Facial Emotion Recognition on a Dataset Using Convolutional Neural Network

Vedat TÜMEN
Munzur University
Tunceli Vocational School
Tunceli, Turkey
vtumen@munzur.edu.tr

Ömer Faruk SÖYLEMEZ
Dicle University
Faculty of Engineering
Computer Engineering,
Diyarbakır, Turkey
osoylemez@dicle.edu.tr

Burhan ERGEN
Firat University
Faculty of Engineering
Computer Engineering,
Elazığ Turkey
bergen@firat.edu.tr

***Özet*— Günümüzde derin öğrenme, bilgisayarla görme uygulamaları ve araştırmalarında sıklıkla kullanılan bir tekniktir. Her ne kadar sıklıkla içerik tabanlı görüntü elde edimi uygulamalarında kullanılsa da farklı bilgisayar görmesi alanlarında da kullanımına imkân vardır. Bu çalışmada, FER2013 veri setinde bulunan yüz ifadelerini otomatik olarak sınıflandırmak üzere Konvolüsyonel Sinir Ağları (CNN) tabanlı bir yüz ifadesi tanıma sistemi geliştirilmiştir. Geliştirmiş olduğumuz CNN, FER2013 veri seti üzerinde % 57.1 başarım sağlamıştır.**

***Anahtar Kelimeler*— Derin Öğrenme, Yüz İfadesi Tespiti, İmge Sınıflandırma, Konvolüsyonel Sinir Ağları**

***Abstract*— Nowadays, deep learning is a technique that takes place in many computer vision related applications and studies. While it is put in the practice mostly on content based image retrieval, there is still room for improvement by employing it in diverse computer vision applications. In this study, we aimed to build a Convolutional Neural Network (CNN) based Facial Expression Recognition System (FER), in order to automatically classify expressions presented in Facial Expression Recognition (FER2013) database. Our presented CNN achieved % 57.1 success rate on FER2013 database.**

***Index Terms*—Deep Learning, Facial Expression Recognition, Image Classification, Convolution Neural Networks.**

I. GİRİŞ

Görüntü işleme, hareketli ve hareketsiz imgeler üzerinde düzenleme, çıkarım veya örüntü tanımı yapmak için çalışılan bir alandır. Günümüzde hızla bir yenisi eklenen görüntü işleme teknolojisi sayesinde gerçek görüntülerin işlenmesi ile görüntülerden anlam çıkarılarak doğruya çok yakın sonuçlar elde edilebilmektedir. Sistemlerin gelişimi ile kısa sürede daha kolay, düşük hata oranı ve zaman kaybına uğramadan sonuçlara ulaşılabilmektedir. Derin öğrenme sınıfına ait teknolojiler gidererek artmakta ve makine öğrenme teknolojisi, farklı uygulama alanlarında hayatlarımızı kolaylaştırmaktadır. Örneğin; e-ticaret yaparken ilgi ve kişi ihtiyaçlarımıza göre ürün önermesi, sosyal medyada otomatik resim etiketleme, daha önce yapılan aramalardan yeni aramalar önerme ve göstermesi, mobil cihazlarda ses kontrolü ve konuşma tanıma fırsatı tanımaktadır. Bu özelliklere ilaveten, insanın yüz ifadesinden duygu analizi, kişi tanıma, nesne algılama ve tanıma, doğal dil işleme, tıbbi uygulamalarda, sürücüsüz otomobillerde ve daha birçok alanda kullanılmaktadır [1].

Yüz ifadeleri, günlük sosyal etkileşimlerimizde hayati bir role sahiptir. Yüz ifadelerini, duygularımızı ifade etmek ve başkalarının bize karşı olan duygu ve tutumlarını anlamakta kullanmaktayız. Yüz ifadelerinden anlamlar çıkarmak insanoğlunun henüz birkaç aylıkken öğrendiği ve sahip olduğu bir yetenektir. Makinelerin de insanlar kadar olmasa bile buna yakınsayacak bir şekilde yüz ifadelerini tanımaları hedeflenmektedir. Özellikle İnsan Bilgisayar etkileşimi konusu üzerinde sıklıkla kullanım alanı bulan otomatik yüz ifade analizi için gerçekleştirilen çalışmalar büyük önem arz etmektedir [2].

