



# Derin Öğrenme ile Yüz Tanıma ve Duygu Analizi

Yaşar Safalı<sup>1\*</sup>, Erdinç Avaroğlu<sup>2</sup>

<sup>1\*</sup> Mersin Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Mersin, Türkiye, (ORCID: 0000-0001-9717-9892), [yasarsafali.01@gmail.com](mailto:yasarsafali.01@gmail.com)

<sup>2</sup> Mersin Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Mersin, Türkiye (ORCID: 0000-0003-1976-2526), [eavaroğlu@mersin.edu.tr](mailto:eavaroğlu@mersin.edu.tr)

(İlk Geliş Tarihi 15 Ekim 2021 ve Kabul Tarihi 10 Aralık 2021)

(DOI: 10.31590/ejosat.1010450)

**ATIF/REFERENCE:** Safalı, Y. & Avaroğlu, E. (2021). Derin Öğrenme ile Yüz Tanıma ve Duygu Analizi. *Avrupa Bilim ve Teknoloji Dergisi*, (31), 764-770.

## Öz

Kişilerin davranışlarına, fiziksel özelliklerine bağlı olarak geliştirilen biyometrik sistemler son yıllarda aktif olarak kullanılmaktadır. Kişinin benzersiz özelliklerine dayanan biyometrik sistemler içerisinde yüz tanıma fiziksel temasa gerek duymaması sebebi ile önemli bir yer kaplamaktadır. Bu çalışmada derin öğrenme tabanlı yüz tanıma ve yüz ifadesi tanıma uygulaması gerçekleştirilmiştir. VGG-16, AlexNet ve ZF Net mimarileri ile geliştirilen modeller eğitilerek başarı oranları karşılaştırılmıştır. En başarılı model %92,03 başarı oranı ile VGG-16 mimarisi referans alınarak geliştirilen model olmuştur.

**Anahtar Kelimeler:** Yüz tanıma, Yüz ifadesi tanıma, derin öğrenme, Biyometrik sistemler.

## Face Recognition and Emotion Analysis with Deep Learning

### Abstract

Biometric systems developed depending on the behavior and physical characteristics of individuals have been actively used in recent years. Facial recognition occupies an important place among biometric systems based on the unique characteristics of the person, as it does not require physical contact. In this study, facial recognition and facial expression recognition based on deep learning were implemented. Models developed with VGG-16, AlexNet and ZF Net architectures were trained and their success rates were compared. The most successful model was the model developed based on VGG-16 architecture with a success rate of 92.03%.

**Keywords:** Face recognition, facial expression recognition, deep learning, Biometric systems.

\* Sorumlu Yazar: [yasarsafali.01@gmail.com](mailto:yasarsafali.01@gmail.com)

## 1. Giriş

İnsan nüfusunun artması ile birlikte kullanılan mobil cihazların sayısında artış görülmektedir. Mobil cihazlara ek olarak dünya genelinde üretilen kamera sayısı da her geçen gün artmaktadır. Kamera ve mobil cihaz sayısının artmasına paralel olarak insanların kameralara verdiği görüntü sayısı artmaktadır. Büyük miktarda oluşan bu görüntüler araştırmacılara bilgisayarlı görü alanında çalışma yapabilecekleri veri tabanları oluşturmuştur.

İnsanların sürekli birbiri ile etkileşim içerisinde olması, birçok sektörün dijital ortamda aktif olarak faaliyet göstermesi, bazı kamu ve özel kurumlara yetkisiz girişlerin engellenmesi istenmesi ortaya güvenlik sorununu çıkarmıştır. Özellikle mobil cihazlarda insanların kişisel verilerinin yer alması, klasik güvenlik önlemlerine ek yeni kimlik doğrulama yöntemlerinin ortaya çıkmasına sebep olmuştur. Parola, PIN (personal identification number) gibi daha önceden kullanılan klasik güvenlik önlemlerinin çalınabilmesi, hatırlanamama gibi dezavantajları vardır. Bu dezavantajlar, kişiye özel kimlik doğrulama yöntemleri olan biyometrik sistemlerin gelişmesine katkı sağlamıştır.

Teknolojinin gelişimine bağlı olarak hızla gelişen biyometrik sistemler, kişilerin fiziksel, kimyasal ve davranışsal özelliklerine dayalı bir tanımlama bilimidir (Vatsa, Singh, and Majumdar 2018). Biyometrik sistemlerin ortak özelliği kişinin benzersiz özellikleri üzerine inşa edilmesidir. Bu benzersiz özellikler sayesinde kişinin kimlik doğrulaması yapılmaktadır. Kimlik doğrulama işlemi yapılırken kişiye ait parmak izi, yüz tanıma, el yazısı, kişinin imzası, yürüyüş, iris vb.. özelliklerinden yararlanan biyometrik sistemler mevcuttur.

Yüz tanıma, dijital görüntüler üzerinde yer alan nesnelerin arasından insan yüzünün konumunu ve boyutunu belirleyen bir bilgisayar teknolojisidir. Bilgisayarlı görü alanında ortaya çıkan yüz tanıma, insanların büyük miktardaki verilerde bilgisayarlar kadar hızlı çalışmama problemlerine çözüm aramaktadır. Bilgisayarlar yüksek hızda, yüksek bellek kullanarak insanlardan daha hızlı hesaplama yapabilmektedir.

Yüz tanıma uygulamaları, kendisine giriş olarak verilen görüntülerde ilk olarak yüzü tespit etmeye çalışır. Giriş verisi olarak verilen görüntüde yüz olup olmadığı doğru bir şekilde tespit edilebilirse ilk aşama başarılı olarak tamamlanmıştır. Eğer görüntüde yüz tespit edilirse daha sonra özellik çıkarımı yapılarak yüzün kime ait olduğu tespit edilir.

