



Düzce Üniversitesi Bilim ve Teknoloji Dergisi

Araştırma Makalesi

Derin Öğrenme ve Destek Vektör Makineleri ile Görüntüden Cinsiyet Tahmini

 Sedat METLEK ^{a,*},  Kıyas KAYAALP ^b

^a *Elektronik ve Otomasyon Bölümü, Teknik Bilimler Meslek Yüksekokulu, Burdur Mehmet Akif Ersoy Üniversitesi, Burdur, TÜRKİYE*

^b *Bilgisayar Teknolojileri Bölümü, Uluborlu Selahattin Karasoy Meslek Yüksekokulu, Isparta Uygulamalı Bilimler Üniversitesi, Isparta, TÜRKİYE*

* Sorumlu yazarın e-posta adresi: sedatmetlek@mehmetakif.edu.tr

DOI: 10.29130/dubited.707316

ÖZET

Yapılan çalışmada günümüzün popüler konularından olan derin öğrenme algoritmaları üzerine bir uygulama geliştirilmiştir. Geliştirilen uygulamada görüntülerden yüz tespiti yapıp sonrasında görüntüdeki kişinin cinsiyet tahmini gerçekleştirilmiştir. Bu uygulamada Wiki görüntü veri tabanından elde edilen 62328 görüntü kullanılmıştır. Kullanılan görüntüler üzerinde, yüz görüntüsü bulunmayanlar veri setinden çıkartılarak yeni bir veri seti oluşturulmuştur. Oluşturulan veri setindeki görüntülerden, ileri derin öğrenme tekniklerinden biri olan Evrişimsel Sinir Ağları (ESA) yöntemi kullanılarak öznelikler çıkartılmıştır. Elde edilen öznelikler Destek Vektör Makinesi (DVM) kullanılarak sınıflandırılmıştır. Sınıflandırma başarısı karmaşıklık matrisi ile gösterilmiş olup, %94,48 başarı oranı ile sınıflandırma işlemi gerçekleştirilmiştir.

Anahtar Kelimeler: *Derin öğrenme, Evrişimsel sinir ağı, Görüntüden cinsiyet tahmini*

Image Gender Prediction with Deep Learning and Support Vector Machines

ABSTRACT

In this study, an application was developed on deep learning algorithms, which is one of the popular topics of today. In the developed application, face detection was made from the images and then the gender prediction of the person in the image was made. In this application, 62328 images obtained from the Wiki image database were used. A new data was created by removing those without facial images on the images used. Features were extracted from the images in the data set, using the Convolutional Neural Networks (CNN) method, one of the advanced deep learning techniques. It is classified with the Support Vector Machine (SVM) in the obtained attributes. The classification success was shown by the complexity matrix and the classification process was performed with a 94.48% success rate.

Keywords: *Deep learning, Convolutional neural network, Image gender prediction*

Geliş: 21/03/2020, Düzeltme: 18/07/2020, Kabul: 25/07/2020

I. GİRİŞ

İnsanlar ilk kez karşılaştıkları birine “Hanımefendi” veya “Beyefendi” diye hitap etmeden önce farkında olmadan bir sınıflandırma işlemi gerçekleştirmektedir. Günlük yaşantıda yapılan bu sınıflandırma işlemine cinsiyet sınıflandırma denilmektedir. Dijital ortamda ise, cinsiyet sınıflandırma, daha çok görüntü işleme uygulamalarında kullanılmaktadır. Cinsiyet sınıflandırmanın en çok kullanıldığı görüntü işleme uygulamaları ise yüz tanıma, yaş tahmini ve ruhsal durum sınıflandırmasıdır [1]. Cinsiyet sınıflandırması genel olarak akıllı ve tıbbi hizmet alanı, interaktif sistemler, güvenlik sistemleri gibi birçok ticari uygulamada kullanılmaktadır. Bu uygulamalardaki sınıflandırma işlemleri için literatürde yapay sinir ağı (YSA), derin öğrenme gibi esnek hesaplama yöntemleri kullanılmaktadır. Yüz tanıma yöntemlerinde genellikle, görüntü üzerinde ölçüm tabanlı işlemler uygulandığından dolayı aydınlatma gibi çevresel şartlardan etkilenebilmektedir. Günümüzde dijital görüntülerden, kişinin saç, göz, ağız, burun gibi özniteliklerine bakılarak yüksek doğruluk oranlarında cinsiyet tahmini yapılabilmektedir. Bununla birlikte literatürde daha temiz görüntü elde etmek için çevresel şartları indirgeyen bazı çalışmalar da bulunmaktadır [2].

Görüntüdeki bir nesneyi tespit etme işleminde genellikle eğriler ve şekli tanımlayan öznitelikler kullanılmaktadır. Bu nedenle nesne tespit etme yöntemleri, iki boyutlu eşleştirme ve bölge tabanlı işlemler olmak üzere ikiye ayrılmaktadır [3]. İki boyutlu eşleştirme işlemi, herhangi bir görüntü üzerindeki şeklin sınır noktalarının, görüntü üzerindeki histogram dağılımına bakılarak elde edilmesidir. Bölge tabanlı işlemler ise bir görüntü üzerinde belli bir filtrenin dolaştırılarak nesnenin tespit edilmesi işlemidir. Bu işlemlerden sonra elde edilen şeklin görüntüsü üzerinden öznitelik vektörü elde edilmektedir. Daha sonrasında, elde edilen öznitelik vektörü de sınıflandırma işleminde kullanılmaktadır. Aynı işlem farklı konulardaki birçok uygulama içinde gerçekleştirilmektedir [4].

Görüntülerden nesne tespit etmek için gelişen teknolojiye bağlı olarak farklı esnek hesaplama yöntemleri de geliştirilmiştir. Bu yöntemlerden biriside derin öğrenmedir. Sınıflandırma problemleri için literatürde daha önceden yapılan çalışmaların önerdiği öznitelik çıkarma işlemi, derin öğrenme algoritmalarında görüntü üzerinde filtreler dolaştırılarak otomatik olarak gerçekleştirilmektedir. Görüntüdeki objenin kayması ve ölçeğinin değişmesi durumlarında bile, diğer öznitelik çıkarma tabanlı çalışan algoritmalara göre daha iyi sınıflandırma sonuçları verdiği görülmektedir [5].

Gerçekleştirilen çalışmada, cinsiyet bilgilerinin elde edilmesini zorlaştıran aydınlatma, çözünürlük, aşırı pozlama gibi dış etkenlerden daha az etkilenen derin öğrenme tabanlı bir sınıflandırma algoritması temel alınmıştır. Geliştirilen algoritmada, yüz bölgesindeki kıvrımları tespit edebilecek yatay, dikey ve çapraz filtreler kullanılmıştır. Elde edilen öznitelikler ile cinsiyet sınıflandırma işlemleri, hedeflendiği gibi yüksek oranda gerçekleştirilmiştir. Sunulan çalışmada 2. bölümde literatürdeki konu ile ilgili yapılan çalışmalar, 3. bölümde materyal ve metod, 4. bölümde uygulamanın geliştirilmesi ve 5. bölümde de elde edilen sonuçlar sunulmaktadır.

II. İLGİLİ ÇALIŞMALAR

Günümüz teknolojisinin hızlı şekilde gelişmesine paralel olarak farklı konularda çok fazla miktarda veri üretilmektedir. Üretilen bu verilerin büyük bir kısmını, ses ve görüntü verileri oluşturmaktadır.

Görüntü verilerinin çoğunluğu ise, kameralardan elde edilen ve anlık olarak dijital ortama aktarılan video görüntüleri oluşturmaktadır.

Bu görüntüler üzerinde en fazla yapılan işlemlerden biriside yüz tespittir. Yüz tespiti kontrollü ortamlar altında fazlasıyla çalışılmış [6] ve yüksek doğruluk oranlarında başarı elde edilmiş olmasına rağmen aşırı aydınlatma ve pozlama gibi kısıtlanmamış ortamlar altında istenilen başarı oranı yakalanamamıştır. Buna ek olarak yüz görüntüsünden cinsiyet tahmini konusu, kontrollü ortamlar altında bile araştırılmaya devam etmektedir. Bu çalışmalar için özel olarak hazırlanmış birçok görüntü veri tabanı da bulunmaktadır [7,8,9].