Tipik bir yüz tanıma sistemi üç aşamadan oluşmaktadır. Bunlar;

Aşama 1: Yüz tespiti ve lokalizasyon,
Aşama 2: Elde edilen yüzlerden özellik çıkarımı,
Aşama 3: Çıkarılan özellikleri kullanarak verilen sınıflara göre yüzleri sınıflandırmak.

Bu çalışmada, çok sınıflı bir yüz ifadesi tespit sistemi için CNN tabanlı bir yaklaşım önerilmiştir. CNN modelini eğitmek ve doğrulamak için FER2013 veri seti tarafından sunulan eğitim ve doğrulama verileri kullanılmıştır. Eğitim ve doğrulama aşamalarından sonra ortaya çıkan CNN, aynı veri setinin test verisiyle doğrulanmıştır. Son aşamada CNN, doğrulama verilerinde % 58.5 ve test verilerinde %57.1 başarım sağlamıştır. Bu veri seti üzerinde insan başarım oranı olan 65 ±5% göz önüne alındığında yeterli bir sınıflandırma başarımı elde edilmiştir [3].

II. DERİN ÖĞRENME

Derin öğrenme, birçok katmanlı ileri beslemeli sinir ağlarının eğitim süreçlerinden oluşmaktadır. Hazırlanan modeller çok sayıda faklı nitelikte gizli katman ile oluşturulduğundan derin öğrenme ismi verilmiştir. Akademik ve özel sektör alanlarında çalışan veri bilimciler hareketli-hareketsiz imge sınıflandırma, imge işleme-düzenleme, video analizi ve sınıflama, ses tanıma ve işleme ve doğal dil öğrenme süreci olmak üzere çeşitli uygulamalarda kullanmaktadır. Derin

978-1-5386-1880-6/17/$31.00 ©2017 IEEE

öğrenme özellikle, büyük miktarlarda, etiketlenmemiş eğitim verilerinden öznitelik çıkarım yöntemleri kullanarak özelliklerin saptamasını yapabilen sistemler oluşturmak için ileri teknoloji yapay sinir ağların kullanılması ile oluşmaktadır.

Derin öğrenme mimarisi birçok katman ve saklı değişkenden oluşur. Derin öğrenmenin en sık kullanılan algoritmaları, Derin Sinir Ağları (Deep Neural Networks), Otomatik Kodlayıcılar (Autoencoders) ve Boltzmann Makinelerinin türevleri olan Kısıtlı Boltzmann Makineleridir [4]. Son zamanlarda özellikle görüntü işleme alanlarında görüntüyü bütün alarak işleyen ve veriyi sınıflara veya özelliklerine ayıran, başarım oranı yüksek olan konvolüsyon yöntemi kullanılmaktadır. Bu çalışmamızda da derin öğrenme yöntemlerinden olan Konvolüsyonel Sinir Ağları (Convolutional Neural Networks – CNN) yöntemi kullanılmıştır.

A. Konvolüsyonel Sinir Ağları(CNN)

1988 yılında Yann LeCun tarafında geliştirilen CNN, çok katmanlı sinir ağlarının özel geliştirilmiş bir türüdür. Farklı mimarilere sahip olmaları yanı sıra klasik yapay sinir ağları gibi ileri yayılım algoritmasına sahiptirler [5]. CNN, önişlem hacmini minimum tutularak piksel görüntülerinden doğrudan görsel kalıpları tanımak için geliştirilmiştir. Ani değişkenliğe sahip desenler ve geometrik dönüşümlere karşı iyi sonuç verir.