Yüz tanıma sistemleri, güvenlik amacı ile kamu kurumlarında kullanılmaktadır. Yüksek güvenlik gerektiren kamu kurumlarına yetkisiz girişlerini engellemek için kullanılmaktadır. Bunun yanı sıra özel sektör ve kamu kuruluşlarında personel takibi için kullanılmaktadır. Yüz tanıma sistemleri elektronik bankacılık, mobil cihazların aktivasyonu gibi dijital ortamlarda da aktif olarak kullanılmaktadır.

Duygu analizi, duyguların tanınması bilgisayar biliminde popüler olarak çalışma yapılan bir alandır. İnsan ve bilgisayar arasındaki etkileşimin daha doğal olması için bilgisayarların insan duygularını tanıması amaçlanmıştır. Duygu analizi çoğunlukla metin madenciliği alanında kullanılsa da son zamanlarda bilgisayarlı görü alanında da çalışmalara konu olmuştur. Metin madenciliği alanında yapılan çalışmalarda genel olarak kişinin bir konu hakkındaki olumlu, olumsuz ve nötr olmak üzere üç duygusu referans alınır. Metin madenciliği alanında yapılan

duygu analizleri reklam, ürün pazarlama, güncel olaylar vb. alanlarda kısa ve anlamlı bilgiler çıkarmak için yapılmaktadır. Bilgisayarlı görü alanında yapılan duygu analizi çalışmaları Otomatik Yüz İfade Tanıma kısaca FER (Face Expression Recognition) olarak adlandırılırlar.

Yüz ifadesi tanıması, yüzdeki belirgin özelliklerden yararlanır. Bu özellikler kaşların yer aldığı bölge, ağız ve dudak bölgesi, göz ve göz kapaklarıdır. İnsanların temel olarak ifade ettiği 6 duygu mevcuttur. Bunlar mutluluk, üzüntü, şaşırma, korku, kızgınlık ve iğrenmedir (Salunke and Patil 2018). Yüz ifadesi tanıma uygulamaları son zamanlarda güvenlik güçleri tarafından sürveyans ve davranış analizi alanında kullanılmaya başlanmıştır. Ayrıca dijital kameralarda gülümseme algılamasında otomatik olarak fotoğraf çekmek için bu tür teknikler kullanılmaktadır (Salunke and Patil 2018).

Literatürde yüz tanıma sistemleri geliştirilirken kullanılan birçok teknik mevcuttur. Bu tekniklerden bir kısmı kendisine verilen görüntüde insan yüzüne ait kaş, göz, burun, ağız gibi özellikler aramaktadır. Bir kısmı ise kendisine verilen görüntüde yüzün özelliklerini aramaktansa görüntünün tamamı üzerinde işlem yapmaktadır. Bu teknikler genel olarak özellik tabanlı ve bütüncül yaklaşımlar olarak gruplandırılmıştır (Kumar, Kaur, and Kumar 2019).

Özellik tabanlı yaklaşımlar insan yüzünde yer alan burun, göz, kulak gibi yüz özelliklerini referans olarak çalışmaktadırlar. Genel olarak Aktif şekil modeli, düşük seviye analiz, özellik analizi başlıklarında yapılan çalışmalardır.

Wiskott ve arkadaşları yapmış oldukları çalışmada (Wiskott et al. 1997) giriş yüzü grafiği olarak ifade edilen prob setini, model yüzü grafiği olarak temsil edilen galeri setiyle eşleştirerek yüzleri tanyan Elastik Demet Grafiği Eşleştirme algoritmasını geliştirmişlerdir. Temel olarak giriş yüzü grafiğinin her düğümü yüzün belirli bir özellik noktasıyla temsil edilir. Örneğin, bir düğüm bir gözü temsil eder ve diğer bir düğüm burnu temsil eder. Bu mantıkla düğümler diğer yüz özelliklerini temsil etmeye devam eder. Düğümler yüzün şekline göre getirilmiş bir graph yapısı oluşturularak birbirine bağlanır (Wiskott et al. 1997).

Anila ve Devarajan 2010 yılında yaptıkları çalışmada (Satish et al. 2010) üç adımdan oluşan bir yüz tanıma yaklaşımı önermişlerdir. İlk adımda görüntülerde yer alan görüntüleri gidermek için medyan filtre uygulanmaktadır. Daha sonra görüntü üzerinde kontrast ayarı yapmak için histogram dengeleme uygulamışlardır. İkinci adımda kenar görüntüleri için Sobel kenar çıkarımı algoritmasını görüntüye uygulamışlardır. Son adımda ise görüntü üzerinde yer alan yüz bölgesini ve diğer arka planları çıkarmak için Geri Yayılım sınır ağı algoritmasını kullanmışlardır.

Viola ve Jones, hesaplama süresini aza indiren, yüksek algılama doğruluğu elde eden bir yaklaşım sundular (Zhang et al. n.d.). Haar benzeri özelliklerin kullanılmasına dayanan bu yaklaşım hızlı sonuç vermektedir. Bu yaklaşım integral görüntü kavramına dayanarak geniş bir dizi özellik üretir ve aşırı kümelenmeyi azaltmak için artırma algoritması AdaBoost'u kullanır. Gri tonlamalı görüntülerde uygulanan AdaBoost dedektör tarama biçiminde çerçevelerle uygulanır.

Kalocsai ve arkadaşları araştırmalarında (Kalocsai et al. n.d.) yüz tanıma özelliğini iyileştirmek için yüzlerin 48 özellik noktasından 40 Gabor çekirdeğinin üretildiği Gabor çekirdeklerinin ağırlıklandırılmasının etkisini araştırdılar. Kafkas yüzleri veri setini kullanarak en ayırt edici yüz özelliklerinin alın

ve gözlerin etrafına yerleştirildiğini buldular. Buna karşılık, en az ayırt edici yüz özellikleri ağız, burun, yanaklar ve yüzün alt ana hattı olduğunu tespit ettiler. En yüksek ağırlıklı çekirdeklerin, yüzlerin daha kompakt bir gösterimini sağlayacağı ve en düşük ağırlıklı çekirdeklere kıyasla en yüksek ağırlıklı çekirdekleri kullanarak daha yüksek tanıma oranları sağlayacağı sonucuna vardılar.

