



Araştırma Makalesi

Yüz Tanımda Derin Öğrenme Mimarilerinin ve Yüz Bulma Yöntemlerinin Karşılaştırılması

Ayşe Merve Büyükbaş^{*1}, Ali Öztürk^{2,3}

¹KTO Karatay Üniversitesi, Ticaret ve Sanayi Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, Bilişim Güvenliği Teknolojisi, Konya, Türkiye

²KTO Karatay Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Bilgisayar Mühendisliği, Konya, Türkiye

³Havelsan A.Ş., Konya, Türkiye

ÖZ

Anahtar Kelimeler:

Derin öğrenme
Yüz bulma
Yüz tanıma
Eigenfaces
Fisherfaces

Bu çalışmada, literatürde yaygın olarak kullanılan görüntü işleme tabanlı yöntemler ve AlexNet, ResNet-18, GoogleNet ve SqueezeNet gibi mimariler kullanılarak performans karşılaştırılması yapılmıştır. Yüzün resim üzerinde belirlenebilmesi için Viola-Jones algoritması kullanılmıştır. Bu algorithmada kaskad obje dedektörü, yüzü algılayıp kare içine alır. Viola-Jones algoritmasının doğruluk oranı %85,71 olarak bulunmuştur. FEI yüz veri tabanındaki sağ, sol ve orta pozlarla veri kümesi oluşturulmuştur. Eigenfaces ve Fisherfaces görüntü işleme yöntemlerinin analizi için Temel Bileşen Analizi (TBA) ve Doğrusal Ayrım Analizi (DAA) kullanılmıştır. Bu yöntemler oluşturulan veri kümesi üzerinde uygulanarak doğruluk oranları elde edilmiştir. Eigenfaces yöntemi veri kümesindeki bazı poz varyasyonları için fisherfaces yönteminden daha iyi sonuç vermiştir. Derin öğrenme metodlarından AlexNet, ResNet-18, GoogleNet ve SqueezeNet kullanılmıştır. Yüz tanıma yöntemlerinden Eigenfaces yönteminin en yüksek doğruluk oranı %76,66 ve derin öğrenmede ResNet-18'in en yüksek doğruluk oranı %100 olmuştur.

Comparison of Deep Learning Architectures and Face Detection Methods in Face Recognition

Keywords:

Deep learning
Face detection
Face recognition
Eigenfaces
Fisherfaces

ABSTRACT

In this study, performance comparisons were made using image processing-based methods widely used in the literature and architectures such as AlexNet, ResNet-18, GoogleNet and SqueezeNet. Viola-Jones algorithm was used to determine the face on the image. In this algorithm, cascade object detector detects the face and frames it. The accuracy rate of Viola-Jones algorithm was found to be 85.71%. A dataset was created with right, left and middle poses in the FEI face database. Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) were used for the analysis of Eigenfaces and Fisherfaces image processing methods. These methods were applied on the created dataset and accuracy rates were obtained. Eigenfaces method gave better results than Fisherfaces method for some pose variations in the dataset. Deep learning methods AlexNet, ResNet-18, GoogleNet and SqueezeNet were used. Among the face recognition methods, the highest accuracy rate of the Eigenfaces method was 76,66% and the highest accuracy rate of ResNet-18 in deep learning was 100%.

*Sorumlu Yazar

^{*}(ayse.merve.buyukbas@karatay.edu.tr) ORCID ID 0000-0002-6534-7764
(ali.ozturk@karatay.edu.tr) ORCID ID 0000-0002-1797-2039

e-ISSN: 2717-8579

Geliş Tarihi: 27/08/2024; Kabul Tarihi: 04/12/2024

Bilgisayar Bilimleri ve Teknolojileri Dergisi

1. GİRİŞ

Yüz tanıma sistemleri, dijital resimler şeklinde kaydedilen insanlara ait yüz resimleri ile yapılan eğitimden sonra, yeni bir yüz resmi verildiğinde, kişinin kimliğinin belirlenmesi için geliştirilen sistemlerdir (Kekül vd., 2018).

Veri toplama, ön işleme, özellik çıkarma, model eğitimi, değerlendirme ve test etme bir makine öğrenimi yaklaşımı kullanılarak geliştirilebilir. Yüz resimlerinden oluşan veri kümesi ile, TBA (Temel Bileşen Analizi), Yerel İkili Örüntü (Local Binary Pattern: LBP), Evrişimsel Sinir Ağları (Convolution Neural Networks: CNN), Destek Vektör Makineleri (Support Vector Machine: SVM), Rastgele Orman (Random Forest) veya Yapay Sinir Ağları (Neural Networks) gibi algoritmalar eğitilebilir. Günümüzde yüz tanıma için kullanılan popüler yöntemler; Eigenface, Fisherface ve Yerel İkili Örüntü Histogramı (Local Binary Pattern Histogram: LBPH) algoritmalarıdır (Balanageshwara vd., 2023).

Yapay görme, makine görme, hesaplamalı görme veya görüntü analizi olarak da adlandırılan bilgisayarlı görme, bir bilgisayarı araç olarak kullanarak görüntülerden bilgi çıkarma işlemidir. Yüz tespiti, görüntüdeki bir veya daha fazla kişinin yüzünün, görüntü içindeki arka planı veya içinde bulunan diğer nesnelere göz ardı ederek bulunmasını sağlayan tekniktir. Başlangıçta bir sınıflandırıcıyı eğitmek için çok sayıda görüntüye ihtiyaç vardır. Örneğin, bir yüz dedektörü geliştirmek için yüz resmi içeren ve içermeyen görüntüler gereklidir (Cadena vd., 2023).

Gerçek zamanlı yüz tanıma, uygulamalarda oldukça popülerdir. Arya ve Tiwari, Eigenface, Fisherface ve LBPH algoritmalarını kullanan gerçek zamanlı otomatik bir yüz tanıma ve tespit sistemi yapmışlardır. Haar Cascade ile Eigenface, Fisherface ve LBPH algoritması kullandıklarında kişiyi -30° ila $+30^{\circ}$, -60° ila $+60^{\circ}$ ve -60° ila $+75^{\circ}$ açılardan başarıyla tespit edip tanıyabildiği sonucuna varmışlardır. Ayrıca kişinin ön yüzünü yukarıdan aşağıya veya tersi yönde döndürdüğü zaman kişiyi tespit edip, ön yüzü eğik olan kişiyi de tespit ederek tanıyabilmişlerdir. Ön yüzü eğik olan kişiyi tespit etmek ve tanımak için LBPH ve Fisherface kullanan sistem $\pm 10^{\circ}$ eğim açısıyla en iyi çalışma performansını vermiştir. Ön yüz açısı tamamen yukarı ve aşağı olduğunda, Eigenface ve LBPH algoritmasını kullanan sistem hem normal ışık koşullarında (gündüz) hem de düşük ışık koşullarında (gece) en iyi sonucu vermiştir (Arya ve Tiwari, 2020).