Bu veri tabanları kullanılarak yüz tanıma, yaş ve cinsiyet tahmini gibi birçok çalışma yapılmıştır. Bunlardan Shan, 7443 yüz görüntüsü üzerinde Adaboost ve DVM algoritmalarını kullanarak cinsiyet tahmini gerçekleştirmiştir [10]. Fakat gerçekleştirilen çalışmada binary görüntülerden histogram elde edildiği için çevresel etkilere çok fazla duyarlıdır.

Bir diğer çalışmada Ren ve Li, 6840 yüz görüntüsü üzerinde, üç farklı özellik türünü kullanarak %98’lik başarı ile cinsiyet tahmininde bulunmuştur [11]. Çalışmada, Real Adabost ve Doğrusal DVM algoritmaları kullanılmıştır. Bunun ile birlikte çalışmada kullanılan öznelikler yine histogram tabanlı olduğundan dolayı çevresel etkilerden çok fazla etkilenecek niteliktedir. Bunların dışında literatürde yapılan diğer çalışmaların özeti Tablo 1’de sunulmuştur.

Tablo 1. Cinsiyet Sınıflandırmada kullanılan yöntemler, veri tabanı özellikleri ve başarı oranları tablosu.

Referans	Yıl	Özellik Çıkarma	Sınıflandırma Algoritması	Veri Tabanı	Görüntü Elde Etme Şartları	Kullanılan Görüntü Sayısı	% Başarı Oranı
14	2014	Gabor Filtresi	DVM	FERET	Kontrollü Ortam	1199	98,18
14	2014	Gabor Filtresi	DVM	FEI	Gülümseme, Gülümsememe	400	96,61
14	2014	Gabor Filtresi	DVM	AR	Aydınlatma, Poz, Arka Fon, Yüz ifadesi	126	96,15
14	2014	Gabor Filtresi	DVM	HINT	Kontrolsüz Ortam	605	93,33
17	2014	Geçici Özellikler	DWT, PCA	FERET	Kontrollü Ortam	1199	92,33
15	2008	Gabor Özellikleri	PCA, DVM	FERET	Aydınlatma, Yüz ifadesi	300	91,91
13	2011	Bölge Yerelleştirme	Doğrusal DVM	YAHOO FLICKER	Kontrolsüz Ortam	26766	90,00
12	2014	Katman Özellikleri	DNN	FERET, CAS-PEAL	Kontrolsüz Ortam	13500	89,63
18	2014	Gabor Filtresi	DVM	LFW	Kontrolsüz Ortam	13010	88,34
16	2010	Kenarlara dayalı	AdaBoost	WEBING	Aydınlatma, Poz, Arka Fon	1948	87,60

18	2016	Bilgilendirici Yama Çıkarma	Yerel DNN	LFW	Kontrolsüz Ortam	13233	77,87
19	2017	Katman Özellikleri	DNN, RoR34	WIKI + IMDB	Kontrolsüz Ortam	440607+ 62328	93,24
20	2019	Katman Özellikleri	CNN	WIKI+ IMDB	Kontrolsüz Ortam	460723 + 62328	89,70

III. MATERYAL ve METOD

A. MATERYAL

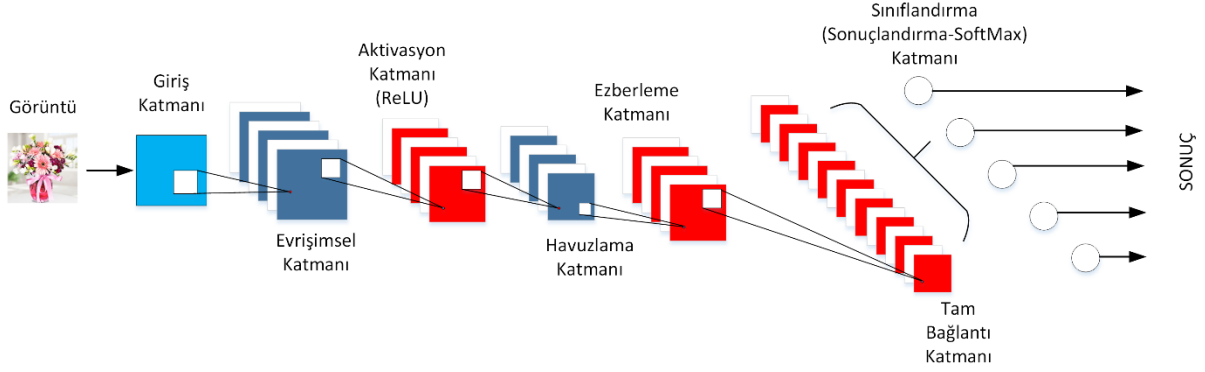
Günümüzde bilginin elde edildiği veri kaynakları oldukça önemlidir. Geliştirilen algoritmaların test edilebilirliğini ve doğruluğunu sağlamak için güvenilir verilere gerek vardır. Bu amaç doğrultusunda yapay sinir ağı ve derin öğrenme algoritmaları gibi esnek hesaplama yöntemlerinin geliştirilmesini sağlamak için özel olarak hazırlanmış birçok görüntü veri tabanı bulunmaktadır [7,8,9]. Bu veri tabanlarında da daha önceden test edilmiş on binlerce görüntü bulunmaktadır. Derin öğrenme algoritmalarında başarı oranını yükseltmek için ağa verilen örnek sayısının da artırılması gerekmektedir. Örnek sayısının artması ile doğru orantılı olarak veri setinin de boyutu artmaktadır. Bu nedenle makalede, cinsiyet sınıflandırması için 62328 görüntü verisinden oluşan Wikipedia'nın sunmuş olduğu kontrolsüz ortamlardan elde edilen Wiki görüntü veri tabanı kullanılmıştır [8]. Hazırlanan algoritma MatLab 2018a versiyonunda uygulanmıştır. Uygulanan algoritma Intel(R) Core(TM) i7-4702MQ 2.2GHz işlemci, 16 GB Ram, NVIIDIA GeForce GT 740M 2 GB ekran kartı bulunan bir bilgisayarda test edilmiştir. Hazırlanan derin öğrenme algoritması 21 katmandan oluşmakta ve sınıflandırıcı olarak DVM kullanılmıştır.

B. METOD

Derin öğrenme, 1960'lı yıllarda ortaya çıkan yapay sinir ağlarının günümüze evrimleşerek geldiği son aşamadır. Derin öğrenmenin gelişim aşamaları ilk olarak Perceptron modeli ile başlayıp sırasıyla Tek Katmanlı Yapay Sinir Ağı, sonrasında Çok Katmanlı Yapay Sinir Ağlarına ve sonrasında da Derin Öğrenme Algoritmaları olarak gelişim göstermiştir. Yapay sinir ağları ilk olarak, insan sinir hücrelerinin yapısı temel alınarak geliştirilmiştir. Bunun sonucunda da bilgisayarlı görü uygulamalarından, sinyal ve ses işleme uygulamalarına kadar pek çok farklı alanda kullanılmaktadır. Veri artışının çok fazla olduğu günümüz teknolojilerinde, bu verilerden anlamlı bilgiler elde etmek zorlaşmaktadır. Bu nedenle YSA gibi esnek hesaplama yöntemlerinin önemi her geçen gün daha da artmaktadır [21,22,23].

Günümüzde bir diğer esnek hesaplama yöntemlerinden olan derin öğrenme ile doğrusal olmayan pek çok probleme çözüm bulunabilmektedir. Derin öğrenme algoritmalarında, çok katmanlı bir ağ yapısı kullanılmaktadır. Katmanlı mimariye sahip bu yapıda, ilk katman girdi katmanını, son katman ise sınıflandırma katmanını temsil etmektedir. Veri girişi işlemi, girdi katmanı olarak adlandırılan ilk katmandan yapılmaktadır. İlk katmandan sonra normalizasyon işlemi gerçekleştirilmektedir. Daha sonra sırasıyla Evrimsel (Convolution), Aktivasyon (ReLU), Havuzlama (Pooling), Ezberleme

(Dropout), Tam Bağlantı (Fullconnected) ve Sonuçlandırma işlemleri gerçekleştirilmektedir [24,25]. Bu yapı özet olarak Şekil 1’de gösterilmektedir.