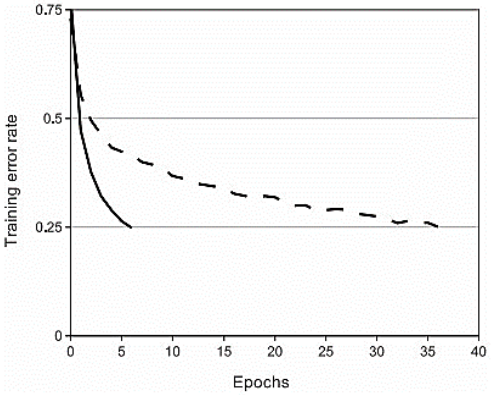
CNN, öznitelik çoğaltma ve özetleme katmanlarını içererek, diğer algoritmalardan farklı bir öznitelik çıkarma işlevini barındırmaktadır. CNN temel olarak 4 katmandan oluşur. Bu katmanlar; konvolüsyon katmanı, aktivasyon fonksiyonu katmanı, pooling katmanı ve normalizasyon katmanları olarak adlandırılmaktadır [6]. Bu katmanlara ek olarak farklı özellikte katmanlar geliştirilmekle beraber son katmanda çok sınıflı sınıflandırma için genellikle softmax katmanı tercih edilmektedir.

1) Konvolüsyonel Katmanlar: Bir görüntüdeki tüm alt bölge için, çıktı özniteliği haritasında tek bir değer üretmek için bir dizi matematiksel işlem gerçekleştirir. Bir imge/videolardaki gerçek görüntüler değişmeme özelliğine sahiptir yani imgeler sabittir. Bu nedenle görüntünün bir bölümünün sayısal değerleri diğer bölümler ile aynı olduğu anlamına gelir. Bu, görüntünün bir bölümünde öğrendiğimiz özelliklerin aynı zamanda resmin diğer bölümlerine de uygulanabileceğimi ve aynı özellikleri tüm konumlarda kullanabileceğimizi göstermektedir.

Gerçek ya da ölçeklendirilmiş bir imgeden rastgele seçilen küçük ölçekli parçalardan daha fazla özellik öğrendikten sonra, öğrenilen bu öznitelik parçasını görüntünün diğer bölgelerine uygulayabiliriz. Bu yöntem ile daha büyük imgeler ile konvolüsyona sokarak görüntüdeki her konumda farklı bir özellik etkinleştirme değeri elde edilebilir.

2) Pooling: Konvolüsyon katmanı kullandıktan sonra elde edilen özellikler sınıflandırılmak istenmektedir [7]. Bu çalışmada kullandığımız FER veri setinde bulunan 48x48 boyutundaki oldukça küçük görüntülere hazırlamış olduğumuz 9x9 girişli ve 1000 özellik öğrenimli bir konvolüsyon katmanını ele aldığımızda; Her konvolüsyon, (48-9 + 1) * (48-9 + 1) = 1600 boyutlarında bir çıktı ile sonuçlanır ve 300 özellik öğrenimli bu ağda örnek başına 1600*1000 = $1.6*10^6$ özellik vektörü oluşturur. Oldukça küçük seçilen görüntü üzerinde bile milyondan fazla özelliği olan girdileri olan bir sınıflandırıcı öğrenmek çok yavaş olur ve aynı zamanda aşırı uyumluluk gösterebilir. Bunun önlenmesi için geliştirilen havuzlama teknikleri ile bu oran azaltılır. Sık kullanılan pooling yöntemi maxPooling ve meanPooling yöntemleridir. maxPooing bir grup piksellerde olan en yüksek piksel değerini alırken meanPooling ise ortalama değeri alarak hesaplama yapar.

3) Aktivasyon Fonksiyonu Katmanı: Belirli bir eşik değerine göre nöronların aktif olup olmamalarını sağlayan fonksiyonlardan oluşur. Aktivasyon işlemi, bir özellik haritasının her bileşenine (diğer bir deyişle noktasal olarak) uygulanan lineer olmayan bir aktivasyon fonksiyonuyla doğrusal bir filtre izlenerek elde edilir. ReLU yöntemi, sigmoid ve tanjant fonksiyonu ile karşılaştırıldığında daha hızlı sonuca ulaşmaktadır. Bu durum özellikle işlem kapasitesi bakımından kısıtlı olan bilgisayarlar için belirgin bir hız artışına olanak sağlamaktadır [8].



Şekil 1. Tanjant fonksiyonu ile ReLU aktivasyon fonksiyonun karşılaştırılması (ReLU, tanjh fonksiyonuna göre 6 kat daha hızlı çalışmaktadır)[8].