Viola-Jones algoritması, bir insan yüzünün özelliklerini arayan pencere ile bir görüntüyü tarar. Özellikleri bulursa ve bir yüz olarak belirli bir değere sahipse, görüntünün belirli penceresinin yüz olduğu tahmin edilir. Farklı boyutlarda yüzlerin olduğu bir durumu çözmek için, pencere her görüntü için tekrarlanan işlemle ölçeklendirilir (Rahmad vd., 2020).

Holat ve Kulaç (2014) yaptığı çalışmada, yüz tanıma sistemi için kameradan alınan anlık görüntüyü ve bilgisayarda kayıtlı olan görüntüyü kullanarak yüz tespiti yapmışlardır. Görüntüyü edinme aşaması; test görüntüsüyle yüzü bulma aşamasından, yüz konum bilgisi üzerinden ön işlemenin yapılması, tanımanın yapılabilmesi için özellik çıkarıcı ve sınıflama işleminden oluşmaktadır. Yüz tanıma için kullanılan TBA ve DAA (Doğrusal Ayrım Analizi) ve HE (Histogram Eşitleme) yöntemleri için Medyan, Gauss ve Laplace filtreleri kullanmışlardır. Yale veri tabanı ile yapılan denemelerde en iyi sonuçlar merkezden aydınlatılmış görüntüler eklendiğinde elde edilmiştir. En iyi yöntem ise LBP+HE+Medyan yönteminde %85 başarı oranına sahiptir. ORL veri tabanı ile yapılan denemelerde de en iyi sonuçlar LBP+HE+Medyan yönteminde %90 başarı oranı olarak bulunmuştur (Holat ve Kulaç, 2014).

Yapay sinir ağları tabanlı derin öğrenme mimarilerinden biri de evrişimsel sinir ağlarıdır. Günümüzde literatürde resim üzerinde sınıflama yöntemlerinde yaygın olarak kullanılmaktadır. Transfer öğrenimi için ön eğitim, özellik çıkarıcı ve kısmi özellik çıkarıcı yöntemler kullanılmıştır (Doğan ve Türkoğlu, 2019). Yapılan çalışmada, veri setindeki sınıflandırmaların AlexNet eğitim setiyle olan benzerliği transfer öğrenimi üzerindeki başarısını arttırmıştır. En yüksek doğruluk oranı %99,02 olarak Mnist veri kümesindedir (Fırıldak ve Talu, 2019).

Derin öğrenmenin gelişmesiyle CNN tabanlı yüz tanıma teknolojisi bu alanda temel yöntem haline gelmiştir. Algoritmalar açısından CNN konvolüsyon katmanı ile diğer konvolüsyon katmanları arasında paylaşım parametreleri vardır. Eğitilecek parametre sayısının bellek gereksinimlerine bağlı olarak azalması da bir avantajdır (Wang ve Li, 2018).

CNN fikri yapay sinir ağlarından geliştirilmiştir. Biyolojik sinir ağı, doğrusal olmama, eşzamanlılık, sağlamlık ve yüksek hata toleransı ile karakterize edilir. Bu özelliklerinden dolayı yapay sinir ağları görüntü işleme gibi alanlarda oldukça yaygın kullanılmaktadır. Wang ve arkadaşları, araştırmasında 8 katmanlı geleneksel kedi türlerinin tanımlanması, 4 evrişimsel katmanı, 2 havuzlama katmanı, tam bağlı katman ve çıkış katmanından oluşan CNN modeli kullanmışlardır.

Konvolüsyon katmanı birden çok katmandan oluşur. Buradaki amaç, giriş verileri üzerinde işlem yaparak ve farklı özellikler oluşturarak konvolüsyonu gerçekleştirmektir. Aynı zamanda görüntü çözünürlüğünü azaltmak ve hesaplamayı kolaylaştırmak için, konvolüsyon katmanı görüntü kolerasyon ilkesine göre bir araya getirilmiştir. Geliştirilen modelde doğruluk testi %68,85'ten %79,41'e yükselmiştir (Wang vd., 2020).

Topal ve arkadaşları (2023) ImageNet üzerinde eğitilmiş CNN görüntülerine karşı evrişimsel algoritma (EA) tabanlı düşmanca saldırı önermişlerdir. EA tabanlı saldırıları, sınıflandırılma

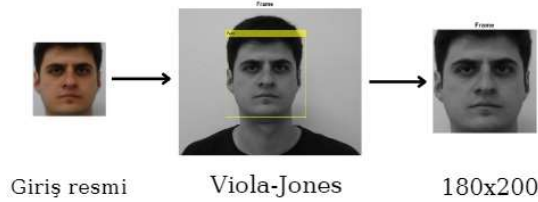
konusunda güven olasılığı en az %75 olarak bulmuşlardır. Önerdikleri EA tabanlı saldırının, başarı oranı ve üretilen rakip görüntülerinin görsel kalitesi açısından rakiplerine göre üstün veya eşit olduğunu ortaya koymuştur (Topal vd., 2023).

2. YÖNTEM

Çalışma kapsamında yürütülen deneyler işlemci olarak Intel Core i7-10750H CPU, ekran kartı özelliği nVIDIA GeForce RTX2060 Intel UHD Graphics, @2.60 GHz temel frekansı ve 16 GB RAM'e sahip 64 bit Windows 10 Home SL işletim sistemine sahip dizüstü bilgisayar üzerinde yapılmıştır. Gerçekleştirilen deneylerde MATLAB R2018b yazılımı kullanılmıştır.

Viola-Jones yöntemi görüntü düzleminin dışında döndürülen ve görüntü düzlemi etrafında döndürülen görüntülerdeki yüz varyasyonlarını inceler. Viola-Jones kaskad obje dedektörü, görüntü üzerinde oluşan pencereyi kaydırarak nesnelere algılamaya yarar. (Jones ve Viola, 2003).

Yüz tanımının yapılabilmesi için veri setindeki resimlerin ilk olarak Viola-Jones algoritmasından başarılı bir şekilde geçmesi gerekmektedir. Yapılan bu çalışma için toplam 686 yüz fotoğrafı bulunmaktadır. Görüntüler siyah beyaz renk uzayına indirgenmiştir. Yeni oluşturulan veri kümesindeki görüntü örneği ve boyutu Şekil 1'de gösterilmiştir.