Şekil 1. Derin öğrenme katman yapısı

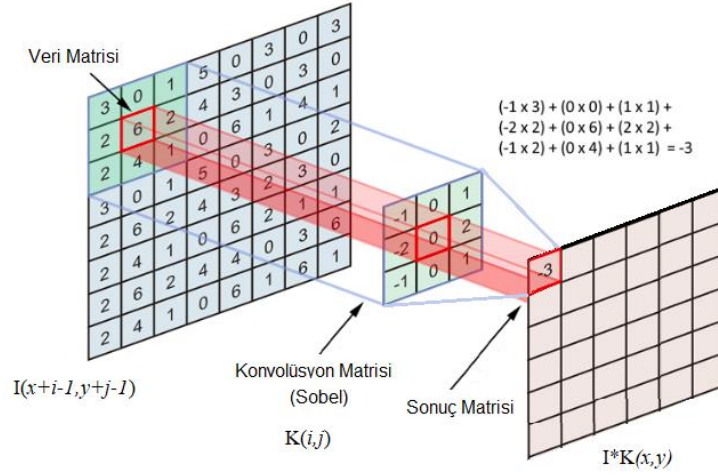
Çok katmanlı yapıda veriye ait öznitelikler ESA ile çıkartılmaktadır. Eğer görüntü işleme uygulaması geliştiriliyorsa çıkartılan öznitelikler içerisinde görüntüye ait kenar ve doku gibi bazı öznitelikler bulunmaktadır. Otomatik olarak çıkartılan bu öznitelikler diğer alt katmanlara sırasıyla iletilir ve en son sonuç katmanından değerler elde edilir. Bu süreçte veri, Şekil 1’de gösterilen ve aşağıda anlatılan katmanlar arasında baştan sona doğru sırasıyla aktarılmaktadır.

B. 1. Giriş Katmanı

Bu katmanı, genellikle derin öğrenme algoritmasına aktarılan görüntüler oluşturur. Bu görüntüler kimi zaman bilgisayarda kayıtlı bir video ya da resim olurken, kimi zamanda kameradan alınan anlık görüntüler de olabilmektedir. Giriş katmanından gelen veriler farklı tür, boyut ve içerikte olabilmektedir. Bundan dolayı ilk olarak görüntüler normalizasyon işlemine tabi tutulurlar. Böylece derin öğrenme algoritmasına sunulan tüm görüntüler belli bir standarda getirilmiş olur.

B. 2. Evrışimsel (Convolution) Katmanı

Derin öğrenmenin ana katmanı konvolüsyon katmanıdır. YSA gibi yöntemlerde öznitelikleri kullanıcının kendisinin çıkarması gerekmektedir. Konvolüsyon sinir ağlarında veya derin öğrenme algoritmalarında, bu katmanda öznitelikler otomatik olarak çıkartılmaktadır. Bu katmana sunulan girdi verisi üzerinde boyutları önceden belirlenen (AlexNette 11x11, ZfNet’te 7x7, Google Net ve VggNet’te 5x5, 3x3, 2x2 gibi) matrisler dolaştırılarak öznitelikler çıkartılmaktadır. Çıkartılan öznitelikler ile başka bir matris oluşturularak giriş verisinden daha küçük boyutta veriler elde edilmektedir [26]. Şekil 2 ve Denklem 1’de ifade edildiği üzere, $(x+i-1),(y+j-1)$ boyutundaki I matrisi üzerinde $[i,j]$ boyutunda bir K matrisi dolaştırılarak, $[x,y]$ boyutunda yeni bir $I*K$ matrisi elde edilmektedir.



Şekil 2. Konvolüsyon katmanı

$$(I * K)_{xy} = \sum_{i=1}^h \sum_{j=1}^w K_{ij} \cdot I_{x+i-1, y+j-1} \quad (1)$$

Görüntü üzerinde dolaştırılacak matrisin, ağıın eğitim ve başarısına doğrudan etkisi bulunmaktadır. Örneğin ağı üzerinde dolaştırılacak matris çok büyük boyutlu ve büyük sayılabilecek rakamlardan oluşması durumunda, ağıın eğitilmesi çok uzun sürebilir, hatta tercih edilen rakamlara göre hata ile karşılaşılma durumu da söz konusu olabilir [27].

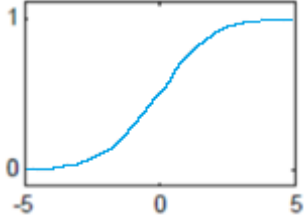
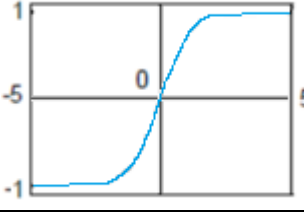

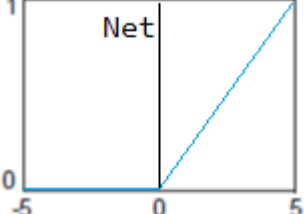
B. 3. Aktivasyon (ReLU) Katmanı

RGB görüntüler 3 boyutlu matris formuna sahiptir. Buradaki her bir boyutta bir renk bileşeni bulunmaktadır. Görüntü işleme uygulamalarında kullanılan, derin öğrenme algoritmalarında da genellikle RGB formatındaki renkli görüntüler kullanılmaktadır. Bu nedenle konvolüsyon katmanında belirtilen Denklem 1, görüntünün her bir boyutuna uygulanmaktadır. Birçok uygulamada da aktivasyon katmanı, konvolüsyon katmanından sonra gelmektedir. Bu nedenle aktivasyon katmanında Denklem 2 uygulanmaktadır.

$$conv(I, K)_{xy} = \sigma \left(b + \sum_{i=1}^h \sum_{j=1}^w \sum_{k=1}^d K_{ij} \cdot I_{x+i-1, y+j-1, k} \right) \quad (2)$$

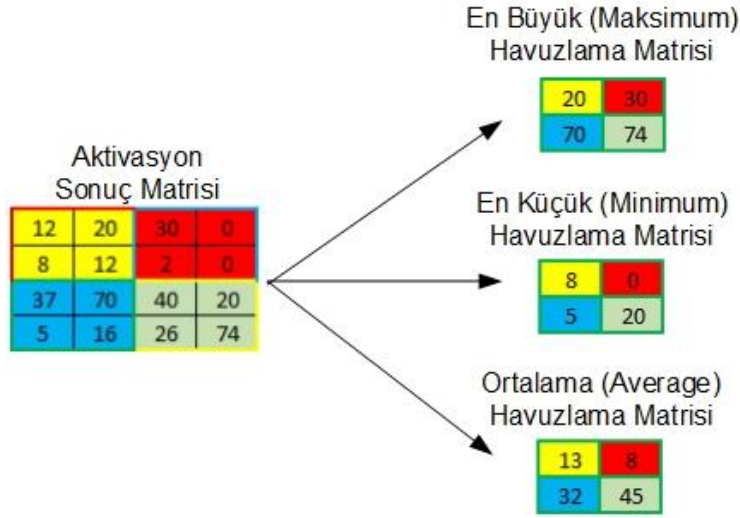
Denklem 2’de, Denklem 1’den farklı olarak bulunan d değeri görüntünün boyutlarını ifade etmektedir. b değeri ise biası ifade etmektedir. Bias değeri genellikle isteğe bağlı olarak kullanılmaktadır. σ değeri ise aktivasyon fonksiyonunu ifade etmektedir. Aktivasyon fonksiyonu olarak da genellikle Tablo 2’deki 3,4,5,6 numaralı fonksiyonlar kullanılmaktadır. Aktivasyon katmanının genel amacı, konvolüsyon katmanından elde edilen negatif değerlerin pozitifeye dönüştürülmesidir [28].

Tablo 2. Aktivasyon fonksiyonları [29]

Sigmoid	$y = \frac{1}{1 + e^{-Net}}$		(3)
Tanjant Hiperbolik	$y = \frac{e^{Net} - e^{-Net}}{e^{Net} + e^{-Net}}$		(4)
Adım Fonksiyonu	$y = \begin{cases} 1 & Net \geq 0 \\ 0 & Net < 0 \end{cases}$		(5)
ReLU	$y = \max(0, Net)$		(6)

B. 4. Havuzlama (Pooling) Katmanı

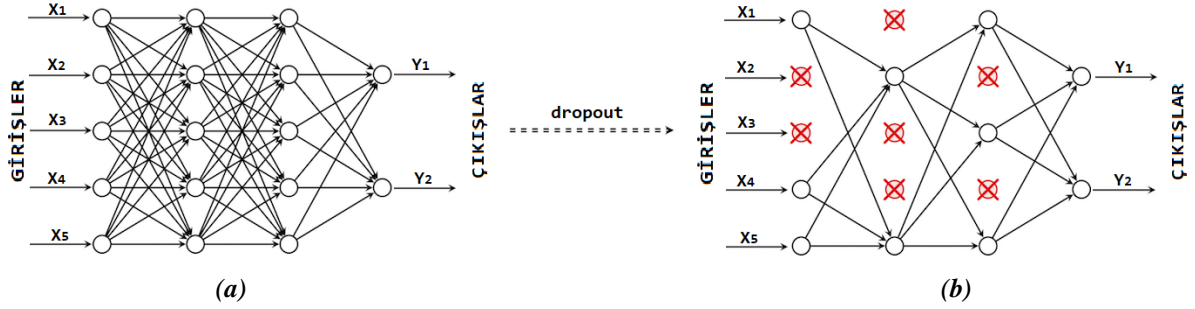
Genellikle aktivasyon katmanından sonra kullanılan havuzlama katmanı, verileri daha küçük boyutlara indirgemek için kullanılmaktadır. Veri indirgeme işleminde ağı daha hızlı çalışması sağlanırken, veri kayıpları yaşanabilmektedir [30]. Havuzlama katmanında veri indirgeme işlemi için tercih edilen matrisler genellikle Şekil 3'de ifade edilen en büyük (max pooling) değer, en küçük (min pooling) değer ve ortalama (average pooling) değer matrisleridir. Böylece kare formundaki bu matrisler, aktivasyon katmanından elde edilen matris üzerinde dolaştırılarak yeni bir veri matrisi elde edilmesini sağlar.