Hjelmas, güçlü yüz tanıma özelliği için geliştirdiği Gabor özelliklerini tanıttı.(Hjelmås n.d.) Hjelmås'a göre, yüz görüntüsünün işlenmesi için (eğitim veya test için), görüntü bir dizi Gabor filtresi ile filtrelenmiştir. Görüntüyü yüzün ortasına odaklamak için 2D Gaussian uygulamıştır. Gabor Filtreli ve Gauss ağırlıklı görüntüde daha sonra yüz tanıma için ilginç özellik noktaları olarak nitelendirilen zirveler aranır. Her zirvede, Habor katsayılarından oluşan bir özellik vektörü çıkarılarak konum ve sınıf etiketi saklanır.

Bütüncül yaklaşımlar görüntü üzerinde insan yüzünün özellikleri aramaktansa görüntünün tamamına dayalı analiz yaparlar. Bütüncül yaklaşımlar, tüm yüz bölgesini bir tanıma sisteminde ham girdi olarak kullanır. Yüz görüntülerinin piksel yoğunluğu değişimlerini istatistiksel olarak hesaplar ve veri tabanındaki resimlerin yoğunluk değerleri ile karşılaştırır. Genel olarak bütüncül yaklaşımlar istatistiksel modelleri ve yapay sinir ağlarını kullanırlar.

Dürr ve arkadaşları Raspberry Pi kullanarak elde ettikleri gerçek zamanlı görüntülerde yüz tanıma gerçekleştirmişlerdir (Dürr et al. 2015). CNN kullanarak inşa ettikleri model ile OpenCV yaklaşımlarını karşılaştırmışlardır. Raspberry Pi kullanarak elde ettikleri yüksek çözünürlüklü görüntüleri 8 bit gri ölçeğe dönüştürmüşlerdir. Daha sonra gri görüntülerin tamamını 640 x 480 piksele yeniden boyutlandırmışlardır. OpenCV içerisinde yer alan Viola-Jones algoritmasını kullanarak yüz tanıma modeli geliştirmişlerdir. CNN kullanarak inşa ettikleri diğer modelde giriş görüntülerini 48 x 48 piksel olarak yeniden boyutlandırmışlardır. Görüntüyü CNN'e beslemeden önce yerel ikili patern (LBP) operatörünü kullanarak 46 x 46 piksel görüntüye çevirmişlerdir. Bu sayede her pikseli kodlayarak sistemin aydınlatma bağımlılığını azaltmışlardır. Veri kümesi üzerinde iki modeli de test ettiklerinde CNN kullanan model saniyede 2 kare performansa ve %97 başarı oranına ulaşmıştır.

Yüz ifadesi tanıma yaklaşımlarında DNN'ler (Derin Sinir Ağları) son zamanlarda öne çıkmaktadır. DNN, büyük miktardaki verilerde başarılı performanslar ortaya koyan sinir ağı mimarileridir. Bilgi işlem gücü gün geçtikçe büyüdükçe, DNN mimarileri metin, resim ve seste karmaşık desenler bulabilen beyin benzeri yapıların geliştirilmesine dayanan bir öğrenme mimarisi sağlamaktadır. Bu nedenle, FER çalışmalarında araştırmacılar derin sinir ağlarını eğiterek özellik tanıma ve çıkarma konusunda yüz ifadesi tanıma sistemini daha kısa sürede tasarlayabilir (Salunke and Patil 2018).

Afshin Dehghan ve arkadaşları (Dehghan et al. n.d.), 2017 yılında yaptıkları çalışmada kişilerin cinsiyetlerini ve yüz ifadelerini tespit eden derin öğrenme tabanlı, derin sinir ağlarına sahip bir boru hattı sistemi tasarlamışlardır. Sistem tasarlanırken 40.000'den farklı kişiye ait toplamda 4.0000.000 görüntü içeren veri seti kullanılmıştır. Yüz ifadesi tanıma işlemi için FER2013 veri setinde yer alan görüntülerden yararlanılmıştır. Veri seti oluşturulurken kullanılan görüntüler yarı denetimli bir şekilde cinsiyet ve duygu ifadelerine göre etiketlenmiştir. Etiketlenen veriler üzerinde yüz hizalaması gerçekleştirilmiştir. Sistemin

başarı oranları ise yaş tahmininde %61,3 olurken yüz ifadesi tanımadada %76,1 olmuştur.

Bu çalışmada derin öğrenme teknikleri kullanılarak yüz tanıma ve yüz üzerinde duygu analizi gerçekleştiren yüz ifadesi tanıma sistemi geliştirilmiştir. Çalışma temel olarak veri kümesi oluşturma ve ön işleme, derin öğrenme modelleri oluşturularak yüz tanıma ve duygu analizi, deneysel sonuçlar olmak üzere 4 adımdan oluşmaktadır.

Yüz tanıma ve duygu analizi işlemleri için 2 farklı veri kümesi oluşturulmuştur. Veri kümesi oluşturulurken açık kaynak verilerden yararlanılıp herhangi bir etik dışı yömeme başvurulmamıştır. Yüz tanıma işlemi için veri kümesi oluşturulurken 4 farklı açık kaynak veri kümesi (CelebA, FFHQ, LFW ve Youtube yüzler veri kümesi) kullanılmıştır. Duygu analizi işlemi için FER-2013 veri kümesi kullanılmıştır. Veri kümeleri oluşturulurken farklı boyuttaki ve çözünürlükteki veriler ön işlemlerden geçirilerek standart hale dönüştürülmüştür.