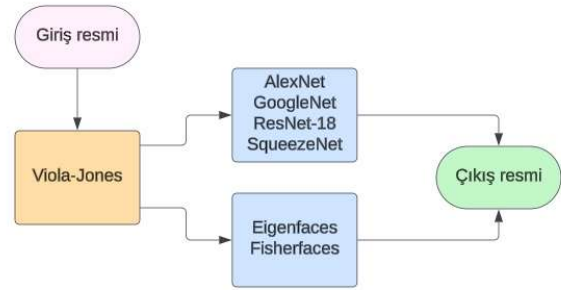
Şekil 1. Viola-Jones algoritması ile tanınan yüz resmi

Giriş görüntüsündeki yüz resmini algılamak amacıyla kaskad obje dedektörü oluşturulur. Dedektör, fotoğraf üzerindeki yüzü arayarak belirlemektedir. Belirlenen yüz nesnesi pencere içine alınmaktadır.

Yüz bulma yöntemleri bilgiye dayalı yöntemler, değişmez özellikli yaklaşım yöntemleri, şablon eşleştirme yöntemleri ve görüntü işleme tabanlı yöntemler olmak üzere dört kategoride incelenmiştir. Bilgiye dayalı yöntemlerde, test görüntüsündeki insan yüzüne ait özellikleri çıkartmak ve bu özelliklere göre aday yüzleri bulmak kolaydır. Ancak arka planda oldukça başarılı sonuçlar üretmesine rağmen, farklı pozlardaki resimler için bu yöntemin uygulanması oldukça zorlayıcıdır. İnsana ait özellikleri kurallara veya kodlara dönüştürmek her zaman mümkün olmayabilir. Değişmez özellikli yaklaşım yöntemlerinde, aday resimlerdeki insan yüzlerini bulmak için renk yoğunluğu ve çeşitliliği, kenar ve desen gibi özellikler kullanılır. Ancak aydınlatma ve

resimlerdeki diğer gürültüler yüzdeki özelliklerin bulunmasını zorlaştırmaktadır. Şablon eşleştirme yöntemleri, yüz resmindeki baskın türde olan özellikleri kullanarak verilen test görüntüsünün üzerindeki yüzleri bulmayı dener. Hesaplaması oldukça kolaydır, ancak yüze yakın yerlerde tarama yapılmazsa maliyetli olabilir. Görüntü işleme tabanlı yöntemlerde öncelikle giriş resmi üstünde ön işlem uygulanır. Test ve eğitim için oluşturulan resimler standart hale getirilir. Sınıflandırmanın yapılabilmesi için pozitif ve negatif algoritmalar sayesinde giriş verilerinin eğitilmesi gerekir. Görüntü işleme tabanlı yöntemler başarı oranı yüksek makine öğrenmesi algoritması kullanırlar. Hızlı ve etkin çalışmalarının yanı sıra başarısı ispatlanmış sonuçlar üretirler (Sütçüler, 2006).

Viola-Jones algoritması ile çerçeve içine alınan yüz resimleri derin öğrenme mimarileri ve yüz bulma algoritmaları olmak üzere iki farklı işlemden geçmiştir. Bu işlemler Şekil 2'de gösterildiği gibidir.



Şekil 2. Viola-Jones ile derin öğrenme mimarileri ve yüz tanıma algoritması

Giriş resmi, Viola-Jones algoritmasında algılandıktan sonra derin öğrenme mimarileri için AlexNet, GoogleNet, ResNet-18 ve SqueezeNet kullanılmıştır. Yüz tanıma yöntemleri için Eigenfaces ve Fisherfaces yöntemleri kullanılmıştır.

2.1. Veri Kümesi

Yapılan bu çalışmada FEI yüz veri setindeki bazı imgeler kullanılmıştır (Thomaz, 2012). FEI yapay zekâ laboratuvarında, 19-40 yaş aralığındaki çalışanların ve öğrencilerin farklı açılardan çekilen pozları veri setini oluşturmuştur. Toplamda 200 farklı kişinin 14 farklı pozlarından oluşmuştur. Her görüntünün boyutu 640x480 pikseldir.

Test ve eğitim klasörlerinden oluşan bu veri kümesinde 140 yüz resmi derin öğrenme mimarileri, Eigenfaces ve Fisherfaces algoritmaları için ayrılmıştır. 546 yüz resmi derin öğrenme mimarileri için ayrılmıştır. Derin öğrenme mimarilerinin özelliklerine göre 224x224 veya 227x227 piksel boyutlarına indirgenerek yeni klasörde toplanmıştır. Toplam 686 adet görüntü yer almaktadır. Görüntü veri setlerinin %25'i test ve %75'i eğitim (train) klasörü olarak ikiye ayrılmıştır.

Veri kümesine ait görüntüler 14 farklı yüz pozlarından oluşmaktadır ve bu görüntüler Şekil 3'te gösterilmiştir.



Şekil 3. FEI bir kişiye ait 14 farklı poz (Thomaz, 2012)

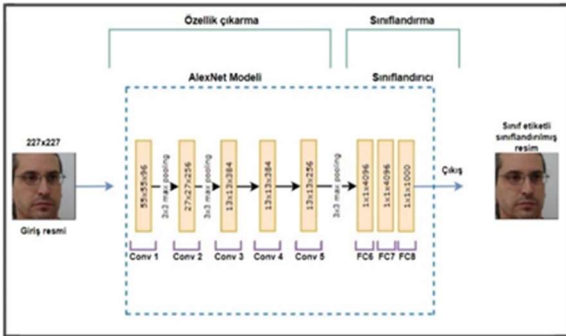
2.2. Derin Öğrenme Mimarileri

Derin öğrenme mimarileri konvolüsyon katmanları, havuzlama katmanları, tam bağlı ve sınıflandırma katmanından oluşmaktadır. Mimarideki ağlar, giriş verisi üzerindeki katmanlar arasında işlemlerden geçerek eğitilmektedir (Sert, 2020). Bu makalede yapılan çalışma için kullanılan mimariler AlexNet, ResNet, GoogleNet ve SqueezeNet olmak üzere dört farklı grupta incelenmiştir.

2.2.1. AlexNet Mimarisi

AlexNet mimarisi ağ ağırlıkları olan sekiz katman içerir; ilk beşi evrişimli ve kalan üçü tamamen bağlı katmanlardır (Krizhevsky vd, 2017).

Giriş katmanı $227 \times 227 \times 3$ boyutunda olup AlexNet mimarisi özellikleri Şekil 4'teki gibidir.



