

ÖZET

1.Giriş

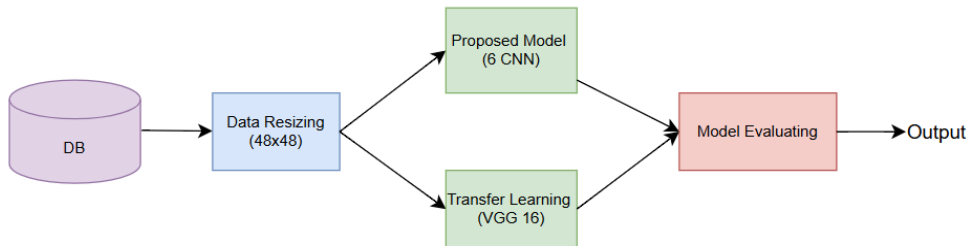
Son yıllarda hızla ilerleyen teknolojik gelişmeler, birçok alanda olduğu gibi yapay zeka ve veri analitiği alanında da önemli dönüşümlere yol açmıştır. Bu gelişmelerin ışığında, yüz ifadelerinden duygu tanıma konusu büyük bir ilgi görmeye başlamıştır. İnsanların ruh halleri ve duygusal durumları, sosyal etkileşimlerden çalışma verimliliğine kadar pek çok alanda kritik bir rol oynamaktadır. Bu nedenle, duyguların doğru ve güvenilir bir şekilde tanınması, hem bireysel hem de toplumsal düzeyde önemli katkılar sağlayabilir.

Büyük veri analitiği ve makine öğrenmesi algoritmalarının birleşimi, yüz ifadelerinden duygu tanıma sistemlerinin geliştirilmesini mümkün kılmıştır. Modern veri işleme teknikleri, büyük miktarda görsel veriyi analiz ederek duygu durumlarını yüksek doğrulukla belirleyebilme yeteneğine sahiptir. Özellikle derin öğrenme algoritmaları, bu alandaki performansı önemli ölçüde artırmış ve insan duygularının otomatik olarak tanınmasını sağlayan sistemlerin geliştirilmesine olanak tanımıştır.

Bu tez çalışmasında, yüz görüntülerinden duygu tanıma üzerine yoğunlaşmış ve çeşitli makine öğrenimi modelleri kullanılarak bu alandaki başarımlar değerlendirilmiştir. Tezimizin amacı, mevcut yöntemleri analiz etmek, performanslarını karşılaştırmak ve duygu tanıma alanında daha etkili ve verimli yöntemler önermektir. Bunun yanı sıra, büyük veri ve derin öğrenme tekniklerinin bu alandaki katkıları da ayrıntılı bir şekilde incelenecektir.

Geliştirilen sistemlerin, psikoloji, insan kaynakları, pazarlama ve güvenlik gibi birçok alanda geniş bir uygulama yelpazesine sahip olması beklenmektedir. Sonuç olarak, bu çalışmanın, hem akademik literatüre hem de pratik uygulamalara değerli katkılar sağlayacağı öngörülmektedir.

CNN ve CNN temelli Transfer öğrenme yöntemi olan VGG16 ile FER2013 isimli yüzden duygu tanıma veri seti üzerinde çalışılmıştır. Şekil 1’de çalışma süreci verilmiştir.



Şekil 1: Çalışmanın Blok Diyagramı

Şekil 2’de görüldüğü resimler önce resize edilmiş sonra TL ve CNN ile model oluşturulup sonuçlar alınmıştır.