ReLU kullanımının en olumsuz yanı, eğitim sırasında bu ünitelerin kırılgan olabilmesi ve bunun da veriye göre olumsuz sonuçlar üretebilmesidir [8]. Örneğin, öğrenme oranı çok yüksek ayarlandığı durumlarda, ağın eğitim veri setinin tamamında etkinleştirilmeyen nöronlar bulunabilir ve bu durumda ReLU üniteleri eğitim sırasında geri döndürülemez. Bu durumun önlenmesi için öğrenme oranının uygun bir şekilde ayarlanması gerekmektedir.

Birçok aktivasyon fonksiyonu bulunmaktadır. Bunlar;

- Maksimum aktivasyon fonksiyonu;
$$f(x) = \max(0;\ x) \quad (1)$$
- Sigmoid aktivasyon fonksiyonu;
$$f(x) = (1 + e^{-x})^{-1} \quad (2)$$
- Hiperbolik tanjant aktivasyon fonksiyonu;
$$f(x) = \tanh(x) \quad (3)$$

olarak verilebilir. Hazırlanan modelde ReLU katmanı için maksimum aktivasyon fonksiyonu kullanılmıştır.

4) Normalizasyon: ReLU katmanları sonucunda oluşabilecek kırılmaları ya da güçlü tepkileri önlemek için normalleştirmeye ihtiyaç duyulmaktadır. Normalleştirme katmanı giriş haritasındaki bütün mekânsal konumdaki özellik kanallarının vektörünü normalleştirmek için kullanılır.

Normalleştirme aşağıdaki denklem 4'de görüldüğü şekilde hesaplanır.

$$b_{x,y}^{i} = a_{x,y}^{i} / \left(k + a_0 + \sum_{j=\max(0,i-n/2)}^{\min(N-1,i+n/2)} \left(a_{x,y}^{j}\right)^2 \right)^{\beta} \qquad (4)$$

Normalizasyonda imgenin düzenlenmesi için 4 parametre ile ayarlama yapılmalıdır. Bu çalışmamızda k = 1, n = 5, $\alpha = 10^{-3}$ ve $\beta = 0.75$ olarak ayarlanmıştır. Bnorm katmanına ise epsilon değeri 10^{-5}, öğrenme oranı 0.05 olarak ayarlanmıştır.

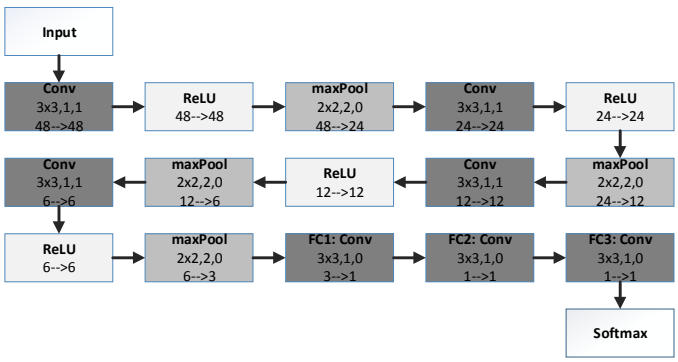
5) Softmax: Softmax regresyonu çok sınıflı verilerin sınıflandırılması için kullanılan başarılı bir sınıflandırma yöntemidir. Özellikle yüksek sayıda sınıf verileri için kullanılır. Normal bir sınıflandırıcıda sınıf sayısı oldukça düşük olmasına karşın Softmax, $y(i) \in \{1, \ldots, K\}$ kadar sınıfa sınıflandırma işlemi yapabilmektedir. Bu çalışmada toplam 7 sınıfta etiketlenmiş veriler bulunmaktadır. Ağımızın son katmanında sınıflandırma için Softmax katmanı kullanılmıştır.