Şekil 4. AlexNet mimarisi özellikleri (Almabdy ve Elrefaei, 2019)

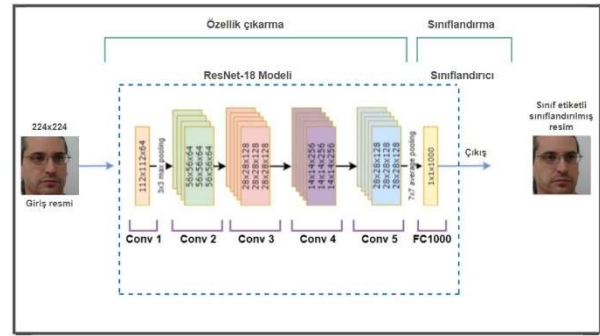
Giriş resmi ilk olarak Viola-Jones algoritmasından geçerek yüzü bulma işlemi tamamlanmıştır. Yüzü bulunan görüntüler, eğitim aşamasından geçerek örnek resimleri ve bunların sınıflandırmaları sağlanmıştır. Epoch değeri, eğitim aşaması boyunca geçecek olan her veri için ağ üzerinde gösterilmeyi sağlar. Mimarideki ağlar eğitilirken, doğruluğu arttırabilmek ve daha iyi sonuçlar elde edebilmek amacıyla epoch sayıları değiştirilerek denemeler yapılmıştır.

2.2.2 ResNet Mimarisi

Derin kalıntı ağı veya ResNet mimarisi, He ve arkadaşları tarafından geliştirilmiş bir modeldir (He

vd., 2016). Derin öğrenme eğitimindeki ikilemleri yenmek için oluşturulmuştur. Çünkü derin öğrenme eğitimi oldukça zaman alır ve belirli sayıda katmanla sınırlıdır. ResNet modelinin diğer mimari modellere kıyasla avantajı, mimari derinleşse bile bu modelin performansının düşmemesidir (Sarwinda vd., 2021).

ResNet18 mimarisi 72 katman ve 18 ağ derinliği, ResNet50 mimarisi 177 katman ve 50 ağ derinliği, ResNet101 mimarisi ise 347 katman ve 101 ağ derinliğine sahip önceden eğitilmiş ağ mimarileridir. Giriş görüntü katmanı $224 \times 224 \times 3$ boyutundadır (Raghu vd., 2020). ResNet-18 mimarisinin özellikleri Şekil 5'te gösterilmiştir.



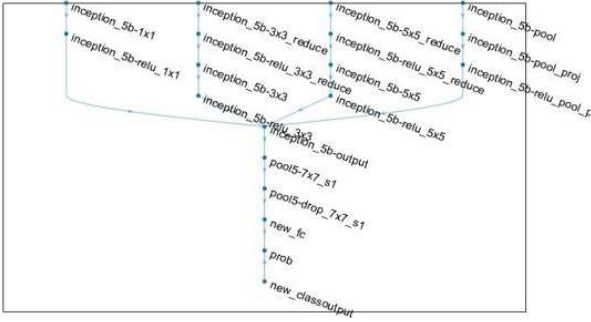
Şekil 5. ResNet-18 mimarisi özellikleri (Almabdy ve Elrefaei, 2019)

$224 \times 224 \times 3$ boyutundaki giriş resmi Viola-Jones algoritmasından geçtikten sonra, eğitime girmiştir. Eğitim aşaması; özellik çıkarma ve sınıflandırma adımlarından oluşmaktadır. Özellik çıkarma ve sınıflandırma adımlarından sonra sınıf etiketi oluşmuş resim elde edilmiştir.

2.2.3 GoogleNet Mimarisi

Szegedy ve arkadaşları tarafından 2014'te ImageNet Büyük Ölçekli Görsel Tanıma Yarışması (ILSVRC) için kurulan GoogleNet adlı takım, iki farklı kategoride birinci olmuştur. İlk kategoride, ek eğitim verileriyle nesne tespitinde, algılama modelleri topluluğu için doğruluk oranı %44,5'tir. Sınıflandırma ve yerelleştirme olarak adlandırılan ikinci kategoride toplam 5 değer skorunda %6,66 hata vermiştir. Hebbian ilkesinden elde edilen çok ölçekli fikrini, evrişimli sinir ağı mimarisiyle birleştiren çalışmadır. Fikirleri birleştirmek toplam parametre sayısını azaltırken, evrişimsel katmanlardaki parametre sayısını önemli derecede attırmayı sağlamıştır (ImageNet Large Scale Visual Recognition Challenge, 2014).

Transfer öğrenimi, yeni problemin temeli olarak onu baştan eğitmek yerine önceden eğitilmiş ağı kullanmayı tercih eder (Şeker A, 2018). Derin öğrenme ile yapılan çalışmaların çoğunda başarı oranını arttırdığı gözlenmiştir. GoogleNet mimarisinde, konvolüsyon ve havuzlama katmanlarından geçen ağlar, transfer öğrenimi için eğitilmiştir. Eğitilmiş ağların son aşaması Şekil 6'da gösterildiği gibidir.



Şekil 6. Eğitilen ağı transfer öğrenimi

2.2.4 SqueezeNet Mimarisi

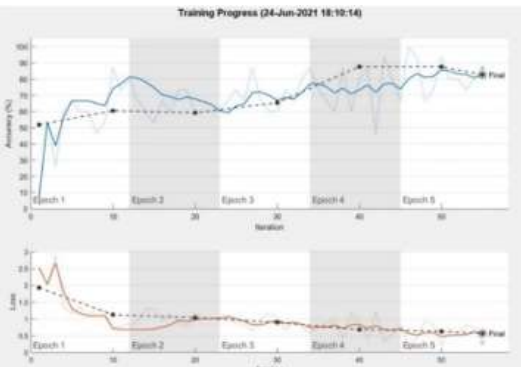
Iandola ve arkadaşları yapmış oldukları çalışmada, evrimsel sinir ağlarında doğruluğu arttırmak için SqueezeNet mimarisini tasarlamışlardır. Tasarım için üç ana strateji kullanmışlardır: ilki 3x3 filtreleri 1x1 boyutlu filtrelerle değiştirmektir. Böylece dokuz kat daha az parametreye sahip filtre kullanmışlardır. İkinci olarak, filtrelere giriş kanalı sayısını 3x3 filtrelere düşürmektir. Mimari bir konvolüsyon katmanı, 8 ateşleme modülü ve son konvolüsyon katmanından oluşmaktadır (Iandola vd., 2016).

Çalışma için oluşturulan veri setindeki giriş yüz görüntüleri, 227x277x3 boyutuna çevrilerek ayrı klasöre eklenmiştir. Orijinal boyutlu görüntüler farklı klasörde tutulmuş olup, bu görüntülerin sınıflandırma örnekleri Şekil 7’de gösterilmiştir.



Şekil 7. SqueezeNet sınıflandırma örnekleri

Örnek sınıflandırma işleminden sonra, ağı tekrar eğiterek transfer öğrenimi gerçekleştirilmiştir. Eğitim aşamasındaki başarı oranı eğrisi Şekil 8’deki gibidir.