2. Literatür Taraması

Bu bölümde çalışmada kullanılan FER2013 veri setine benzer içerikli çalışmalar anlatılmıştır. Amal ve arkadaşlarının FER2013’te yapılan çalışmada %59.1 acc almışlardır [1]. Oguine ve diğerleri yaptıkları çalışmada %70 Acc almışlardır [2]. Mazen ve arkadaşları yaptıkları çalışmada %64 acc almışlardır [3]

3. Materyal Metot

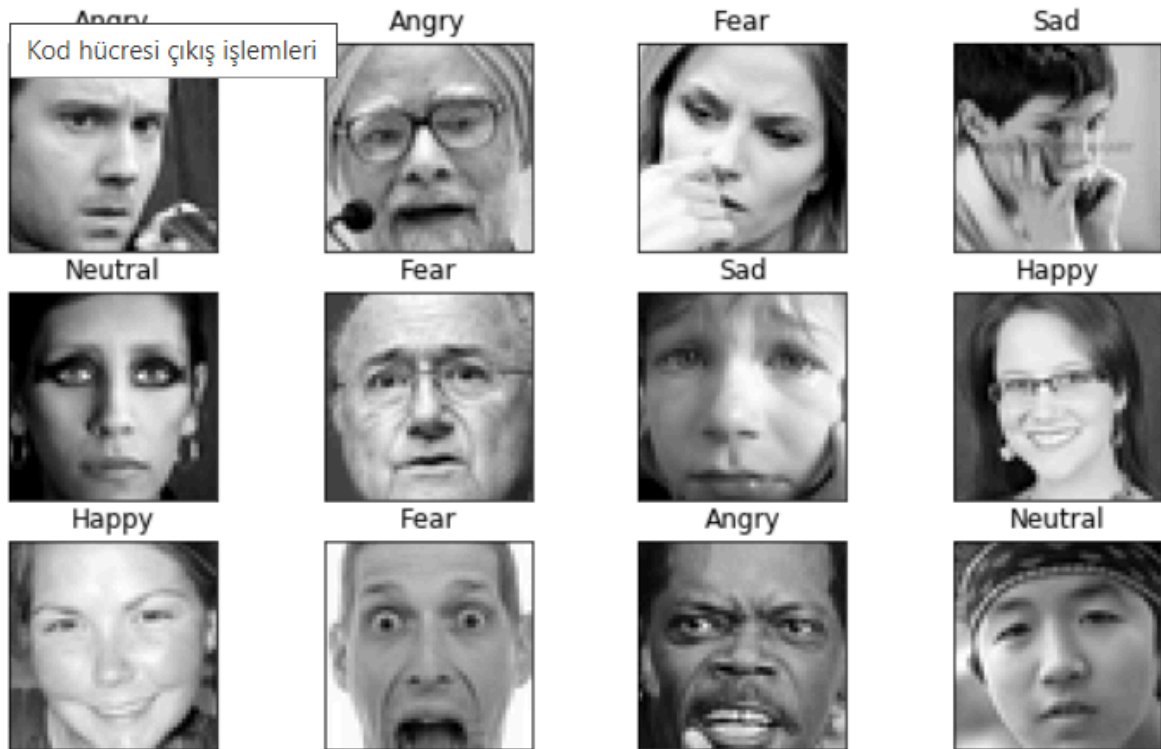
Bu bölümde veri seti, CNN ve CNN temelli transfer öğrenme yöntemleri, model için ilgili değerlendirme parametreleri: Transfer Learning

3.1 Dataset

Veriler yüzlerin 48x48 piksel gri tonlamalı görüntülerinden oluşmaktadır. Yüzler otomatik olarak kaydedilmiştir, böylece yüz aşağı yukarı ortalanır ve her görüntüde yaklaşık aynı miktarda yer kaplar.