Şekil 3. Havuzlama katmanı

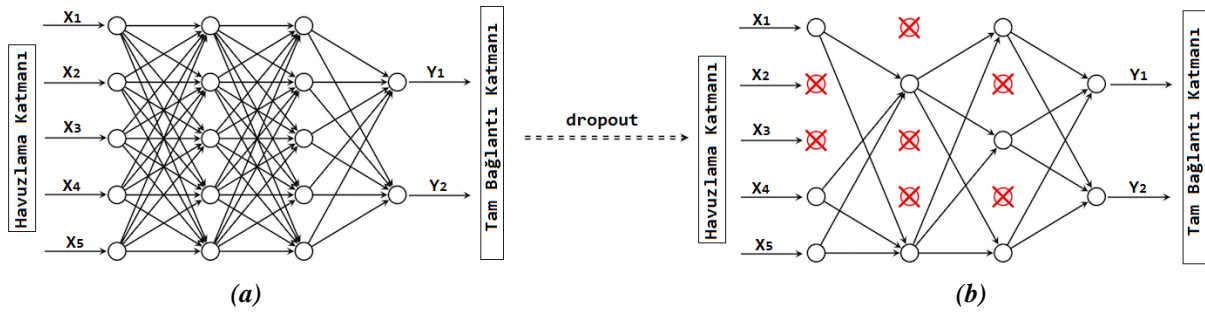
B. 5. Ezberleme (Dropout) Katmanı

Çok katmanlı yapay sinir ağı modellerinin eğitim işleminde çok az veri kullanılması durumunda, ağ eğitim setini ezberleyebilmektedir. Bu nedenle eğitim setindeki bazı değerlerin ağa unutturulması gerekebilir. Derin öğrenme algoritmalarında da kullanılan bu yapı sayesinde ağın ezberlemesinin önüne geçilmektedir. Bunun sonucunda da ağın öğrenme süresi uzayabilmektedir. Şekil 4(a)'da ezberleme yapmış çok katmanlı YSA modeli gösterilerken, Şekil 4(b)'de ezberleme katmanı kullanılmış çok katmanlı YSA modeli gösterilmektedir.



Şekil 4. (a) Ezberleme yapmış YSA modeli, (b) Ezberleme katmanı kullanılmış YSA modeli

Şekil 5(a)'da ise ezberleme yapmış derin öğrenme modeli gösterilirken, Şekil 5(b)'de ezberleme katmanı kullanılmış derin öğrenme modeli gösterilmiştir. Şekil 4 ve 5'den de anlaşılacağı üzere derin öğrenme modelinde ezberleme işlemi havuzlama katmanı ile tam bağlantı katmanı arasında gerçekleşmektedir.

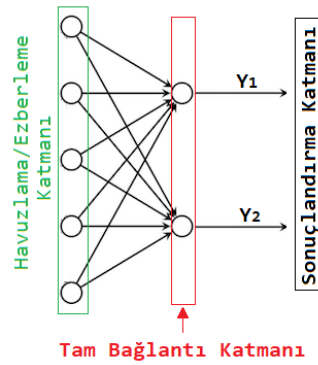


Şekil 5. (a) Ezberleme yapmış derin öğrenme modeli, (b) Ezberleme katmanı kullanılmış derin öğrenme modeli

Literatürdeki bazı çalışmalarda, ağın çıkışını etkileyen girişlerden bazılarının sonucu çok fazla etkilediği görülmüştür. Böyle bir durumda ağı girişindeki etki değeri küçük olan girişlerdeki değişimleri algılayamamaktadır ve güçlü olan giriş değerlerine göre ağı çıkışı şekillenmektedir. Bundan dolayı, giriş değerlerine bağlı olarak ezberleme katmanı kullanılarak ağı girişindeki tüm değerlerin çıkışa daha fazla etki etmesi sağlanmaktadır. Bir diğer durumda, ESA'larda sadece girişteki büyük değişimlere cevap verilmesi istenilebilmektedir. Bu durumlar için literatürdeki bazı çalışmalarda ezberleme katmanı kullanılmamıştır [31].

B. 6. Tam Bağlantı (Fullconnected) Katmanı

Havuzlama veya ezberleme katmanından sonra kullanılan tam bağlantı katmanı, Şekil 6'da gösterildiği üzere kendisinden önceki katmandan gelen tüm bağlantıların toplandığı katmandır. Bu katmandan çıkan veriler tek boyutlu vektör olarak sonuçlandırma katmanına aktarılmaktadır. Bu duruma örnek verilecek olursa, tam bağlantı katmanından önce gelen veriler $16 \times 16 \times 100$ ise ve tam bağlantı katmanının boyutu da 2048×1 olarak seçilirse, toplamda 25600×2048 boyutunda bir ağırlık matrisi oluşur. Bu durumda tam bağlantı katmanının 2048 adet nöron girişine 25600 adet nöron tam olarak bağlanmış olmaktadır.



Şekil 6. Tam bağlantı katmanı

B. 7. Sonuçlandırma Katmanı

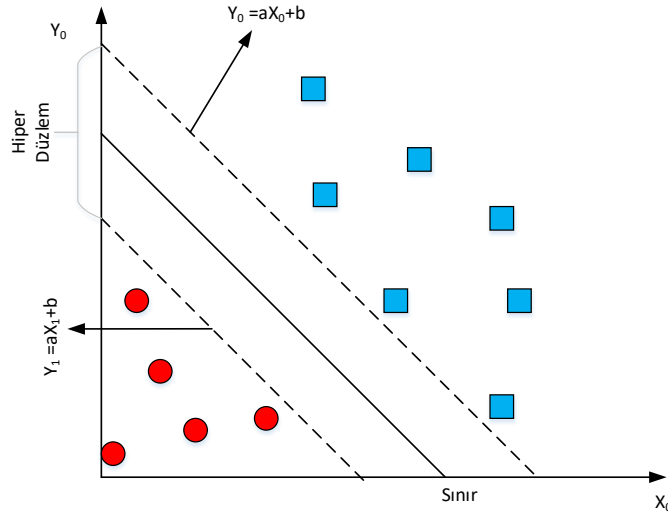
Derin öğrenme algoritmalarının son katmanı sonuçlandırma katmanıdır. Bu katmanda tam bağlantılı katmandan gelen veriler değerlendirilerek ağı çıkışına aktarılmaktadır. Ağı çıkışı sınıflandırma bilgisi olabileceği gibi eğri uydurma yöntemlerindeki polinom katsayıları da olabilmektedir. Eğer gerçekleştirilecek uygulamada sınıflandırma işlemi yapılacaksa, bu katmanın çıkışı sınıf sayısına eşit

olmalıdır. Örneğin 5 farklı nesnenin sınıflandırılması gerçekleştirilecekse, bu katmanın çıkış sayısı 5 olmalıdır. Tam bağlantılı katmandan, sonuçlandırma katmanına gelen değer sayısı, çıkış sayısı ile çarpılarak sınıflandırma ağırlık matrisi elde edilir. Sonuçlandırma katmanı için birçok sınıflandırma fonksiyonu kullanılabilir. Genelde literatürdeki derin öğrenme uygulamalarında, olasılıksal hesaplama gerçekleştiren SoftMax sınıflandırıcısı kullanılmaktadır. SoftMax sınıflandırıcısı ile sınıflandırmada sonucunda 0-1 aralığında değerler üretilmektedir. Üretilen bu değerlerden 1'e en yakın olan değer, sınıf bilgisini vermektedir.