Şekil 8. ImageNet ağı eğitimi

2.3 Yüz Tanıma Yöntemleri

2.3.1 Eigenfaces ile Yüz Tanıma

Eigenfaces ile yüz tanıma metodunda, veri setindeki tüm görüntüler siyah beyaz renk uzayına çevrilmiştir. Görüntünün piksel boyutları 180x200 olarak indirgenmiştir. Görüntü veri tabanı matrisini ortalama görüntüden farkı alınmış görüntü matrisi oluşturur. Görüntüler $N^2 \times M$ boyutlu veri matrisi haline getirilmiştir. Eigenfaces ile elde edilebilir olması için, görüntü veri tabanı matrisiyle kovaryans matrisi işlemi gerçekleştirilir.

Eğitim setindeki 140 görüntüden 120 tanesi Viola-Jones algoritmasından başarıyla geçmiştir. Yüz tabanındaki kişilerin 14 farklı görüntüsü yerine, algorithma algılanan 12 farklı yüz görüntüleri kullanılmıştır. Yeni oluşan eğitim setindeki bir kişinin 12 farklı görüntüsü Şekil 9’daki gibidir.



Şekil 9. Eğitim setindeki örnek görüntü kümesi

Kaskad obje dedektörü aracılığıyla kare içine alınan resimler kırılarak, 431x360 örnek piksel boyutlu resim 180x200 piksel boyutuna düşürülmüştür.

10 ayrı kişiden oluşturulan 120 farklı fotoğrafın 30’u test, 90 tanesi eğitim klasörü olmak üzere iki klasöre ayrılmıştır. Eğitim ve test görüntüsünün karşılaştırması Şekil 10’da gösterilmiştir.



Şekil 10. Veri setindeki eğitim ve test görüntüsü

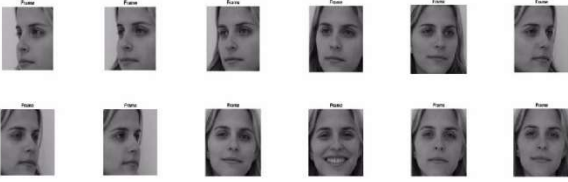
Veri setindeki iki görüntü birbiriyle karşılaştırıldığında, aynı kişiye ait olduğu görülmektedir.

2.3.2 Fisherfaces ile Yüz Tanıma

Fisherfaces yaklaşımı DAA’ya dayanmaktadır. DAA sınıf içi değişkenliği en aza indirirken, sınıflar arası değişkenliği en üst düzeye çıkaran verilerin doğrusal kombinasyonlarını bulmayı amaçlayan denetimli bir boyutluluk azaltma yöntemidir (Çakmakoğlu, 2018). Giriş verilerindeki farklı sınıflar arasında en iyi ayrımı sağlayan yeni

indirgenmiş alt uzayı bulmaya çalışır. Bu temel fikir, TBA uygulamasına benzer şekilde yüz tanıma uygulanır. Her yüz görüntüsü daha yüksek boyutlu bir uzayda nokta olarak kabul edilir. Ardından, fisherfaces adı verilen temel vektörleri elde etmek için verilere DAA uygulanır. Yüz görüntüleri daha sonra eşleştirmenin gerçekleştirildiği bu temelde yansıtılır (Aly M., 2006).

Eğitim setindeki 140 görüntünün 120 tanesi yüz bulma dedektöründen geçmiştir. Görüntüler siyah beyaz renk uzayında 180x200 piksel boyutuna indirgenerek oluşturulmuştur. Eğitim setindeki bir kişiye ait yüz görüntü kümesi örneği Şekil 11'de gösterilmiştir.



Şekil 11. Eğitim setindeki bir kişinin görüntüleri

Yüz görüntülerinde farklı yüz ifadeleri ve farklı açılardan çekimler olduğu görülmektedir.

3. BULGULAR

Bu bölümde ilk olarak Viola-Jones algılama dedektörü ile yüz görüntülerinin ne kadarının algılanıp algılanmadığı, buna bağlı olarak doğruluk oranları gösterilmiştir. Viola-Jones algoritmasından algılanarak geçen yüz resimleri yüz tanıma yöntemlerinde kullanılan veri setini oluşturmaktadır. Veri setindeki görüntüler eğitim ve test olmak üzere iki kategoriye ayrılmıştır. İkinci olarak, yüz tanıma Eigenfaces ve Fisherfaces yöntemleri kullanılmıştır. Eigenfaces yönteminde, giriş görüntü resmi için ortalaması ve matris boyutlarıyla özdeğer ve özvektörler yardımıyla ağırlık matrisi bulunmuştur. Öklid mesafesine göre karşılaştırması yapılarak görüntünün eş değeri test görüntüsünde elde edilmiştir. Fisherfaces yönteminde, giriş resmi üzerinden resmin ortalaması ve sınıf içi ortalaması aracılığıyla öz vektör matrisi bulunmuştur. Test ve eğitim görüntülerinin öklid mesafesi aracılığıyla eşleştirmeleri yapılmıştır. Çalışmanın sonraki adımlarında, derin öğrenme modellerinden evrimsel sinir ağları olan AlexNet, ResNet-18, GoogleNet ve SqueezeNet mimarileri kullanılmıştır. Her bir mimari için sınıflandırma örnekleri ve eğitim aşaması uygulanmıştır. Yüz tanıma doğruluk oranları gösterilmiştir. Son olarak derin öğrenme modellerinde kullanılan hiper parametreler ve bu parametrelere bağlı olarak elde edilen bulgular gösterilmiştir.

Viola-Jones algoritması ile elde edilen yüz tespitindeki doğruluk oranları Tablo 1'de gösterilmiştir.

Tablo 1. Viola-Jones algoritmasının performansı

	Toplam yüz resmi	Tespit edilen yüz resmi	Tespit edilemeyen yüz resmi	Doğruluk oranı (%)
Viola-Jones	140	120	20	85,71
Viola-Jones	546	462	84	84,61

Viola-Jones algoritmasında toplam 686 görüntü içinde tespit edilen yüz resimleri 582, tespit edilemeyen yüz resimleri 104 tanedir. Veri setindeki en yüksek doğruluk oranı %85,71 olarak bulunmuştur.

3.1 Yüz Tanıma Yöntemleriyle İlgili Bulgular

Yüz tanıma yöntemlerinde kullanılan Eigenfaces ve Fisherfaces algoritmaları için sağ, sol ve orta poz olmak üzere üç farklı görüntü incelenmiştir. Veri setindeki görüntülerin hepsi test ve eğitim klasörü olarak ikiye ayrılmıştır. Eigenfaces algoritmasındaki doğruluk oranları Tablo 2'de gösterilmiştir.