Veri kümesi oluşturulduktan sonra derin öğrenme mimarilerinden AlexNet, Vgg-16 ve ZFNet mimarileri kullanarak yüz tanıma ve duygu analizi gerçekleştirilen modeller tasarlanmıştır. Modellerin başarı oranları karşılaştırılarak en başarılı sonucu veren Vgg-16 modeli ile çalışmanın yapılmasına karar verilmiştir.

Çalışmanın diğer bölümleri şu şekilde organize edilmiştir. İkinci bölümde yüz tanıma ve duygu analizi için geliştirilen modellerin tasarlanması ve tasarım sürecinde kullanılan teknolojiler açıklanmıştır. Üçüncü bölümde elde edilen deneysel sonuçların karşılaştırılması, dördüncü bölümde ise sonuçlara değinilmiştir.

## 2. Materyal ve Metot

### 2.1. Veri Kümesi Oluşturma ve Ön İşleme

Yüz tanıma sistemi tasarlanırken açık kaynak olarak araştırmacıların kullanımına açılmış 4 farklı veri tabanından veri kümesi oluşturulmuştur. Veri kümesi oluşturulurken araştırmacılar tarafından kullanılan CelebA (büyük ölçekli CelebFaces), FFHQ (Flickr-Faces-HQ Veri Kümesi), LFW (Wild Home veri kümesi) ve Youtube Yüzler Veri kümesi kullanılmıştır. Bu veri kümelerinde yer alan görüntülerden toplamda 300.000 görüntü eğitim amaçlı kullanılırken 60.000 görüntüde test amaçlı kullanılmıştır. Veri kümesinin özellikleri Tablo 1'de yer almaktadır.

Tablo 1. Veri setinde kullanılan veri tabanlarının özellikleri

Veri Tabanı	İçerdiği Görüntü Sayısı	Veri Tabanı Boyutu	Kullanılan Görüntü Sayısı
CelebA	200.000	8 GB	180.000
FFHQ	70.000	7 GB	50.000
LFW	13.000	173 MB	10.000
Youtube Yüzler Veri kümesi	155.560	10 GB	120.000

Yüz ifadesi tanıma işlemi için Fer-2013 veri seti kullanılmıştır. Veri seti, sinirli, iğrenme, üzgün, korku, mutlu, şaşkın ve nötr olmak üzere 7 insan duyusundan oluşan görüntüler içermektedir. Fer-2013 veri seti içerisinde 39.892 adet görüntü bulunmaktadır. Bu görüntüler veri seti içerisinde eşit olarak dağıtılmamıştır. Çalışmada yüz ifadesi tanıma modeli geliştirilirken ilk olarak Fer-2013 veri seti üzerinde sistem eğitilmiştir. Daha sonra veri kümesine görüntüler eklenerek veri kümesinin her sınıf için eşit sayıda görüntü içermesi amaçlanmıştır. Veri kümesinin dengeli ve dengesiz olduğu durumlarda sistemin başarı oranları karşılaştırılmıştır. Fer-2013 veri setinin özellikleri Tablo 2’de açıklanmıştır.

Tablo 2. FER-2013 veri seti özellikleri

Duygu Sınıfı	İçerdiği Görüntü Sayısı
İğrenme	1766
Üzgün	6176
Korku	6096
Mutlu	7849
Sinirli	6989
Nötr	7011
Şaşkın	4005

Veri kümeleri oluşturulduktan sonra görüntülerin eğitim aşamasında kullanılabilmesi için görüntüler üzerinde yeniden boyutlandırma ve kırpma işlemleri gerçekleştirilmiştir. Geliştirilecek yüz tanıma modellerinin giriş katmanı 224 x 224 x 1 boyutunda olduğu için görüntüler 224 x 224 piksel olarak yeniden boyutlandırılmıştır. Veri kümeleri hazır hale getirildikten sonra yüz tanıma işlemi için gerekli modeller evrişimsel sinir ağları kullanılarak tasarlanmıştır.

## 2.2. Yüz Tanıma ve Duygu Analizi Modellerinin Tasarlanması

Yüz tanıma işlemi için derin öğrenme tabanlı AlexNet, VGG-16, ZF Net mimarilerini temel alan 3 model geliştirilmiştir. Geliştirilen 3 modelde de evrişimsel sinir ağları kullanılmıştır. Modellerin başarı oranı veri kümesi üzerinde test edilerek en başarılı olan model tespit edilmiştir. Ayrıca yüz ifade analizi için 3 model eğitilerek modellerin başarı oranları karşılaştırılmıştır.

Evrişimsel (konvolüsyonel) sinir ağları, yapay sinir ağlarının çok katmanlı türüdür (Tüfekçi and Fatih Karpat n.d.). Genellikle bilgisayarlı görü alanında yapılan çalışmalarda yüksek başarı oranı yakalayan evrişimsel sinir ağları, görüntü üzerinde nesne tanıma çalışmalarında kullanılmaktadır. Evrişimli sinir ağı temel olarak giriş, konvolüsyon (evrişim), havuzlama, tam bağımlı katman ve sınıflandırma katmanlarından oluşmaktadır. Konvolüsyon ve havuzlama katmanları özellik çıkarımının yapıldığı, tam bağımlı ve sınıflandırma katmanları ise çıktının hangi sınıfa ait olduğunu tahmin eden katmanlardır.

Giriş katmanı, evrişimsel sinir ağına verinin özellik çıkarımı yapılmadan ham olarak verildiği katmandır (Koushik 2016). Giriş verisinin boyutu sinir ağının derinliğini belirler. Giriş boyutunun

yüksek olması sistemin eğitim süresini uzatmaktadır. Ayrıca daha fazla belleğe ihtiyaç duymasına sebep olmaktadır. Giriş boyutunun küçük olduğu sistemlerde ise sınıflandırma performansı düşmektedir (Koushik 2016).