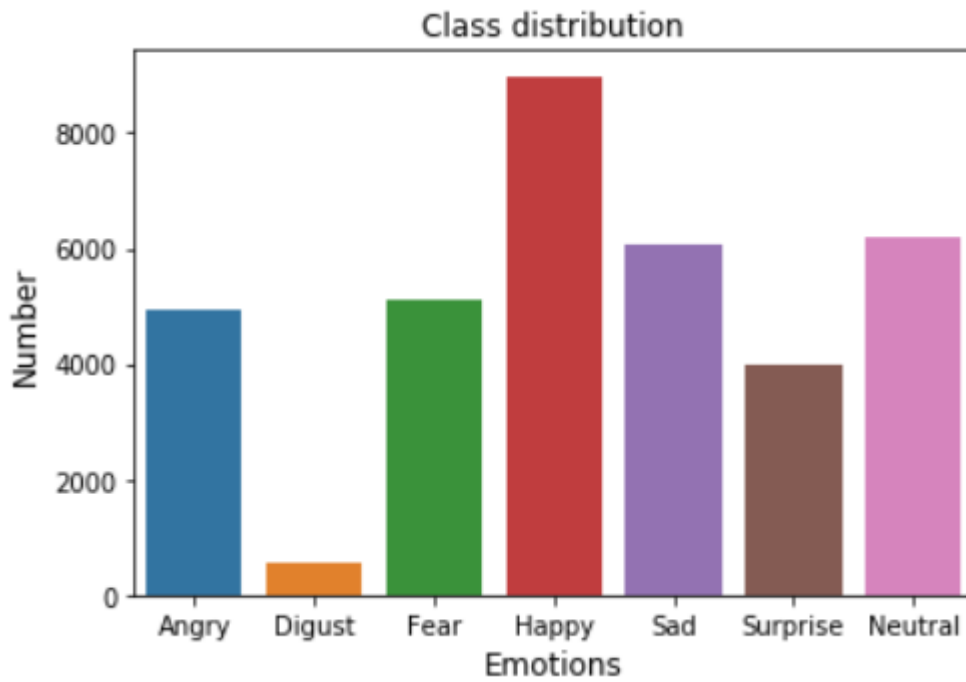
Görev, yüz ifadesinde gösterilen duyguya göre her yüzü yedi kategoriden birine (0=Kızgın, 1=İğrenme, 2=Korku, 3=Mutlu, 4=Üzgün, 5=Şaşkın, 6=Nötr) sınıflandırmaktır. Eğitim seti 28.709 örnekten ve genel test seti 3.589 örnekten oluşmaktadır [5].

Şekil 2’de veri seti görülmektedir



Şekil 2: FER2013 veri seti

Şekil 2’de görülen duygu resimlerinin barplot diagramı şekli 3’te verilmiştir.



Şekil 3: class distribution

Şekil 3’teki gibi bu veri imbalanced db dir. Bu çalışmada ihmal edilmiştir.

Çalışmada kullanılan veri seti train/test ayrımı tablo 1’de verilmiştir

Tablo 1: veri setini train test value ayrımı

| | sad | fear | happy | disgust | neutral | angry | surprise |
|-------|------|------|-------|---------|---------|-------|----------|
| train | 4830 | 4097 | 7215 | 436 | 4965 | 3995 | 3171 |
| test | 1247 | 1024 | 1774 | 111 | 1233 | 958 | 831 |

3.2 Deep learning

Deep learning, makine öğrenmesinin bir alt dalıdır ve özellikle yapay sinir ağlarına dayanan algoritmalarla çalışır. Derin öğrenmenin temel amacı, büyük ve karmaşık veri setlerinden otomatik olarak anlamlı özellikler çıkarabilmektir. Bu yöntem, çok katmanlı sinir ağlarını kullanarak, verilerin temsil edilme şeklini optimize eder.

3.2.1 CNN

Convolutional Neural Networks, derin öğrenmenin bir türüdür ve özellikle görüntü ve video işleme gibi iki boyutlu veriler üzerinde mükemmel sonuçlar verir. CNN'ler, verilerden özellikleri çıkarmak için evrişim (convolution) ve havuzlama (pooling) katmanlarını kullanır.

3.3 Transfer learning

Transfer öğrenimi, CNN gibi bir derin öğrenme modelinin eğitilmesinden elde edilen bilginin alınarak farklı veya benzer bir alanda uygulanması sürecidir. Genellikle büyük boyutta ve farklı veri kümeleriyle eğitilmiş, önceden eğitilmiş bir modelin alınmasını ve yeni bir veri kümesi veya göreve göre ince ayar yapılmasını içerir [6].