Tablo 2. Eigenfaces ile yüz tanıma doğruluk oranı

Eigenfaces	Sağ poz	Sol poz	Orta poz
Toplam yüz görüntüsü	120	120	120
Toplam test görüntüsü	30	30	30
Toplam eğitim görüntüsü	90	90	90
Doğru yüz görüntüsü	13	6	23
Yanlış yüz görüntüsü	17	24	7
Doğruluk oranı (%)	43,33	20	76,66

Tablo 2'de elde edilen bulgularda en fazla yanlış bulunan yüz görüntüsü sol pozda, en fazla doğru bulunan yüz görüntüsü orta pozda olmuştur. Doğruluk oranları kıyaslandığında, en yüksek değer %76,66 bulunarak orta pozda gerçekleşmiştir.

Fisherfaces algoritmasında görüntüler sağ, sol ve ortadan olmak üzere üç ayrı kategoride incelenmiştir. Elde edilen yüz tanıma performansları Tablo 3'te gösterilmiştir.

Tablo 3. Fisherfaces ile yüz tanıma doğruluk oranı

Fisherfaces	Sağ poz	Sol poz	Orta poz
Toplam yüz görüntüsü	120	120	120
Toplam test görüntüsü	30	30	30
Toplam eğitim görüntüsü	90	90	90
Doğru yüz görüntüsü	18	6	22
Yanlış yüz görüntüsü	12	24	8
Doğruluk oranı (%)	60	20	73,33

Orta pozda en yüksek doğru yüz görüntüsü elde edilmiştir. Sol poz için en düşük doğruluk oranı %20 olduğu gözlemlenmiştir. Doğruluk oranlarında en yüksek başarımlar %73,33 olarak orta pozda bulunmuştur.

3.2 Derin Öğrenme Mimarileriyle İlgili Bulgular

Derin öğrenme mimarileriyle yüz tanıma doğruluk oranlarını artırabilmek için epoch sayı aralıkları ile çalışmalar yapılmıştır. AlexNet mimarisine yüz tanıma metodunda 3-10 epoch sayı aralığında denemeler yapılmıştır. Yapılan çalışmada elde edilen bulgular Tablo 4'te gösterilmiştir.

Tablo 4. AlexNet ile yüz tanıma için epoch sayıları ve doğruluk oranları

AlexNet	Görüntü Sayısı				
Epoch Değeri	120	462	30 sağ	30 sol	30 orta
3	77,78	76,45	87,27	88,46	82,69
4	83,95	81,61	90,91	88,46	88,46
5	91,36	87,10	90,91	92,31	86,54
6	83,95	88,06	94,55	90,38	92,31
7	86,42	89,68	92,73	90,38	90,38
8	79,01	92,26	90,91	92,31	88,46
9	90,12	92,90	90,91	88,46	92,31
10	90,12	92,58	96,36	94,23	96,15

AlexNet mimarisi doğruluk oranlarına göre; 120 görüntü için epoch değeri 5, 462 görüntü için 8, 30 sağ görüntü için 10, 30 sol görüntü için 10 ve 30 orta görüntü için 10 olarak belirlenmiştir. En yüksek doğruluk oranı %96,36 olarak sağ görüntüde bulunmuştur.

ResNet-18 mimarisine yüz tanıma daha iyi bulgular elde edebilmek için epoch sayı aralıkları ile denemeler yapılmıştır. Bulunan doğruluk oranları Tablo 5'te gösterilmiştir.

Tablo 5. ResNet-18 ile yüz tanıma için epoch sayıları ve doğruluk oranları

ResNet-18	Görüntü Sayısı				
Epoch Değeri	120	462	30 sağ	30 sol	30 orta
3	86,42	91,61	84,62	84,62	86,54
4	93,83	72,26	86,54	86,54	96,15
5	88,89	81,29	92,31	88,46	86,54
6	96,30	88,39	90,38	92,31	92,31
7	98,77	92,26	98,08	92,31	96,15
8	96,30	92,90	94,23	94,23	100
9	97,53	98,71	94,23	94,23	98,08
10	97,53	95,81	92,31	92,31	96,15

Derin öğrenme mimarileri için oluşturulan 462 görüntü ve yüz tanıma yöntemleri için oluşturulan 120 görüntünün tamamı ve bu görüntülerin pozlara göre üç farklı görüntü kümeleri kullanılmıştır. Epoch değer aralığında 8 olarak belirlenen denemede en yüksek doğruluk oranı orta pozda %100 olarak gözlemlenmiştir.

GoogleNet mimarisine yüz tanımanın sağlanabilmesi için epoch değerleri ile denemeler yapıldığında bulunan doğruluk oranları için bulgular aşağıdaki Tablo 6'daki gibidir.

Tablo 6. GoogleNet ile yüz tanıma için epoch sayıları ve doğruluk oranları

GoogleNet	Görüntü Sayısı				
Epoch Değeri	120	462	30 sağ	30 sol	30 orta
3	71,60	69,03	80,77	84,62	84,62
4	71,60	76,77	88,46	88,46	86,54
5	76,54	75,81	88,46	88,46	82,69
6	79,01	86,15	88,46	88,46	84,62
7	86,42	87,74	88,54	88,46	86,54
8	86,42	86,45	88,46	92,31	90,38
9	90,12	90,97	94,23	84,62	92,31
10	93,83	89,35	94,23	94,23	90,38

Verilen değer aralıklarına göre en yüksek doğruluk %94,23 olarak sağ ve sol pozda, 9 ve 10 epoch değerlerinde bulunmuştur.

Derin öğrenme mimarilerinden SqueezeNet ile yüz tanıma için epoch değerlerinde elde edilen bulgular Tablo 7'de verilmiştir.

Tablo 7. SqueezeNet ile yüz tanıma için epoch sayıları ve doğruluk oranları

SqueezeNet	Görüntü Sayısı				
	120	462	30	30	30
Epoch Değeri			sağ	sol	orta
3	61,73	63,87	83,64	84,62	82,69
4	69,14	75,48	83,64	84,62	84,62
5	75,31	79,35	83,64	86,54	82,69
6	80,25	82,58	85,45	86,54	86,54
7	81,48	88,39	87,27	92,31	90,38
8	77,78	88,06	89,09	92,31	88,46
9	91,36	90	87,27	82,69	84,62
10	97,53	85,16	94,55	84,62	84,62

SqueezeNet mimarisiyle yüz tanımda epoch sayılarının seçilmesinde 120 görüntüde 10, 462 görüntüde 9, sağda 10, solda 7 ve orta pozda 7 değerlerinde başarı oranlarının arttığı görülmüştür. En yüksek doğruluk oranı %97,53 olarak bulunmuştur.

Derin öğrenme mimarileri için Tablo 4, Tablo 5, Tablo 6 ve Tablo 7'de verilen doğruluk değerleri veri kümesinin %75 eğitim ve %25 test olarak ayrılması ile yapılan eğitim sonucunda, test veri kümesindeki performanslarını göstermektedir.