III. DENEYLER

Yüz ifadelerinin tespiti için açık veritabanı olarak paylaşılan toplamda 35887 adet yüz imgeleri [3] kullanılarak veriler 7 yüz ifadesine sınıflandırılmaya çalışılmıştır. FER2013 veri seti üzerinde yapılan literatür taramasında, Liu ve arkadaşlarının [2] dataset üzerinde birçok önişlem yapıldıktan sonra belirlenen CNN modeli üzerinden verilerin tümü eğitilmiş ve doğruluk oranı en yüksek 62.44% olarak tespit edilmiştir. Yine Hong-Wei ve arkadaşlarının yapmış olduğu çalışmasında [9] bir CNN modeli ile aynı veriler üzerinde uygulanmış ve 10% test verisinde 55,6%, doğrulama da ise 48,5% bir başarı oranı elde edilmiştir. Bu çalışmalar incelendiğinde görüntüler model uygulanmadan önce bir dizi ön işlemden geçirildikten sonra hazırlanan modele uygulanmıştır. Bu başarının büyük çoğunluğu verilerin önişlemden geçirilmesinden kaynaklanmaktadır. Bu çalışmamızda [3] verileri üzerinde önişlem yapılmadan gri formatta bulunan 48x48 boyutunda imgeler MatConvNet [10] sistemi üzerinde Şekil 2 de görüldüğü gibi hazırlanan CNN modeline uygulanmış ve başarımları değerlendirilmiştir.

A. Konvolüsyon Sinir Ağı Yapısı(CNN)

CNN modeli Konvolüsyon, ReLU, Pooling ve sonuç olarak softmax katmanlarında oluşan model şekil 2'de gösterilmiştir.



Şekil 2. Hazırlanan ve kullanılan CNN modeli

Şekil 2'de görüldüğü gibi ağ 15 katman ve 4 katman aralara yerleştirilmiş normalizasyon olmak üzere toplamda 19 katmandan oluşmaktadır. Bu model de 4'er konvolüsyon, ReLU, maxPool katmanı ve 3 Full Connect katmanlarından oluşmaktadır. Hazırlanan modelin son katmanında softmax ile imgeler sınıflandırılmaktadır. Bu model kullanılarak elimizdeki kamuya açık gerçek imgeler test edilerek başarımları değerlendirilemeye çalışılmıştır.

B. Kullanılan Yazılım ve Donanımlar

İmge veri tabanında belirlenen imgelerin eğitim ve test işlemler FER2013 veri tabanından sağlanmıştır. Burada MATLAB ortamında imgeler sayısal formata dönüştürülmüştür. MatConvNet sistemi taban alınarak geliştirmiş olduğumuz Şekil 2'deki model kullanılmıştır. Yapılan bu çalışma 2.4 GHz Intel i7 işlemcili, GPU olmayan ve 8 GB RAM bulunan bir bilgisayarda gerçekleştirilmiştir.

C. Veritabanı

FER2013 dataseti insan yüz ifadeleri barındıran gerçek imgelerden oluşmaktadır. Kamuya açık olarak kullanılan bir veri tabanıdır. FER datasetinde toplamda 35.887 adet rastgele bir araya getirilmiş yüz imgesi ve bu imgelerin etiket kısımları bulunmaktadır. Bu yüz imgeleri incelendiğinde içinde 7 sınıfa ait rastgele dizilmiş imgeler bulunmaktadır. Yüz ifadelerinde kızgın, iğrenme, korkma, sevinç, üzgün, şaşırma ve doğal yüz bulunmaktadır. Buradaki verinin 80%'i eğitim, 10%'u doğrulama ve geriye kalan son %10'luk bölümü ise test edilmek üzere seçilmiştir. Şekil 3'te veri setinden örnek imgeler görünmektedir.

Şekil 3.Fer2013 veri setinde bulunan örnek görüntüler.

