 11, pp. 5391-5420, 2017.
- [42] I. Goodfellow, Y. Bengio, and A. Courville, "Deep learning," 2016: *MIT press*.
- [43] S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv preprint arXiv:1609.04747*, 2016.
- [44] M. D. Zeiler, "Adadelata: an adaptive learning rate method," *arXiv preprint arXiv:1212.5701*, 2012.

- [45] D. P. Kingma, and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [46] J. Jiao, M. Zhao, J. Lin, and K. Liang, "A comprehensive review on convolutional neural network in machine fault diagnosis," *Neurocomputing*, vol. 417, pp. 36-63, 2020.