Evrişim katmanı, belirli bir filtrenin görüntü üzerine uygulanması ile aktivasyon haritasının oluşturulduğu katmandır. Aktivasyon haritası, görüntü üzerinde yer alan bölgelerin önemini belirlediği haritalardır. Görüntünün giriş boyutuna göre 1x1, 2x2, 3x3 gibi çeşitli filtreler uygulanmaktadır. RGB formatındaki bir görüntünün derinliği 3 tür. Örnek olarak 5x5x3 boyutundaki bir görüntüye 3x3 boyutunda bir filtre uygulandığında 3x3 boyutunda bir aktivasyon haritası elde edilir. Aktivasyon haritası, görüntünün kırmızı, yeşil ve mavi renk kanallarındaki önem arz eden bölgelerini ve kenar özelliklerini tespit eder. Filtre uygulanırken katsayılar araştırmacılar tarafından belirlenebilir. Bir filtre görüntü üzerinde uygulanırken her indiste yer alan kat sayılar görüntü matrisinde yer alan indisteki değerler ile çarpılarak toplanır (Goodfellow et al. 2016).

Havuzlama katmanı, tasarlanan evrişimsel sinir ağı modellerinde isteğe bağlı olarak kullanılan bir katmandır. Evrişim katmanından sonra kullanılan bu katmanın temel amacı görüntünün genişlik ve yükseklik boyutlarını düşürerek bir sonraki katman için daha küçük boyutta çıktı üretmektir (Pervan 2019). Görüntü boyutunun küçültülmesi ağın ezberleme durumunu engellediği gibi bir sonraki katmanında iş yükünü azaltır (Özkan, Araştırma, and 2017 n.d.).

Havuzlama katmanında genel olarak maksimum havuzlama ve ortalama havuzlama yöntemleri kullanılmaktadır. Maksimum havuzlama yöntemi, görüntüye uygulanan filtrenin uygulandığı indislerdeki maksimum değeri alır. Ortalama havuzlama yöntemi ise filtrenin uygulandığı indislerdeki değerler toplamının ortalamasını alır. Filtreler görüntüye uygulanırken kaç adım gezdirileceğini ifade eden adımlama sayısı araştırmacı tarafından seçilir. Havuzlama katmanında elde edilecek çıktının boyutu katmana girdi olarak verilen görüntünün boyutuna, filtre boyutuna ve adımlama sayısına bağlıdır.

Tam bağımlı katman, kendisinden önce gelen evrişim ve havuzlama katmanları bağı olarak çalışmaktadır. Bu katmanlar sonucu elde edilen çıktılar tam bağımlı katmanda giriş olarak kullanılmaktadır. Havuzlama katmanı sonucunda elde edilen nöronlar tam bağımlı katmanda birbirine bağlanır. Tam bağımlı katmanda ağın ezberleme yapmasını engellemek için DropOut kullanılır. Birbirine bağlanmış nöronlar DropOut yardımı ile birbirinden koparılarak ezberleme yapılmasının önüne geçilir (Özkan, Araştırma, and 2017 n.d.).

Tam bağımlı katmandan sonra gelen sınıflandırma katmanı, sinir ağının son çıktı üreten katmanıdır. Sınıflandırma katmanında elde edilen çıktı sayısı sınıflandırılacak nesne sayısına eşittir. N tane sınıf için sınıflandırma işlemi yapılacaksa n tane 0 ile 1 arasında çıktı üretilir. 1’e en yakın olan çıktı nesnenin ait olduğu sınıf olarak kabul edilir. Sınıflandırma katmanında genel olarak SVM, Softmax, k en yakın komşu algoritması gibi sınıflandırıcılar kullanılmaktadır.

### 2.2.1 Vgg-16 Mimarisi ile Geliştirilen Yüz Yanıma Modeli

VGG-16 modeli 13 evrişim katmanı, 3 tam bağımlı katman olmak üzere toplamda 16 katmandan oluşan bir nesne tanıma mimarisidir (Simonyan and Zisserman 2015). 224 x 224 x 3 boyutundaki görüntüleri giriş olarak almaktadır. Evrişim, havuzlama ve tam bağımlı katmanlardan oluşmaktadır.

Geliştirdiğimiz modelde evrişim katmanlarında 3 x 3 boyutunda filtreler uygulanarak aktivasyon haritaları oluşturulmuştur. Bu aktivasyon haritaları giriş görüntüsü üzerinde ağırlıklı çıkartımı yapmasını sağlamıştır. Havuzlama işlemi yapılırken 2 x 2 boyutundaki filtreler ve maksimum havuzlama yöntemi kullanılmıştır. Modelde gerçekleştirilen havuzlama işlemlerinde adımlama sayısı 2 olarak alınmıştır. Modelde tam bağımlı katmana 7 x 7 x 512 boyutunda giriş görüntüsü verilmiştir. Tam bağımlı katmanda ise bu görüntüye karşılık 25.088 nöron oluşturulmuştur. Tam bağımlı katmanda Dropout fonksiyonu kullanılarak bazı nöronlar serbest bırakılmıştır. Sınıflandırma katmanında ise softmax sınıflandırıcı kullanılarak geliştirilen model veri kümesi üzerinde eğitilmiştir.

### 2.2.2 AlexNet Mimarisi ile Geliştirilen Yüz Yanıma Modeli

Krizhevsky ve arkadaşları tarafından tasarlanan Alexnet modeli, 1000 nesneyi %15,4 hata oranı ile sınıflandırarak 2012 yılında düzenlenen ImageNet yarışmasını kazanmıştır. Alexnet modeli, temel olarak 5 evrişim katmanı, 3 havuzlama katmanı ve 3 tam bağımlı katman içermektedir (Krizhevsky, Sutskever, and Hinton 2017). Tasarlanan sinir ağına 224 x 224 x 3 boyutundaki görüntü giriş olarak verilmiştir.