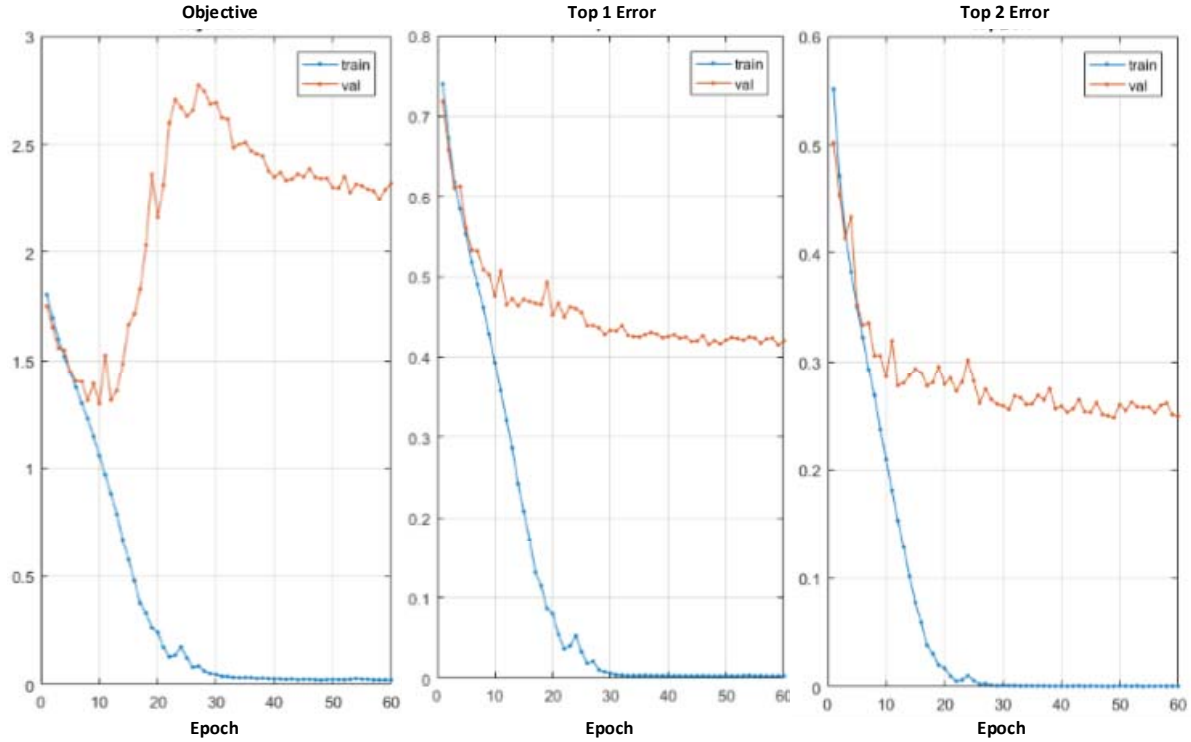
D. Uygulama Sonuçları

Yapılan bu çalışmada yüz imgeleri için toplam da 35887 adet 48x48 boyutlu gri formatta imgeler CNN modelinden geçirildikten sonra elde dilen özellikler ağın son katmanı olan softmax sınıflandırıcı ise sınıflandırılmıştır. Veri setinde farklı yüz ifadelerinden oluşan 7 tür imge aynı zamanda tüm verilerin sınıf bilgileri de bulunmakta ve test işlemi bu değerlere göre yapılmıştır. Aşağıdaki Tablo 1' de bu veri tabanındaki yüz ifadelerin gerçek sayıları ve tüm verilere oranları görülmektedir.

Tablo 1. Datasetteki imge türlerin sayı ve oranları

No	*Yüz İfadesi*	*Sayısı*	*Oranı*
1	Kızgın (Anger)	4.953	0.1380
2	Tiksinme (Disgust)	547	0.0152
3	Korkma (Fear)	5.121	0.1427
4	Mutluluk (Happiness)	8.989	0.2505
5	Üzüntü (Sadness)	6.077	0.1693
6	Şaşkınlık (Surprise)	4.002	0.1115
7	Doğal (Neutral)	6.198	0.1727
	Toplam	35.887	1

Tablo 1 incelendiğinde veri setinde bulunan yüz ifadesi türlerinin eşit oranda dağılmadığı tespit edilmiştir. Tiksinme türündeki imgelerin sayısı 547 ve tüm verilere oranı 0.0152 iken diğer taraftan en çok bulunan imge türü 8989 sayısı ile mutluluk olarak görülmekte ve bunun da veri tabanına oranı 0.2505 olarak ölçülmüştür. Şekil 4'te eğitim ve doğrulama aşamalarında oluşan Top 1 Error ve Top 2 Error grafikleri gösterilmektedir.