3.3.1 VGGNET

VGG, 2014 yılında Oxford Üniversitesi'nden Visual Geometry Group tarafından geliştirilen bir dizi konvolüsyonel sinir ağı (CNN) modelini ifade eder. VGG modelleri, ağırlıklı olarak küçük 3x3 konvolüsyon filtreleri kullanarak derin katmanlı ağlar oluşturur. Bu yaklaşım, modelin derinliğini artırırken aynı zamanda hesaplama verimliliğini de sağlar. VGG19, toplamda 19 katmandan oluşan bir CNN modeliyken VGG16 ise 16 katmandan oluşur [7]

Bu çalışmada VGG16 kullanılmıştır.

3.4 Performans Metriği

Accuracy, precision, recall, ve f-score gibi performans değerlendirme metotları, görüntü işleme gibi sınıflandırma problemleri için türetilmiş olan modelleri değerlendirmek için kullanılıyor. Bu metodlar karışıklık matrisinden elde ediliyor. Karışıklık matrisi Tablo 2'de gösteriliyor [8]

Tablo 2: Confusion Matrix

| | | Actual Values | |
|------------------|--------------|---------------|--------------|
| | | Positive (1) | Negative (0) |
| Predicted Values | Positive (1) | TP | FP |
| | Negative (0) | FN | TN |

Modelin doğruluğu doğru sınıflandırılmış örneklerin tüm örneklere oranıyla gösterilir. Yüksek doğruluk modelin tutarlı tahmin yapmasındaki becerisini gösterir, Düşük doğruluk ise modelin performansı için geliştirmelerin yapılması gerektiğini gösterir. Doğruluk değerinin hesaplanması Denklem (1)'de verilmiştir [9].

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

Doğruluk performans metriği Denklem (5) teki gibi hesaplanmıştır.

Kesinlik(Precision), örneklerden pozitif tahmin edilenlerle gerçekten pozitif olanlar arasındaki orantıdan bahseder. Yüksek kesinlik yanlış tahminlerin az ve pozitif olarak belirlenmiş olan örneklerinin çoğunun gerçekten de pozitif olduğu anlamına gelir. Kesinlik denklem (2)'deki gibi hesaplanır [9].

$$\text{Precision} = T_P / (T_P + F_P) \quad (2)$$

“Recall” sınıflandırma problemlerinde kullanılan bir performans ölçüsüdür. Bütün doğru pozitif örneklerin doğru şekilde pozitif olarak tahmin edildiğindeki orana işaret eder. “Sensitivity” yanlış negatiflerin sayısını düşürmek için ve bulunamayan gerçek pozitif örneklerinin minimuma indirmek için kullanılan önemli bir ölçüdür. Recall Denklem(3)'teki gibi hesaplanır [9].

$$\text{Recall} = T_P / (T_P + F_N) \quad (3)$$

F-score, bilgi kurtarma ve makine öğrenmesi gibi sınıflandırma görevlerinde sıklıkla kullanılan bir ölçüdür. Bir modelin performansını ölçmek, hem precision hem de recall u matematiksel olarak hesaba alan bir dengeyi sağlar. F-score, precision ve recall'un harmonik ortalaması olarak tanımlanır. F-score hesabı Denklem (4)'te verilmiştir. [10].

$$\text{F-score} = (2 * \text{precision} * \text{recall}) / (\text{precision} + \text{recall}) \quad (4)$$

Bu pozitif tahminlerin ne kadar tutarlı olduğunu ölçmeyi amaçlar. Precision'a hassasiyet veya gerçek pozitif oranı da denir. Modelin tüm pozitif örnekleri tespit etmede ne kadar iyi olduğunu gösterir.

F-score bu iki metriği precision ve sensitivity dahil ederek dengeleyerek tek bir değer ortaya çıkartır . Bu sınıf dağılımı dengesizken veya yanlış pozitifler ve yanlış negatifler farklı sonuçlar çıkarınca oldukça kullanışlı olur. F-score özellikle tıbbi tespitle ve veri kurtarma sistemleri gibi hem precision ve sensitivity'nin önemli olduğu durumlarda mühimdir [10].