Tablo 4, Tablo 5, Tablo 6 ve Tablo 7'ye göre her bir derin öğrenme modeli için en yüksek doğruluk değerinin elde edildiği maksimum epoch değeri kullanılarak, derin öğrenme modelleri üzerinde 5 katlı çapraz doğrulama uygulanmıştır. Çapraz doğrulamada, derin öğrenme mimarileri için kullanılan yığın boyutu (mini batch size), maksimum epoch değeri (max epoch) ve başlangıç öğrenme katsayısı (initial learning rate) değerleri Tablo 8'de verilmiştir.

Tablo 8. Derin öğrenme mimarilerinde kullanılan hiper parametreler

Mimari Modeli	Yığın Boyutu	Maksimum Epoch	Öğrenme Katsayısı
AlexNet	15	10	0,0001
ResNet-18	10	3	0,0001
GoogleNet	15	9	0,0003
SqueezeNet	15	5	0,0003

Tablo 8'de verilen yığın boyutu ve başlangıç öğrenme katsayısı değerleri, aynı zamanda veri kümesinin %75 eğitim ve %25 test şeklinde ayrımı ile yapılan performans değerlendirmesinde de kullanılmıştır.

Tablo 9. Derin öğrenme mimarileri için çapraz doğrulama sonuçları

Mimari Model	Görüntü Sayısı				
	120	462	30	30	30
Çapraz Doğrulama			sağ	sol	orta
AlexNet	73,6	84,6	73,8	67,8	60,8
ResNet-18	96,8	97	84,4	73,2	78,2
GoogleNet	84,4	86,6	62,8	60,8	64
SqueezeNet	88,8	93,6	76,6	58,8	60

Tablo 9'a göre en yüksek doğruluk değeri tüm mimariler için 462 resmin bulunduğu veri kümesi ile elde edilmiştir. AlexNet için en yüksek doğruluk değeri %84,6. ResNet için %97, GoogleNet için %86,6 ve SqueezeNet için %93,6 olarak bulunmuştur.

Bu çalışmada bulunan sonuçların, literatürde FEI veri kümesi kullanılarak yapılan çalışmalarla karşılaştırılması Tablo 10'da verilmiştir.

Tablo 10. FEI veri kümesiyle yapılan çalışmalar

Referans	Mimari Model	Doğruluk Oranı
Almabdy ve Elrefaei (2019)	AlexNet+SVM	97,50
Curtidor vd., (2021)	RLD	93,57
Hassan vd., (2021)	LDA	97,84
Win vd., (2021)	MTCNN	86,80
Ayata ve Çavuş (2022)	ESA	98,86
Yapılan Çalışma	ResNet-18	97

Bu çalışmada kullanılan Resnet-18, Curtidor vd. ile Win vd.'nin çalışmalarından daha iyi sonuç vermiştir. Almabdy ve Elrefaei ile Hassan vd.'nin çalışmaları ile hemen hemen aynı sonucu vermiştir. Ancak Ayakta ve Çavuş'un çalışmasından biraz daha düşük performanstadır.

4. SONUÇLAR

Bu çalışmada, yüz tanıma algoritması görüntü tabanlı yöntemler ve derin öğrenme mimarileri kullanılarak eğitilmiş, doğruluk oranları kıyaslanmıştır. Yüz tanımanın ön adımında kullanılan yüz algılama için Viola-Jones algoritması tercih edilmiştir. Kare içinde bulunan yüz görüntüleri veri tabanı olarak kaydedilmiştir.

Yüz tanıma teknikleri içerisinde Temel Bileşen Analizi ile Eigenfaces yöntemi ve Doğrusal Ayrım Analizi ile Fisherfaces yöntemi kullanılmıştır. Öncelikle Eigenfaces algoritması kullanılarak sağ, sol ve ön olmak üzere üç farklı yüz görüntüsü

incelenmiştir. Test görüntüsünden seçilen resim bu metotlar aracılığıyla eğitime giren resimlerle karşılaştırılmıştır. Aynı kişiye ait olan görüntüler doğru, farklı kişiye ait olan görüntü eşleşmeleri yanlış olarak belirlenmiştir. Ön pozisyonlarda yüz tanıma sıklığının arttığı gözlemlenmiştir. Bununla birlikte, Fisherfaces algoritması incelendiğinde ön pozun yüz tanıma oranı en iyi sonucu vermiştir. İki algoritma birbiriyle kıyaslandığında, Eigenfaces algoritması %76,66 yüksek doğruluk oranına sahiptir.

Derin öğrenme yöntemlerinden AlexNet, ResNet-18, GoogleNet ve SqueezeNet mimarileri kullanılmıştır. Görüntüler sağ, sol, ön ve hepsi olarak dört farklı kategoride incelenmiştir. İlk olarak veri kümesi %75 eğitim ve %25 test olarak ayrılarak derin öğrenme yöntemlerinin performansları değerlendirilmiştir. En yüksek doğruluk oranını veren epoch sayısının belirlenebilmesi için 3-10 aralığında epoch değerleri için denemeler yapılmıştır. Elde edilen sonuçlarda AlexNet'te 10 epoch, ResNet-18'de 8 epoch, GoogleNet'te 9 ve 10 epoch değerlerinde, SqueezeNet'te 10 epoch değerlerinde en yüksek doğruluk oranları elde edilmiştir. Mimariler birbirleriyle karşılaştırıldığında en yüksek başarımlar ResNet-18 için %100 olarak bulunmuştur.

Ayrıca, derin öğrenme mimarileri üzerinde 5 katlı çapraz doğrulama uygulanmıştır. Elde edilen sonuçlara göre, en yüksek doğruluk oranı ResNet-18 için %97 olarak elde edilmiştir.

120, 462, 30 sağ, 30 sol ve 30 orta görüntü içerisinden en düşük performans Eigenfaces ve Fisherfaces için 30 sol pozda olmuştur. Görüntülerdeki sağ sol poz oranlarının doğruluk sayısını etkilediği gözlemlenmiştir. Derin öğrenme mimarilerinde ise AlexNet, ResNet-18 ve SqueezeNet'te 30 sol pozda en düşük sonuçlar elde edilmiştir. GoogleNet mimarisinde ise 30 sağ pozda en düşük sonucu vermiştir. Görüntü sayısı arttıkça ve epoch değeri düştükçe elde edilen doğruluk oranlarının azaldığı görülmüştür. Çapraz doğrulama için en düşük sonuç 30 sol pozda çıkmıştır. Doğruluğun artırılması ve görüntülerdeki parametre değerlerine bağlı değişikliklerin en iyi performansı vermesi için daha farklı çalışmalar yapılabilir.