Şekil 4. Modelin imgeler üzerinde eğitim ve doğrulama sonuçları

Top 1 Error, hedef sınıfın ilk tahmin edilen sınıf ile uyuşmama oranını göstermektedir. Top 2 Error ise hedeflenen sınıfın tahmin edilen ilk 2 sınıf arasında bulunmama oranını ifade etmektedir. Şekil 4 'te görülmektedir ki 40. epok'tan sonra öğrenme çok az olarak gerçekleşmiştir. Bu nedenle 60. epok'tan sonra eğitim sonlandırılmıştır. Eğitim aşaması sonucunda Top 1 Error %42, Top 2 Error ise %25 olarak gerçekleşmiştir.

Şekil 4 te elde edilen sonuçtan sonra test imgeleri hazırlanan CNN modelinden geçirildikten sonra elde edilen sonuçlar tablo 2'de görülmektedir.

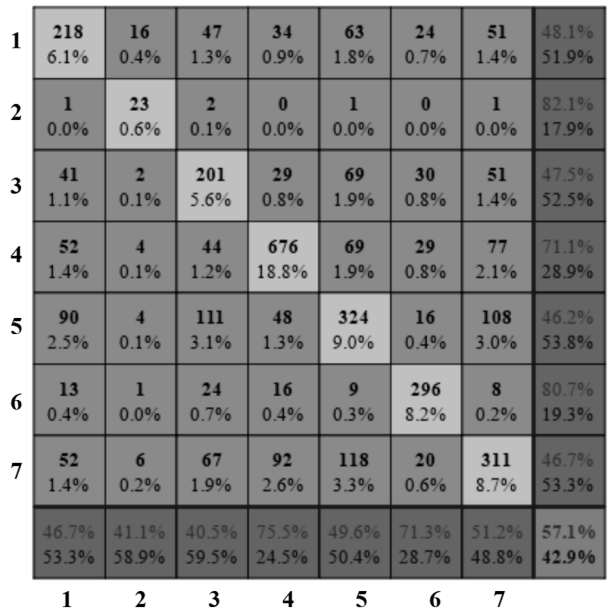
Tablo 2. Test veri seti (%10) için kullanılan imge sayıları ve başarım oranları

No	*Yüz ifadesi*	*Sayısı*	*Kendi veri sayısına oranı*	*Başarım Oranı*
1	Kızgın (Anger)	453	0.091	48.1%
2	Tiksinme (Disgust)	28	0.051	82.1%
3	Korkma (Fear)	423	0.082	47.5%
4	Mutluluk (Happiness)	951	0.105	71.1%
5	Üzüntü (Sadness)	701	0.115	46.2%
6	Şaşkınlık (Surprise)	367	0.091	80.7%
7	Doğal (Neutral)	666	0.107	46.7%

Tablo 2 incelendiğinde FER2013 veri setinde üzerinde hazırlanan yöntem ile eğitilen modelimiz, verilerin kalan son %10'luk kısmı için test edilmiştir. Tabloda da görüldüğü gibi imgeler rastgele dağıtılmış ve sayıları bu doğrultuda çıkmıştır. Burada oranlar baktığımızda en yüksek 0.115 ile 5 nolu üzüntü ifadesi belirten imgeler olurken en düşük 28 tane tiksinme ifadesi belirten imgeler bulunmaktadır. Ortalama dağılım bakıldığında çok büyük farklılar görülmemektedir. Dağılımlar incelendiğinde hemen hemen verilerin %10'una karşılık gelmektedir.

Önerilen model imgelerin komşular arası ilişkilere bakarak tanımaya çalışmaktadır. Model kendini eğitirken ilk önce bu ilişkileri çıkarmakta daha sonra test aşamasında bu belirlemiş olduğu öğrenmeleri kullanmaktadır.

Test işlemi sonucunda elde ettiğimiz sonuç verileri şekil 5'teki karılaştırma matrisine dönüştürülerek sonuçlar daha anlaşılır yapılmaya çalışılmıştır.