AlexNet mimarisi referans alınarak geliştirilen model 11 katman olarak tasarlanmıştır. Evrişim katmanlarında 11 x 11, 5x5 ve 3x3 boyutunda filtreler kaydırılarak uygulanmıştır. Havuzlama katmanlarında maksimum havuzlama yöntemi uygulanarak sisteme girdi katmanında verilen 224 x 224 x 3 boyutundaki görüntü 6x6x256 boyutuna düşürülmüştür. Tam bağımlı katmana 9.216 nöron girdi olarak verilirken Dropout kullanılarak 4096 nörona düşürülmüştür. Son olarak sınıflandırma işlemi için softmax sınıflandırıcısı kullanılmıştır.

AlexNet mimarisi kullanılarak geliştirilen bu modelde her evrişim katmanından sonra Relu aktivasyon fonksiyonu kullanılmıştır. Bu fonksiyon katmanlardan çıkan değerlerin negatif değere eşit olmamasını sağlar. Çıktıları sıkıştırarak negatif değerlere sıfır atar.

### 2.2.3 ZFNet Mimarisi ile Geliştirilen Yüz Yanıma Modeli

Zeiler ve arkadaşları tarafından 2014 yılında geliştirilen ZF Net modeli 9 katmandan oluşmaktadır (Zeiler and Fergus 2014). 224 x 224 x 3 boyutundaki renkli görüntüler sinir ağına giriş olarak verilmiştir. Mimaride yer alan 5 adet evrişim katmanında AlexNet den farklı olarak 7 x 7 boyutundaki filtreler 2 adım kaydırılarak uygulanmıştır (Zeiler and Fergus 2014).

ZF Net mimarisi kullanılarak tasarlanan 3. Model AlexNet modeli ile benzerlik göstermektedir. 224x224x3 boyutundaki görüntü 13x13x256 boyutuna düşürülmüştür. Havuzlama yöntemi olarak maksimum havuzlama yöntemi kullanılmıştır. Tam bağımlı katmanda üretilen 43.264 adet nöronlardan 4096 tanesi tam bağımlı katmanda işleme tabi tutulmuştur. ZF Net tabanlı modelde sınıflandırıcı olarak softmax fonksiyonu kullanılmıştır

## 3. Araştırma Sonuçları ve Tartışma

Yüz tanıma işlemi için Tablo 1’de belirtilen 4 farklı veri tabanından toplamda 360.000 görüntü üzerinde eğitim işlemi gerçekleştirilmiştir. Veri kümesinin %80’i eğitim verisi, %20’si test verisi olarak kullanılmıştır. Modeller eğitilirken eğitimin kaç kez gerçekleşeceği epoch (tur sayısı) ile belirlenmektedir. İlk

olarak modeller 10 tur eğitilmişlerdir. Eğitimin her adımında modelin ağırlıkları güncellenerek başarı oranları değişmektedir. Başarı oranları sabit kaldığı zaman modellerin eğitimi tamamlanmıştır. Eğitimin her adımında modelin ağırlıkları güncellenerek başarı oranları değişmektedir. Başarı oranları sabit kaldığı zaman modellerin eğitimi tamamlanmıştır. Modellerin başarı oranı hesaplanırken binary crossentropy fonksiyonu kullanılmıştır. Tablo 3’ de modellerin epoch değerlerine göre oluşan yüz tanıma başarı oranları gösterilmiştir.

Tablo 3. Geliştirilen yüz tanıma modellerinin başarı oranları

Tur Sayısı	Vgg-16 Modeli Başarı Oranı	AlexNet Modeli Başarı Oranı	ZFNET Modeli Başarı Oranı
10	%55,77	%47,35	%41,11
20	%67,01	%56,74	%50,47
30	%78,57	%65,35	%71,98
40	%90,14	%88,49	%80,63
50	%92,03	%88,49	%89,98

VGG-16 mimarisi kullanılarak inşa edilen model 1. eğitim adımında %55,77 başarı oranı yakalamıştır. Başarı oranı 43 epoch değerinden sonra sabit kalmıştır. 43. eğitim adımında model %92.03 başarı oranına ulaşmıştır.

AlexNet mimarisi ile tasarlanan 2. model 1. eğitim aşamasında %47,35 başarılı olmuştur. Model 38. epoch değerinde maksimum %88,49 değerini yakalamıştır. Geriye kalan eğitim adımlarında başarı oranı artmamıştır.

ZF Net ile tasarlanan model ise ilk eğitim adımında model %41,11 başarı oranı yakalamıştır. Tur sayısı arttıkça model 46.turda %89,98 başarı oranını yakalamıştır. 46. turdan itibaren başarı oranı sabit kalmıştır.

Modeller incelendiğinde en başarılı sonucu %92.03 başarı oranı ile VGG-16 modeli vermiştir. Modellere ait başarı oranlarının yüksek olmasının en önemli sebebi veri kümesinin sadece yüz görüntüleri içermesidir. Ayrıca veri kümesi üzerinde ön işleme adımı uygulanarak görüntüler modellerin giriş boyutuna uygun hale getirilmiştir. Bu da başarı oranını arttırmıştır.

Yüz ifadesi tanıma işlemi için AlexNet, VGG-16 ve ZF Net mimarileri referans alınarak 3 farklı model tasarlanmıştır. Modeller FER-2013 veri seti üzerinde eğitilmiştir. FER-2013 veri seti 7 farklı sınıftan farklı sayılarda görüntüler içermektedir.