Bazı durumlarda bağlantı katmanından gelen öznitelikler, kullanılan algoritmaya bağlı olarak, farklı bir sınıflandırma algoritmasına gönderilerek de sınıf tahmin bilgisi elde edilebilmektedir. Uygulamada kullanılan sınıf tahmin algoritması ise DVM'dir.

B. 8. Destek Vektör Makinesi (DVM)

DVM, istatistiksel öğrenme teorisi alanında geliştirilmiş bir öğrenme yöntemidir. DVM, ilk olarak verileri lineer olarak ayrılabilirliği yüksek bir boyuta aktarmaktadır. Sonrasında da bunlar arasındaki maksimum sınırın bulunmasını sağlamaktadır [32]. Genel olarak DVM'ler problemleri, lineer olan ve olmayan olmak üzere ikiye ayırmaktadır. Günlük hayattaki birçok problem ise lineer olarak ayrılamamaktadır. Lineer problemlerde DVM'nin kullanım amacı özellikler arasından geçen bir hiper düzlemi bulmaktır. Bu hiper düzlem sınıflara ait özelliklerin birbirine en uzak olduğu iki doğrudan oluşmaktadır. Şekil 7'de bu hiper düzlem üzerinde bulunan doğrular gösterilmektedir.



Şekil 7. Hiper düzlem ve destek vektörleri

Doğrusal bir denklem $ax+b$ şeklinde düşünüldüğünde burada amaç sınıfları ayıracak denklem için a ve b değerlerini hesaplamaktır. Düzlemin ortasından geçen lineer doğru, sınıflar için ayırıcı görev görmektedir [33].

IV. GELİŞTİRİLEN UYGULAMA

A. VERİ ÖN İŞLEME

Gerçekleştirilen uygulama da Wiki Görüntü veri tabanından alınan 62328 adet görüntü üzerinde çalışılmıştır. Görüntü veri tabanından alınan görüntüler ile beraber bir adette MatLab veri dosyası bulunmaktadır. Bu veri dosyasında Tablo 3’de sunulan bilgiler bulunmaktadır.

Tablo 3. Wiki görüntü veri tabanı

Alan Adı	Açıklama
dob	Doğum tarihi
photo_taken	Fotoğrafın çekildiği yıl
full_path	Dosya yolu
gender	Kadın için 0 ve Erkek için 1, Bilinmiyorsa NaN
name	Ünlünün adı
face_location	Yüzün koordinatları
face_score	Dedektör skoru (ne kadar yüksekse o kadar iyidir). Inf, görüntüde hiçbir yüz bulunmadığını
second_face_score	İkinci en yüksek puana sahip yüzün dedektör skoru. Bu, birden fazla yüzü olan görüntüleri yok saymak için kullanılır
celeb_names	Tüm ünlülerin isimleri
celeb_id	Ünlü adının dizini

Veri tabanı ile birlikte kullanılan veri dosyasında ki yüz koordinatlarının bilgisini içeren “face_location” alanındaki bilgilerin tutarsız olduğu tespit edilmiştir. Bu sorunun çözümü için MatLab’de geliştirilen aşağıdaki “face_dedector” algoritması kullanılmış ve elde edilen sonuç Şekil 8’de gösterilmiştir. Veri tabanında ki görüntüler ve veri dosyasından alınan bilgiler doğrultusunda MatLab’de kullanılmak üzere, yeni bir veri dosyası oluşturulmuştur.

```
...  
faceDetector = vision.CascadeObjectDetector();  
bbox = step(faceDetector, res);  
res2 = insertShape(res, 'Rectangle', bbox);  
...
```



Şekil 8. Wiki dosyasından elde edilen yüz koordinat bilgileri ile uygulamadan elde edilen yüz koordinatlarının karşılaştırılması

Oluşturulan yeni veri setine, özellikle görüntüde bulunan yüz koordinatları ve görüntüdeki yüz ifadesi alınan kişinin cinsiyet bilgisi aktarılmıştır.

Veri dosyasında, “face_score” değeri “inf” olan görüntülerde yüz ifadesi bulunmamaktadır, büyüklüğü 20 pikselden küçük olan değerlerde ise yüz olarak tanımlansa bile gerçek görüntüde yüzün olmadığı tespit edilmiştir. Bu nedenle wiki de bulunan veri dosyasından, “face_score” alanındaki değeri “inf” olanlar veya görüntünün büyüklüğü 20 pikselden küçük olan görüntüler, aşağıdaki algoritma ile elimine edilerek yeni veri seti oluşturulmuştur.

```

...
if((face_score~=inf)&(büyüklük>20))
    database(sira,:)=[sira (face_score) wiki.full_path(i) name (bboxPoints) büyüklük (cinsiyet)];
    sira=sira+1;
end
...

```

Tablo 4. face_score değeri “inf” olan veya boyutu 20 pikselden küçük olan görüntüler

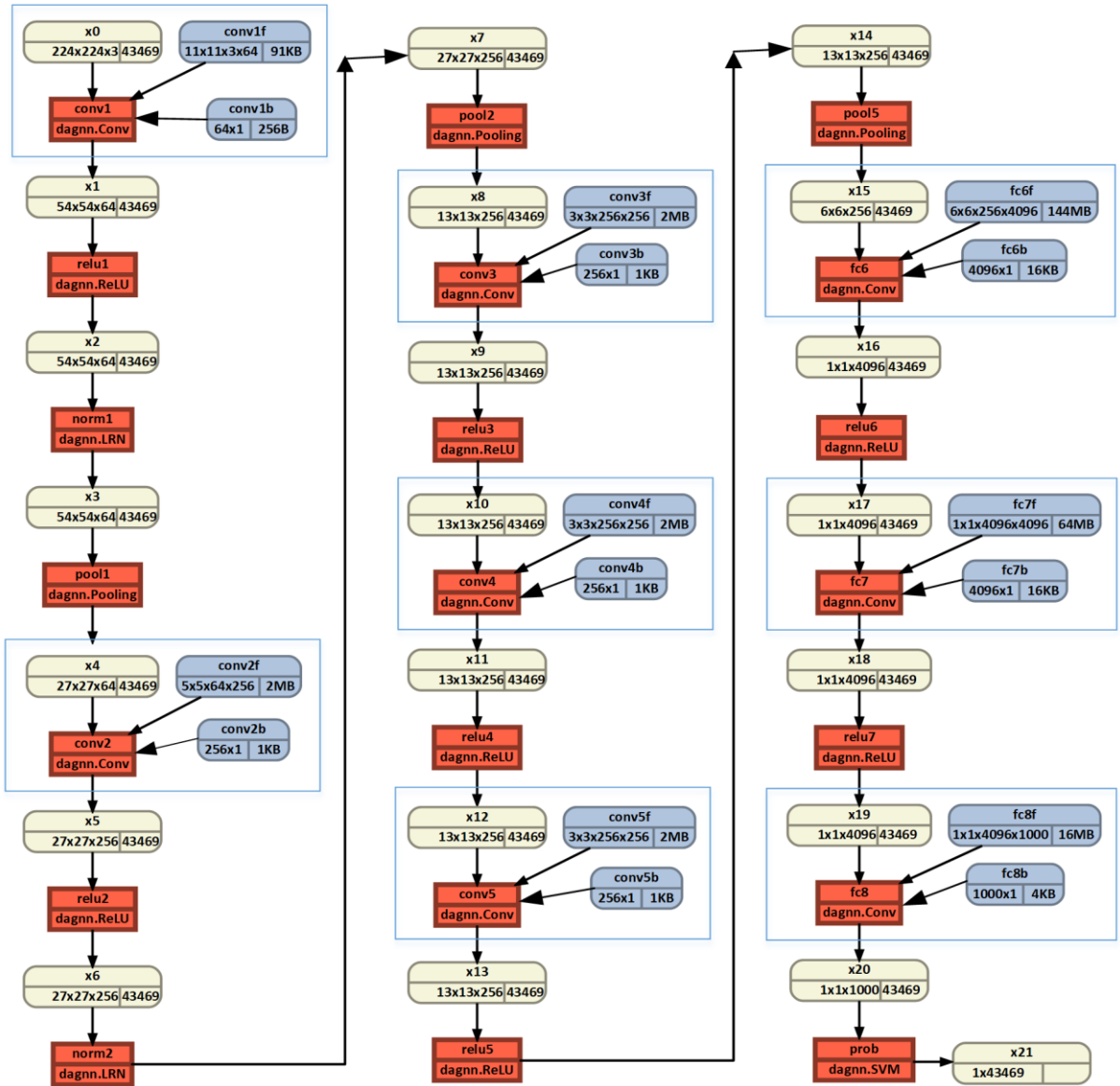
Sıra No	face score	full path	name	face location	piksel	gender	resim
9	-Inf	'13/10004113_1946-08-26_2007.jpg'	'William W. Destler'	[1 1;2 1; 2 2;1 2]	1	1	.
508	1	'02/40059402_1992-03-31_2013.jpg'	'User talk:Kellzchris'	[1 1;19 1; 19 19;1 19]	18	1	

Veri dosyasında, “face_score” değeri “inf” olan veya büyüklüğü 20 pikselden küçük olan örnekler Tablo 4’de gösterilmiştir. Yukarıdaki kod parçacığı ile veri tabanındaki kayıt sayısı 62328’den 43469’e düşürülmüştür. Veri sayısı azalmasına rağmen hazırlanan derin öğrenme algoritmasına yüksek doğruluktaki görüntülerin aktarılması amaçlanmıştır.