Şekil 5. Test işlemi soncunda oluşan karşılaştırma matrisi

Şekil 5 teki karşılaştırma matrisini incelediğimizde en yüksek başarım oranı 2 yani iğrenme türünde imgelerin olduğu tespit edilmiştir. Test için ayrılan veri sayısı oldukça düşük olmasına rağmen başarımı %82.1 olarak tespit edilmiştir. İğrenme imgelerinde yüz hatları biraz daha belirgin olduğu için doğru sınıflandırdığı düşünülmektedir. Yine 4 nolu mutluluk imgelerinde ve 6 nolu şaşkınlık imgelerinde oldukça yüksek başarı oranı tespit edilmiştir. 3 nolu korkma ve 5 nolu üzüntü imgelerinde başarı oranı diğerlerine oranla daha düşük çıkmıştır. Doğal görünüm verileri baktığımızda bu görüntüler insan yüz ifadesine göre değişebilmektedir. Bazı insanların normal yüz ifadesi üzgün ya da mutlu olarak ta görünebilir. Bu yüz ifadesi kişiye özgü bir durumdur. Toplamda 666 adet imgenin 311(%46.7) tanesini doğru sınıflandırırken 118 adet imgeyi üzüntü olarak, 92 adetini de mutluluk olarak yanlış tespit etmiştir. 5 nolu üzüntü imgelerine baktığımızda doğal veriler ile olan benzerlikleri ortaya çıkmış olur. Burada grupta bulunan toplam imgelerin 701 imgenin 108 tanesi doğal yüz ifadesi ve 111 tanesini de korkma yüz ifade olarak yanlış sınıflandırdığı tespit edilmiştir. Korkma yüz ifadesi aynı zamanda üzüntüde belirttiği için burada korlu ifadesini yüksek oranda üzüntü olarak tespit etmesi olağan görülmektedir.

IV. SONUÇ

Bu çalışmamızda gri formatta bulunan kamuya açık gerçek görüntüler üzerinde yüz ifadelerine ait imgeler tasarlanan CNN yöntemine göre sınıflandırması yapılmaya çalışılmıştır. Yapılan çalışmada geliştirilen CNN yönteminin yüz ifadelerini herhangi bir ön işlemden geçirmeden tespit etmesinde iyi sonuçlar verdiği görülmüştür. Özellikle mutluluk, şaşkınlık ve tiksinme yüz ifadelerinde oldukça yüksek başarım elde edilmiştir. Tüm bu başarımlar incelendiğinde % 10'luk verilerin test edilmesi ile %57.1 oranında bir başarı elde edilmiştir.

İleriki zamanda kendimize ait veri setleri oluşturarak yine geliştirilecek olan farklı CNN yöntemleri ile yüksek başarımlı uygulama yapılmaya çalışılacaktır. İmgeler üzerinde yapılan bu çalışmalar hareketli görüntüler üzerinde de uygulanmaya çalışılacaktır.

KAYNAKLAR

[1] LeCun, Y., Bengio, Y., & Hinton, G. 2015. Deep Learning. Nature, 521(7553), 436-444.

[2] Liu, Kuang, Mingmin Zhang, and Zhigeng Pan. "Facial Expression Recognition with CNN Ensemble." Cyberworlds (CW), 2016 International Conference on. IEEE, 2016.

[3] İntetnet Erişimi: https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge

[4] Ian J. Goodfellow, Dumitru Erhan, Pierre Luc Carrier, Aaron Courville, Mehdi Mirza, Ben Hamner, Will Cukierski, Yichuan Tang, David Thaler, and Dong Hyun Lee. Challenges in representation learning: A report on three machine learning contests. *Neural Networks*, 64:117–124, 2014.

[5] Gehring J, Miao, Y, Metze F., "Extracting deep bottleneck features using stacked auto-encoders", Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on May 2013.

[6] Internet Erişimi: http://cs231n.github.io/convolutional-networks/

[7] Internet Erişimi: http://ufldl.stanford.edu/tutorial/ supervised/ Pooling/

[8] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems. 2012.

[9] Ng, Hong-Wei, et al. "Deep learning for emotion recognition on small datasets using transfer learning." Proceedings of the 2015 ACM on international conference on multimodal interaction. ACM, 2015.

[10] İntermet Erişimi: http://www.vlfeat.org/matconvnet/