Fer -2013 veri seti üzerinde geliştirilen modellere ait başarı oranları Tablo 4 ‘de ifade edilmiştir. Başarı oranları hesaplanırken Categorical Cross-Entropy loss fonksiyonu kullanılmıştır. Bu fonksiyon çok sınıflı sınıflandırmada kullanılmaktadır (Lau and Baldwin 2016). Tablo 4 incelendiğinde başarı oranı en yüksek olan model % 63,17 ile VGG-16 modeli olmuştur.

Tablo 4. Geliştirilen duygu analizi modellerinin başarı oranları

Tur Sayısı	Vgg-16 Modeli Başarı Oranı	AlexNet Modeli Başarı Oranı	ZFNET Modeli Başarı Oranı
10	%31,17	%28,14	%30,14
20	%38,23	%34,21	%37,10
30	%45,77	%40,36	%44,82
40	%54,39	%48,17	%50,37
50	%63,17	%53,84	%55,71

Modellerin başarı oranını arttırmak için Fer-2013 veri setine eklemeler yapılmıştır. Her sınıfta eşit sayıda görüntü olacak şekilde veri seti tekrardan düzenlenmiştir (Mungen et al. 2020). Her sınıf için 5000 adet görüntü yer alacak şekilde veri seti toplamda 35.000 görüntüden oluşturulmuştur. Verilerin %80'i eğitim, %20'si test aşamasında kullanılmıştır.

Veri kümesi dengeli hale getirildikten sonra tekrar eğitilen modeller içerisinde en başarılı sonucu %87,18 ile VGG-16 modeli vermiştir. Veri kümesinin dengeli olması modellerin başarı oranını minimum %25 arttırmıştır. Başarı oranları Tablo 5' de ifade edilmiştir.

Tablo 5. Geliştirilen duygu analizi (yüz ifade tanıma) modellerinin başarı oranları

Tur Sayısı	Vgg-16 Modeli Başarı Oranı	AlexNet Modeli Başarı Oranı	ZFNET Modeli Başarı Oranı
10	%64,38	%49,25	%52,31
20	%71,48	%58,67	%60,51
30	%75,11	%67,14	%68,76
40	%81,69	%74,89	%74,32
50	%87,18	%84,36	%85,61

#### 4. Sonuç

Bu çalışmada derin öğrenme teknikleri kullanılarak yüz tanıma ve yüz ifadesi analizi yapılmıştır. Yüz tanıma uygulamasında başarı oranını etkileyen en önemli faktör modellerin giriş katmanındaki girdi boyutlarına göre eğitim kümesinin oluşturulmasıdır. Ayrıca veri kümesi hazırlanırken kişilerin farklı poz ve çözünürlükteki görüntülerinin kullanılması da başarı oranını arttırmıştır. AlexNet, VGG-16 ve ZF Net mimarileri referans alınarak inşa edilen modellerin başarı oranları ile eğitim adım sayıları arttıkça artmıştır. VGG-16 modeli ile geliştirilen yüz tanıma modeli ve yüz ifadesi tanıma modelleri diğer modellere göre daha başarılı sonuçlar vermiştir

#### 5. Teşekkür

Bu çalışma, Mersin Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimi tarafından 2019-2-TP2-3532 numaralı proje olarak desteklenmiştir. Destekleri için Mersin Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimi'ne teşekkür ederim.

#### 6. Kaynakça

- Dehghan, Afshin, Enrique G Ortiz Guang, Shu Syed, and Zain Masood. DAGER: Deep Age, Gender and Emotion Recognition Using Convolutional Neural Networks. <https://www.sighthound.com/products/cloud> (December 14, 2020).
- Dürr, Oliver et al. 2015. "Deep Learning on a Raspberry Pi for Real Time Face Recognition TubeCam: A New System to Detect Small Mammals (Foremost Mustelids and Dormice) View Project Speaker Diarization View Project Deep Learning on a Raspberry Pi for Real Time Face Recognition." researchgate.net. <https://www.researchgate.net/publication/279537625> (December 18, 2020).
- Goodfellow, I, Y Bengio, A Courville, and Y Bengio. 2016. "Deep Learning." <https://doi.org/10.4258/hir.2016.22.4.351> (January 3, 2021).
- Hjelmås, Erik. hig.no Feature-Based Face Recognition. <http://www.hig.no/~erikh/papers/nobim2000.pdf> (December 18, 2020).
- Kalocsai, P, C von der Malsburg, J Horn - Image and Vision Computing, and undefined 2000. "Face Recognition by Statistical Analysis of Feature Detectors." Elsevier. [https://www.sciencedirect.com/science/article/pii/S0262885699000517?casa\\_token=H2O15WOWp8AAAAAA:LjVzRWDnOUdLj-oM7Pq4-JOM6jPmDySyCEictAB\\_iKVYec6n-aMMtmJXNAkJ1muYdQFJ1zKUES2q](https://www.sciencedirect.com/science/article/pii/S0262885699000517?casa_token=H2O15WOWp8AAAAAA:LjVzRWDnOUdLj-oM7Pq4-JOM6jPmDySyCEictAB_iKVYec6n-aMMtmJXNAkJ1muYdQFJ1zKUES2q) (December 18, 2020).
- Koushik, Jayanth. 2016. "Understanding Convolutional Neural Networks." <http://arxiv.org/abs/1605.09081> (January 3, 2021).
- Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. 2017. "ImageNet Classification with Deep Convolutional Neural Networks." Communications of the ACM.
- Kumar, Ashu, Amandeep Kaur, and · Munish Kumar. 2019. "Face Detection Techniques: A Review." Artificial Intelligence Review 52: 927–48. <https://doi.org/10.1007/s10462-018-9650-2> (December 18, 2020).
- Lau, Jey Han, and Timothy Baldwin. 2016. "An Empirical Evaluation of Doc2vec with Practical Insights into Document Embedding Generation." In Association for Computational Linguistics (ACL), 78–86.
- Mungen, A. A., Aygün, İ., & Mehmet, K. A. Y. A. (2020). Finding the Relationship Between News and Social Media Users' Emotions in the COVID-19 Process. Sakarya University Journal of Computer and Information Sciences, 3(3), 250-263.
- Özkan, İ, E Ülker - Gaziosmanpaşa Bilimsel Araştırma, and undefined 2017. "Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri." pdfs.semanticscholar.org. <https://pdfs.semanticscholar.org/e2b2/4cc0c4c529d341c450bc18f949968c7cac8d.pdf> (January 3, 2021).
- Pervan, Nergis. 2019. ANKARA ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ YÜKSEK LİSANS TEZİ DERİN