Hazırlanan yeni veri setindeki görüntüler, ikinci bir algoritma ile bir dizinin altına “kadın” ve “erkek” olarak iki farklı dosya halinde otomatik olarak kayıt edilmesi sağlanmıştır. Böylece 10757 adet kadın ve 32712 adet erkek görüntüsü elde edilmiştir.

B. DERİN ÖĞRENME MİMARİSİ

Görüntüleme teknolojilerinin gelişimi ile birlikte elde edilen görüntülerin boyutları ve çözünürlükleri de artmaktadır. Görüntüden nesne tanıma, sınıflandırma, hareket algılama gibi işlemlerin gerçekleştirilmesi esnasında birçok görüntü işleme tekniğinin aynı anda çalıştırılması gerekmektedir. Özellikle gerçek zamanlı uygulamalarda hız çok önemlidir. Bu işlemlerin normal bir işlemcide gerçekleştirilmesi uzun zaman almaktadır. Bundan dolayı, grafik kartının gücünü kullanarak, bilgisayarların performansını yüksek oranda arttıran bir paralel programlama platformu olan CUDA geliştirilmiştir. Bu çalışmada, Oxford Üniversitesinin geliştirmiş olduğu [34,35], birçok farklı uygulamada da kullanılan [36,37,38] ve CUDA gibi farklı mimarilerde de çalıştırılabilme imkanına sahip olan MatLab açık kaynak kodlu, MatConvNet-1.0-beta15 mimarisi temel alınmıştır.



Şekil 9. 21 katmanlı derin öğrenme maskeli bölgesel evrimsel sinir ağı.

Kullanılan bu mimaride ki “vl_simplenn” fonksiyonunda aşağıdaki iyileştirmeler yapılmıştır.

- vl_simplenn fonksiyonuna temel normalizasyon katmanının eklenmesi;

İlk olarak sistemde kullanılan görüntülerin üzerinde yüz tespit etme işlemi gerçekleştirilmiştir. Bu aşamadan sonra elde edilen temel yüz görüntüsünde belli bir yerde kümelenmiş renk değerlerinin resimdeki cinsiyeti tespit etmeyi zorlaştırmasından dolayı, vl_simplenn fonksiyonunun başına temel bir histogram eşitleme prosedürü eklenmiştir.

- vl_simplenn fonksiyonunda kullanılan filtrelerin cinsiyet tespitine uygun hale getirilmesi;

Şekil 9’da gösterilen mimaride, 21 katmandan oluşan maskeli bölgesel evrişimsel sinir ağı modeli kullanılmıştır. Çalışmada kullanılan MatConvNet-1.0-beta15 mimarisi genel olarak kullanılan bir mimari olmasından dolayı, farklı uygulamalarda (elma sınıflandırma, karakter tanıma, vb.) farklı filtre içeriklerinin kullanılması gerekmektedir. Bu nedenle, evrişimsel katmanında dolaştırılan filtre boyutları ve içerikleri daha iyi öznitelikler elde edilebilmesi için özelleştirilmiştir. Gerçekleştirilen sistemde kullanılan konvolüsyon filtrelerinden (matris) bazıları Şekil 10’da gösterilmiştir.

8	0	0	0	8
0	8	0	8	0
0	0	8	0	0
0	8	0	8	0
8	0	0	0	8

Conv2f

1	0	1
1	0	1
1	0	1

Conv3f

-1	-1	-1
0	0	0
-1	-1	-1

Conv4f

Şekil 10. Çalışmada kullanılan konvolüsyon filtrelerinden örnekler

Özellikle yüz hatlarındaki kıvrımların tespit edilebilmesi için geliştirilen filtreler (yatay, dikey, çapraz) ile görüntünün elde edildiği ortamdaki aşırı aydınlatma değişimleri gibi görüntü kalitesini bozan çevresel etkilerden, çok daha az etkilenebilecek bir yapı oluşturulmuştur. Bunun sonucu da Tablo 1’de gösterilen aynı veri setini de kapsayan ve hatta içerisinde daha fazla görüntü bulunan, görüntü veri setleri [19, 20] ile eğitilmiş algoritmalara göre daha yüksek başarı oranı elde edilmiştir. Burada literatüre yapılan en önemli katkı ve vl_simplenn fonksiyonunda yapılan en büyük iyileştirme; Geliştirilen sistemin çok fazla sayıda görüntü ile eğitilmesini sağlamak yerine çok daha az görüntüden cinsiyet bilgisini temsil edebilecek özniteliklerin daha iyi filtreler ile tespit edilebilmesini sağlamaktır. Bunun için de sırasıyla,

- 11×11 lik 64 adet (conv1f),
- 5×5 lik 64×256 adet (conv2f),
- 3×3 lük 256×256 adet (conv3f),
- 3×3 lük 256×256 adet (conv4f),
- 3×3 lük 256×256 adet (conv5f) farklı filtre kullanılmıştır.

Kullanılan mimaride beş adet temel konvolüsyon katmanı bulunmaktadır. Konvolüsyon katmanları arasında dönüşümlü olarak da ReLU ve havuzlama işlemleri gerçekleştirilmiştir. Kullanılan mimaride; konvolüsyon katmanlarına ek olarak sekiz adet hesaplama, tam bağlantı, ezberleme ve sonuç katmanı bulunmaktadır.

Her havuzlama katmanında $H(s_1, s_2)$ şeklinde bir filtre kullanılmaktadır. Çalışmada, $H(2,2)$ 'lik en büyük değer (max pooling) matrisi kullanılmıştır. Katman sayısı ise Denklem 7'de d ile ifade edilmiştir.

$$y_{i,j,d} = \max_{1 \leq i' \leq s_1, 1 \leq j' \leq s_2} x_{i+i'-1, j+j'-1, d} \quad (7)$$

Denklem 7'ye göre x giriş matrisinden y çıkış matrisi elde edilmiştir. Buradan elde edilen çıkış matrisi (y), kendisinden sonra gelen matrisin giriş matrisi (x) olmaktadır. ReLU katmanına gelen $x_{i,j,d}$ matrisinin Denklem 8'e göre aktivasyon sonucu hesaplanmaktadır.

$$y_{i,j,d} = \max\{0, x_{i,j,d}\} \quad (8)$$

Temel alınan MatConvNet mimarisinde üç tip normalleştirme tekniği uygulanmaktadır. Bunlar; toplu normalleştirme, özel normalleştirme ve yerel yanıt normalleştirmesidir.

İşlem sonunda 20. katmanda cinsiyet sınıf tahmini için veri bilgisi bulunmaktadır. Gerçekleştirilen uygulamada, 20. katmanda her bir veri için 1000 adet öznitelik olmak üzere toplam $1 \times 1 \times 1000 \times 43469$ adet öznitelik bilgisi bulunmaktadır.

Derin öğrenme algoritması sonucunda elde edilen öznitelikler ve sınıf bilgisi verileri, doğrusal bir sınıflandırıcı olan DVM'ye gönderilerek sınıflandırma işlemi sağlanmıştır. Üretilen sınıf bilgileri 0 ile 1 arasındadır. Bu değerlerin en yüksek olanı görüntünün ait olduğu sınıf bilgisini vermektedir. Bu aşamada DVM'ye girilen verileri güvenilir bir şekilde değerlendirmek için k-fold değeri 5 olarak test edilmiştir. Yapılan tüm testler sonucunda %94,48 başarı oranı ile sınıflandırma işlemi gerçekleştirilmiştir.