Sonuç olarak derin öğrenme mimarilerinin, görüntü işleme tabanlı yöntemlere kıyasla daha iyi sonuçlar verdiği gözlemlenmiştir. Görüntü işleme alanında çok popüler olan bu yöntemlerin gelecekte daha iyi sonuçlar vereceğine inanılmaktadır.

KAYNAKÇA

Aly M., Face Recognition using SIFT Features, CNS 186 Term Project Winter, 2006.

Almabdy S., ve Elrefaei L., Deep Convolutional Neural Network-Based Approaches for Face Recognition, *Applied Sciences*, 9(20), 4397, 2019.

Arya Z., ve Twiari, V. (2020). Automatic Face Recognition and Detection Using OpenCV, Haar Cascade and Recognizer at Different Angle of Face, *International Journal of Engineering Research and Applications*, 10(6), 2020.

Ayata F., ve Çavuş H., (2022). Yüz Tanıma Sistemlerinde Kullanılan ESA, YGH-DVM ve DSA Algoritmalarının Performans Testleri, *Fırat Üniversitesi Fen Bilimleri Dergisi*, 34(1), 39-48.

Balanageshwara, S., Kareem, A., ve Kumara, V. (2023, June). Machine Learning Approach for a Novel Facial Recognition System. In *2023 8th International Conference on Communication and Electronics Systems (ICCES)* (pp. 1178-1183). IEEE.

Cadena, J., Villa, M., Martínez, M., Acurio, J., ve Chacón, L. (2023). An Efficient Technique for Global Facial Recognition using Python and OpenCV in 2D Images. *WSEAS Transactions on Systems and Control*, 18, 47-57.

Curtidor A., Baydyk T., ve Kussul E., (2021). Analysis of Random Local Descriptors in Face Recognition, *Electronics*, 10(11), 1358. <https://doi.org/10.3390/electronics10111358>

Çakmakoğlu A., (2018). *Finans Sektöründe Veri Madenciliği Teknikleri Kullanılarak Kampanya Modelleme*, Yüksek Lisans Tezi, İstanbul Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

Doğan F., & Türkoğlu İ., Derin Öğrenme Modelleri ve Uygulama Alanlarına İlişkin Bir Derleme, *DÜMF Mühendislik Dergisi*, 10(2), 419-424, 2019.

Fırıldak K. ve Talu F. M., (2019). Evrimsel Sinir Ağlarında Kullanılan Transfer Öğrenme Yaklaşımlarının İncelenmesi, *Computer Science*, 4(2), 88-95.

Hassan M. M., Hussein H. I., Eesa A. S., ve Mstafa R. J., (2021). Face redognition based on gabor feature extraction followed by fastica and LDA, *Computers, Materials & Continua*, 68(2), 1637-1659. <https://doi.org/10.32604/cmc.2021.016467>

Holat R., ve Kulaç S., (2014). *Yüz Bulma ve Tanıma Sistemleri Kullanarak Kimlik Tespitinin Yapılması*, Yüksek Lisans Tezi, Düzce Üniversitesi, Fen Bilimleri Enstitüsü, Düzce.

Iandola F. N., Han S., Moskewicz M. W., Ashraf K., Dally W. J., ve Keutzer K., SqueezeNet: AlexNet-Level Accuracy with 50x Fewer Parameters and <1MB Model Size, ArXiv, 1-5, 2016.

- ImageNet Large Scale Visual Recognition Challenge 2014 (ILSVRC2014), 14 Ağustos 2024 tarihinde <https://image-net.org/challenges/LSVRC/2014/results> adresinden erişildi.
- Jones, M., ve Viola, P. (2003). Fast Multi-view Face Detection, Mitsubishi Electric Research Laboratories, TR-20003-96.
- Kekül, H., Bircan, H., ve Arslan, H. (2018). Yüz Tanıma Uygulamalarında Özyüzler ve Yapay Sinir Ağlarının Karşılaştırılması, *Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi*, 2(1), 52.
- Krizhevsky A., Sutskever I., ve Hinton G. E., (2012). ImageNet Classification with Deep Convolutional Neural Networks, *Advances in Neural Information Processing Systems*, 25(2),1097-1105.
- Raghu S., Sriraam N., Temel Y., Rao S. V., ve Kubben P. L., EEG based multi-class seizure type classification using convolutional neural network and transfer learning, *Neural Networks: the Official Journal of the International Neural Network Society*, 206, 2020.
- Rahmad C., Asmara R. A., Putra D. R. H., Dharma I, Darmono H., ve Muhiqqin I., (2020). Comparison of Viola-Jones Haar Cascade Classifier and Histogram of Oriented Gradients (HOG) for face detection, *IOP Conference Series: Materials Science and Engineering*, 732(1), 012038.
- Sarwinda D., Paradisa R. H., Bustamam A., ve Anggia P., (2021). Deep Learning in Image Classification using Residual Network (ResNet) Variants for Detection of Colorectal Cancer, *Procedia Computer Science*, 423-431(179), 1877-0509.
- Sert Z., (2020, 27 Aralık). ESA (Evrşimsel Sinir Ağları) [Blog yazısı]. 13 Ağustos 2024 tarihinde erişim adresi: <https://zeysert.medium.com/esa-evri%C5%9Fimsel-sinir-a%C4%9Flar%C4%B1-87d9bd986579>
- Sütçüler E., (2006). *Gerçek Zamanlı Video Görüntülerinden Yüz Bulma ve Tanıma Sistemi*, Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.
- Şeker A., (2018). Evaluation of Fabric Defect Detection Based on Transfer Learning with Pre-trained AlexNet, *2018 International Conference on Artificial Intelligence and Data Processing (IDAP)*, Malatya, Turkey, pp. 1-4.
- Thomaz C. E., (2012). *FEI Face Database*, 13 Ağustos 2024 tarihinde erişim adresi: <https://fei.edu.br/~cet/facedatabase.html>
- Topal A. O., Chitic R., ve Leprévost F. (2023). One evolutionary algorithm deceives humans and ten convolutional neural networks trained on ImageNet at image recognition, *Applied Soft Computing*, 143, 110397. doi: <https://doi.org/10.1016/j.asoc.2023.110397>
- Wang J. ve Li Z., (2018). Research on Face Recognition Based on CNN, *IOP Conference Series: Earth and Environmental Science*, 170(3), 032110.
- Wang D., Yu H., Wang D. ve Li G., (2020). Face Recognition System Based on CNN, *2020 International Conference on Computer Information and Big Data Applications (CIBDA)*, 470-473.
- Win H. P. P., Khine P. T. T., ve Tun K. N. N., (2021). Face Recognition System based on Convolution Neural Networks, *International Journal of Image Graphics and Signal Processing (IJIGSP)*, 13(6), 23-29.