- ÖĞRENME YAKLAŞIMLARI KULLANARAK TÜRKÇE METİNLERDEN ANLAMSAL ÇIKARIM YAPMA. <http://dspace.ankara.edu.tr/xmlui/handle/20.500.12575/41714> (January 3, 2021).
- Salunke, Vibha V., and C. G. Patil. 2018. "A New Approach for Automatic Face Emotion Recognition and Classification Based on Deep Networks." In 2017 International Conference on Computing, Communication, Control and Automation, ICCUBEA 2017, Institute of Electrical and Electronics Engineers Inc.
- Satish, Anila, Nanjundappan Devarajan, S Anila, and N Devarajan. 2010. researchgate.net Simple and Fast Face Detection System Based on Edges IoT Based Automatic Farm Management System Using Wireless Sensor Networks View Project Powerful and Dominating Woman View Project Simple and Fast Face Detection System Based on Edges. <https://www.researchgate.net/publication/225292501> (December 18, 2020).
- Simonyan, Karen, and Andrew Zisserman. 2015. "Very Deep Convolutional Networks for Large-Scale Image Recognition." In 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, International Conference on Learning Representations, ICLR.
- Tüfekçi, Mustafa, and Doç Fatih Karpat. "Derin Öğrenme Mimarilerinden Konvolüsyonel Sinir Ağları (CNN) Üzerinde Görüntü İşleme-Sınıflandırma Kabiliyetinin Arttırılmasına Yönelik Yapılan Çalışmaların İncelenmesi." www.set-science.com (January 2, 2021).
- Vatsa, M, R Singh, and A Majumdar. 2018. "Deep Learning in Biometrics." [https://books.google.com/books?hl=tr&lr=&id=rGhQDwAAQBAJ&oi=fnd&pg=PP1&dq=Vatsa,+M.,+Singh,+R.+and+Majumdar,+A.,+2018.+Deep+Learning+in+Biometrics,+CRC+Press+Taylor+%26+Francis+Group,+New+York&ots=e69JCxR4g3&sig=J4s94pKBX3s\\_HGKCMsSzsyy3pU](https://books.google.com/books?hl=tr&lr=&id=rGhQDwAAQBAJ&oi=fnd&pg=PP1&dq=Vatsa,+M.,+Singh,+R.+and+Majumdar,+A.,+2018.+Deep+Learning+in+Biometrics,+CRC+Press+Taylor+%26+Francis+Group,+New+York&ots=e69JCxR4g3&sig=J4s94pKBX3s_HGKCMsSzsyy3pU) (December 12, 2020).
- Wiskott, Laurenz, Jean Marc Fellous, Norbert Krüger, and Christoph Der Von Malsburg. 1997. "Face Recognition by Elastic Bunch Graph Matching." IEEE Transactions on Pattern Analysis and Machine Intelligence 19(7): 775–79.
- Zeiler, Matthew D., and Rob Fergus. 2014. "Visualizing and Understanding Convolutional Networks." In Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Springer Verlag, 818–33.
- Zhang, H, Y Xie, C Xu - Proceedings 2011 International, and undefined 2011. "A Classifier Training Method for Face Detection Based on AdaBoost." ieexplore.ieee.org. <https://ieexplore.ieee.org/abstract/document/6199306/> (December 18, 2020).
- Learning." <https://doi.org/10.4258/hir.2016.22.4.351> (January 3, 2021).
- Koushik, Jayanth. 2016. "Understanding Convolutional Neural Networks." <http://arxiv.org/abs/1605.09081> (January 3, 2021).
- Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. 2017. "ImageNet Classification with Deep Convolutional Neural Networks." *Communications of the ACM*.
- Özkan, İ, E Ülker - Gaziosmanpaşa Bilimsel Araştırma, and undefined 2017. "Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri." [pdfs.semanticscholar.org/e2b2/4cc0c4c529d341c450bc18f949968c7cac8d.pdf](https://pdfs.semanticscholar.org/e2b2/4cc0c4c529d341c450bc18f949968c7cac8d.pdf) (January 3, 2021).
- Pervan, Nergis. 2019. ANKARA ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ YÜKSEK LİSANS TEZİ DERİN ÖĞRENME YAKLAŞIMLARI KULLANARAK TÜRKÇE METİNLERDEN ANLAMSAL ÇIKARIM YAPMA. <http://dspace.ankara.edu.tr/xmlui/handle/20.500.12575/41714> (January 3, 2021).
- Simonyan, Karen, and Andrew Zisserman. 2015. "Very Deep Convolutional Networks for Large-Scale Image Recognition." In 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, International Conference on Learning Representations, ICLR.
- Tüfekçi, Mustafa, and Doç Fatih Karpat. "Derin Öğrenme Mimarilerinden Konvolüsyonel Sinir Ağları (CNN) Üzerinde Görüntü İşleme-Sınıflandırma Kabiliyetinin Arttırılmasına Yönelik Yapılan Çalışmaların İncelenmesi." www.set-science.com (January 2, 2021).
- Zeiler, Matthew D., and Rob Fergus. 2014. "Visualizing and Understanding Convolutional Networks." In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer Verlag, 818–33.