C. BAŞARIM ORANI

Hazırlanan yazılım, matematiksel işlemlerin kolaylıkla yapılabildiği ve aynı zamanda açık kaynak kodlu kütüphaneleri bulduran MatLab ortamında geliştirilmiştir. Geliştirilen yazılımda açık kaynak kodlu bir derin öğrenme algoritması temel alınmıştır. Uygulanan derin öğrenme algoritmasının öznitelik çıkarma aşamasında, bölgesel evrimsel sinir ağı modeli kullanılmıştır. Gerçekleştirilen uygulamada MatConvNet-1.0-beta15 mimarisi temel alınıp, filtre yapısında değişiklikler yapılarak yeni bir model geliştirilmiştir. Bu mimari 21 katmandan oluşup her bir örüntü için 1000 adet öznitelik çıkarılmaktadır. Tasarlanan sistemin eğitilebilmesi için Wiki görüntü veri tabanında bulunan görüntüler kullanılmıştır. Çalışmanın konusu görüntü üzerinden cinsiyet tahmini olduğu için konu ile alakası olmayan görüntüler çıkarılarak, sadece “kadın” ve “erkek” görüntülerinin bulunduğu 43469 adet görüntü içeren, yeni bir veri seti hazırlanmıştır. Hazırlanan veri setinde 10757 kadın, 32712 adet erkek görüntüsü bulunmaktadır. Oluşturulan veri setinin yüksek başarı oranı için %80 eğitim ve %20'si test olarak ikiye ayrılmıştır.

Sistemden elde edilen öznitelikler DVM ile sınıflandırılmıştır. Karmaşıklık matrisi (Confusion Matrix) özellikle sınıflandırma algoritmalarının tahmin sonuçlarının doğruluğunu ölçmek için geliştirilen bir yöntemdir. Bu nedenle sınıflandırma sonucunun başarısını ölçmek için karmaşıklık matrisi kullanılmıştır. Elde edilen sonuçlarda Tablo 5'de sunulmuştur.

Tablo 5. Karmaşıklık matris sonuçları

		Tahmin		
		Erkek	Kadın	
Gerçek	Erkek	(TP) 30931	(FN) 1781	32712
	Kadın	(FP) 617	(TN) 10140	10757
		31548	11921	

Denklem 9,10 ve 11'deki;

- True Positive (TP): Algoritma ERKEK çıktısı verdi, gerçek durumda ise ERKEK olmasını,
- False Positive (FP): Algoritma ERKEK çıktısı verdi, gerçek durumda ise KADIN olmasını,
- False Negative (FN): Algoritma KADIN çıktısı verdi, gerçek durumda ise ERKEK olmasını,
- True Negative (TN): Algoritma KADIN çıktısı verdi, gerçek durumda ise KADIN olmasını,

ifade etmektedir.

Doğruluk: Genel olarak, sınıflayıcıların ne sıklıkta doğru tahmin ettiğinin bir ölçüsüdür. Geliştirilen algoritmanın doğruluk oranı Denklem 9'a göre hesaplanmaktadır.

$$\text{Doğruluk Oranı} = (TP + TN) / \text{Toplam Veri Sayısı} \quad (9)$$

$$\text{Doğruluk Oranı} = (30931 + 10140) / 43469 = 0.9448$$

Denklem 9'a göre sistemin doğruluk sonucu %94,48 olarak hesaplanmıştır.

Hassasiyet: Tüm sınıflardan, birinci sınıfın ne kadar doğru tahmin edildiğinin bir ölçüsüdür. Hassasiyet oranı Denklem 10' a göre hesaplanmaktadır.

$$\text{Hassasiyet Oranı} = TP / (TP + FP) \quad (10)$$

$$\text{Hassasiyet Oranı} = 30931 / (30931 + 617) = 0.9804$$

Denklem 10'a göre sistemin hassasiyet sonucu %98,04 olarak hesaplanmıştır.

Özgüllük: Tüm sınıflardan, ikinci sınıfın ne kadar doğru tahmin edildiğinin bir ölçüsüdür Özgüllük oranı Denklem 11' e göre hesaplanmaktadır.

$$\text{Özgüllük Oranı} = TN / (TN + FN) \quad (11)$$

$$\text{Özgüllük Oranı} = 10140 / (10140 + 1781) = 0.8505$$

Denklem 11'e göre sistemin hassasiyet sonucu %85,05 olarak hesaplanmıştır.

Uygulamanın yanlış cinsiyet tahmininde bulunduğu örnek görüntüler Tablo 6’da gösterilmiştir. Geliştirilen uygulamanın hatalı bulduğu görüntü örneklerinin muhtemel sebepleri;





- Tablo 6 örnek 1,2 ve 4’de gösterildiği gibi görüntülerin kontrolsüz olarak alındığı, dış ortam kaynaklı çevresel etkenler (aydınlatma patlaması, aşırı veya çok az aydınlatma, vb.),
- Tablo 6 örnek 2 ve 6’da gösterildiği gibi görüntü de yüz bölgesinde farklı cisimlerin bulunması (mikrofon, el, şapka, gözlük, vb.,)
- Tablo 6 örnek 3 ve 5’de gösterildiği gibi kişilerin baş bölgesinde bulunan bazı (saç, göz, kaş, ağız, burun yapısı gibi) özelliklerin her iki cinsiyette de birbirine yakın olması, sistemin bazı örnekleri hatalı bulmasına neden olmaktadır.
- Bunlara ek olarak örnek 2’de olduğu gibi birden fazla olumsuz etkende (hem yüz bölgesinde bir cisim bulunması hemde aydınlatma kaynaklı olumsuzluk) yanlış sınıflandırmaya neden olabilmektedir.









Bu sorunların çözümü içinde;

- Görüntünün alındığı ortamdaki çevresel şartları, en azından uç noktalardan kabul edilebilir noktalara çekilebilir,
- Kişinin baş bölgesi görüntüsü ile tüm vücut görüntüsünü birlikte değerlendirilerek hatalı sınıflandırma sorunları çözülebilir. Böylece kişinin yüz görüntüsünün içerisinde farklı cisimler dahi bulunsu başarı oranında bir artış olacağı aşikardır.

Ancak her iki durumda farklı bir çalışma konusudur. Gerçekleştirilen çalışma, kontrolsüz ortamlardan elde edilen aynı veri setini kapsayan daha yüksek eğitim verisi ile eğitilmiş sistemler ile kıyaslandığında, %94,48’lik başarı oranı ile yüksek bir sınıflandırma oranına sahiptir.

Tablo 6. Sistemin hatalı bulduğu görüntü örnekleri

Örnek no	Wiki’den alınan görüntü	Görüntüden yüzün tespit edilmiş hali	Uygulamanın bulduğu sonuç	Gerçek sonuç
1			Kadın	Erkek
2			Erkek	Kadın

3			Kadın	Erkek
4			Erkek	Kadın
5			Erkek	Kadın
6			Kadın	Erkek

V. SONUÇ

Gerçekleştirilen uygulamada, kontrolsüz ortamlardan elde edilen görüntüler ile oluşturulan Wiki veri setindeki görüntüler kullanılarak, cinsiyet tahmini üzerine bir çalışma yapılmıştır. Yapılan çalışmada, ilk olarak kullanılan görüntü veri tabanındaki veriler, bir önışlemden geçirilerek, kişilerin yüz görüntülerinin bulunduğu yeni bir veri seti hazırlanmıştır. Derin öğrenme algoritmasında ise, bölgesel evrışimsel sinir ağı modeline dayanan MatConvNet-1.0-beta15 mimarisi temel alınarak, uygulamaya özel yeni filtre yapısı geliştirilmiştir. Cinsiyet tahminin de yüksek başarıım oranı sağlanabilmesi için kullanılan mimaride bulunan filtre boyutları ve içerikleri deęiştirilerek, görüntüdeki yüz bölgesinden cinsiyet sınıflandırılmasında kullanılabilecek öznitelikler elde edilmiştir. Derin öğrenme algoritmasından elde edilen öznitelikler DVM ile sınıflandırılmış ve karmaşıklık matrisi hesaplamasına göre %94,48 başarıım oranı ile cinsiyet sınıflandırma işleminde gerçekleştirilmiştir.

Bundan sonraki yapılacak uygulamalarda, MatConvNet-1.0-beta15 mimarisinin katman sayısı ve farklı sınıflandırma algoritmalarının hibrit olarak kullanılması ve belirli ortamlar altında alınan yüz görüntüleri ile başarıım oranının daha da yukarıya çekilmesi hedeflenmektedir. Buna ek olarak, baş bölgesi ile birlikte tüm vücut görüntüsü alınarak yapılacak bir derin öğrenme modeli ile cinsiyet sınıflandırmadaki hatalarının daha aşağıya indirilmesi de öngörülmektedir. Türkiye koşullarında yüz görüntülerinin belli ortam şartları altında alınmasını sağlamak üzere, hareketli bir platform tasarımının gerçekleştirilmesi de planlanmaktadır.

VI. KAYNAKLAR

[1] Eidinger, E., Enbar, R., & Hassner, T. (2014). "Age and gender estimation of unfiltered faces." *IEEE Transactions on Information Forensics and Security*, vol:9, no.12, pp. 2170-2179.

- [2] Levi, G., & Hassner, T. (2015). "Age and gender classification using convolutional neural networks." *In Proceedings of the IEEE conference on computer vision and pattern recognition workshops* pp. 34-42.
- [3] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). "You only look once: Unified, real-time object detection." *In Proceedings of the IEEE conference on computer vision and pattern recognition* pp. 779-788
- [4] Khalid, S., Sajjad, S., Jabbar, S., & Chang, H. (2017). "Accurate and efficient shape matching approach using vocabularies of multi-feature space representations." *Journal of Real-Time Image Processing*, 13(3), 449-465.
- [5] Chen, Y., Jiang, H., Li, C., Jia, X., & Ghamisi, P. (2016). "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks." *IEEE Transactions on Geoscience and Remote Sensing*, vol: 54, no:10, pp 6232-6251.
- [6] Gao, W., Cao, B., Shan, S., Chen, X., Zhou, D., Zhang, X., & Zhao, D. (2007). "The CAS-PEAL large-scale Chinese face database and baseline evaluations." *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol:38, no: 1, pp 149-161.
- [7] Cas-Peal Face Database [Çevrimiçi]. Erişim: <http://www.jdl.ac.cn/peal/index.html>, Erişim Tarihi: 05.01.2020.
- [8] Imdb-Wiki-500k+ face images with age and gender labels [Çevrimiçi]. Erişim: <https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki/>, Erişim Tarihi:07.01.2020.
- [9] Coco Common Objects in Context [Çevrimiçi]. Erişim:<http://cocodataset.org/#download>, Erişim Tarihi: 07.01.2020.
- [10] Shan, C. (2012). "Learning local binary patterns for gender classification on real-world face images." *Pattern recognition letters*, vol: 33, no:4, pp 431-437.
- [11] Ren, H., & Li, Z. N. (2014, August). "Gender recognition using complexity-aware local features." *In 2014 22nd International Conference on Pattern Recognition* (pp. 2389-2394). IEEE.
- [12] Zhang, N., Paluri, M., Ranzato, M. A., Darrell, T., & Bourdev, L. (2014). "Panda: Pose aligned networks for deep attribute modeling." *In Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1637-1644.
- [13] Hu, S. Y. D., Jou, B., Jaech, A., & Savvides, M. (2011, October). "Fusion of region-based representations for gender identification." *In 2011 International Joint Conference on Biometrics (IJCB) (pp. 1-7). IEEE.Joint Conference on Biometrics (IJCB)*, 2011. IEEE (2011)
- [14] Rai, P., & Khanna, P. (2014). "A gender classification system robust to occlusion using Gabor features based (2D) 2PCA." *Journal of Visual Communication and Image Representation*, 25(5), 1118-1129.

- [15] Leng, X., & Wang, Y. (2008, October). "Improving generalization for gender classification." *In 2008 15th IEEE International Conference on Image Processing* (pp. 1656-1659). IEEE.
- [16] Chen, D. Y., & Lin, K. Y. (2010). "Robust gender recognition for uncontrolled environment of real-life images." *IEEE Transactions on Consumer Electronics*, 56(3), 1586-1592.
- [17] Biswas, S., & Sil, J. (2014). "Gender recognition using fusion of spatial and temporal features." *In Advanced Computing, Networking and Informatics-Volume 1* (pp. 109-116). Springer, Cham.
- [18] Mansanet, J., Albiol, A., & Paredes, R. (2016). "Local deep neural networks for gender recognition." *Pattern Recognition Letters*, 70, 80-86.
- [19] Zhang, K., Gao, C., Guo, L., Sun, M., Yuan, X., Han, T. X., ... & Li, B. (2017). "Age group and gender estimation in the wild with deep RoR architecture." *Special Section on Recent Advantages of Computer Vision Based on Chinese Conference on Computer Vision (CCCV) 2017*, 5, 22492-22503.
- [20] Agbo-Ajala, O., & Viriri, S. (2019). "Face-Based Age and Gender Classification Using Deep Learning Model." *In Pacific-Rim Symposium on Image and Video Technology* (pp. 125-137). Springer, Cham.
- [21] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). "Imagenet classification with deep convolutional neural networks." *In Advances in neural information processing systems* (pp. 1097-1105).
- [22] LeCun, Y., Bengio, Y., & Hinton, G. (2015). "Deep learning. nature," 521(7553), 436-444.
- [23] Goodfellow, I., Bengio, Y., & Courville, A. (2016). "Deep learning." *MIT press*.
- [24] Schmidhuber, J. (2015). "Deep learning in neural networks: An overview." *Neural networks*, 61, 85-117.
- [25] Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S., & Lew, M. S. (2016). "Deep learning for visual understanding: A review." *Neurocomputing*, 187, 27-48.
- [26] Ciresan, D. C., Meier, U., Masci, J., Gambardella, L. M., & Schmidhuber, J. (2011, June). "Flexible, high performance convolutional neural networks for image classification." *In Twenty-Second International Joint Conference on Artificial Intelligence*.
- [27] Pang, Y., Sun, M., Jiang, X., & Li, X. (2017). "Convolution in convolution for network in network." *IEEE transactions on neural networks and learning systems*, vol: 29, no:5, pp: 1587-1597.
- [28] Arora, R., Basu, A., Mianjy, P., & Mukherjee, A. (2016). "Understanding deep neural networks with rectified linear units." *arXiv preprint arXiv:1611.01491*.

- [29] Metlek, S., Özkan, T., (2019). “Analysis of Perceived Service Quality and Customer Satisfaction in the Aviation Sector with Artificial Neural Networks.” *Techno-Science, 2nd Internatioanl Conference on Technology and Science* pp.853-864, 14-16 November, Burdur, Turkey
- [30] Hinton, G. E.“A practical guide to training restricted Boltzmann machines.” *In Neural networks: Tricks of the trade* , pp. 599-619, 2019.
- [31] <https://www.kdnuggets.com/2018/09/dropout-convolutional-networks.html>
- [32] Demirci, D. A. (2007). “Vektör makineleri ile karakter tanıma” (*Yüksek lisans tezi, Bilgisayar Mühendisliği Bölümü, Yıldız Teknik Üniversitesi, İstanbul, Türkiye*).
- [33] Demir, H., Erdoğan, P., & Kekeçoğlu, M. “Destek Vektör Makineleri, YSA, K-Means ve KNN Kullanarak Arı Türlerinin Sınıflandırılması.” *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, c:6, s:1, ss 47-67.
- [34] About MatConvNet [Çevrimiçi]. Erişim:http://www.wisdom.weizmann.ac.il/~vision/courses/2016_2/FCN/matconvnet-fcn-master/matconvnet/doc/site/docs/about.md, Erişim Tarihi: 08.02.2020.
- [35] Dr Andrea Vedaldi [Çevrimiçi]. Erişim:<http://www.robots.ox.ac.uk/~vedaldi/code.html>, Erişim Tarihi: 07.02.2020.
- [36] Verma, G. K., Lather, J. S., & Kaushal, A. “MatConvNet-Based Fast Method for Cervical MR Images Classification.” *In Computational Intelligence: Theories, Applications and Future Directions-* vol:2 pp. 669-679, 2109 Springer, Singapore.
- [37] Hossain, M. A., & Ali, M. M. (2019). “Recognition of handwritten digit using convolutional neural network (CNN).” *Global Journal of Computer Science and Technology*.
- [38] Suthakaran, A., & Premaratne, S. (2020). “Detection of the affected area and classification of pests using convolutional neural networks from the leaf images”. *International Journal of Computer Science Engineering (IJCSE)*.