Çalışma kapsamında TL'den bizim katman sayımıza en yakın olan model ile karşılaştırıldı.

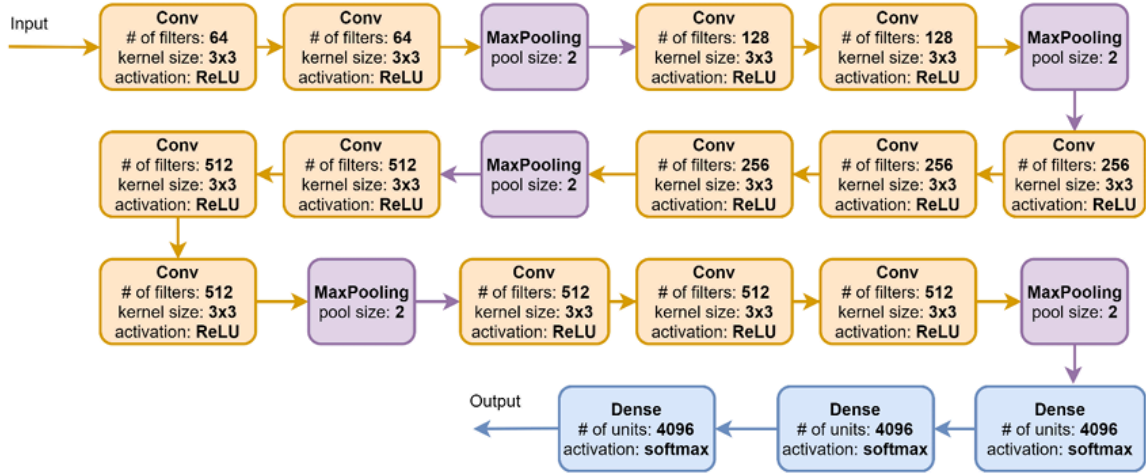
5. DENEYSEL KURULUM VE SONUÇLAR

%80 eğitim %20 test ayrımı sonrası CNN ve VGG16 ile yüzden duygu sınıflandırması için 2 adet model oluşturulmuştur.

```
img_width, img_height = 48, 48 # Size of images
batch_size = 64
epochs = 10
num_classes = 7
```

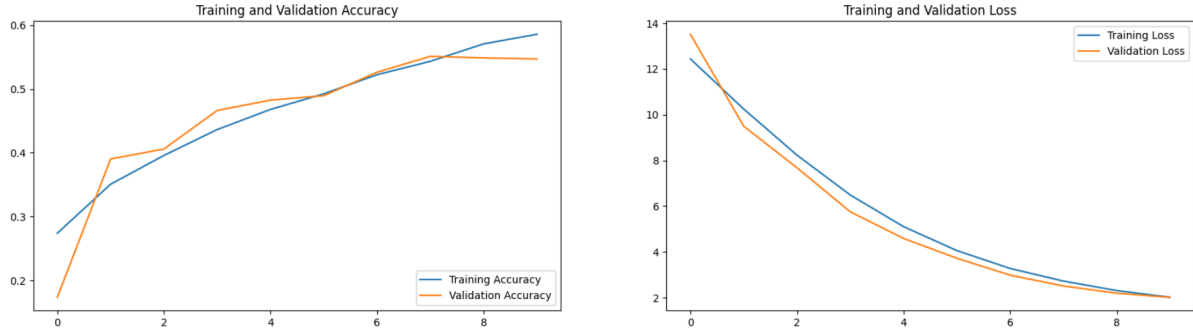
CNN modeli sonuçları

6 katmanlı 2 boyutlu CNN ve sonrasında Flaten, dense dropout ve Batch Normalization katmanında oluşan toplamdan 20 katmanlı bir yöntem önerdik. ŞEkil 3 'te görülmektedir

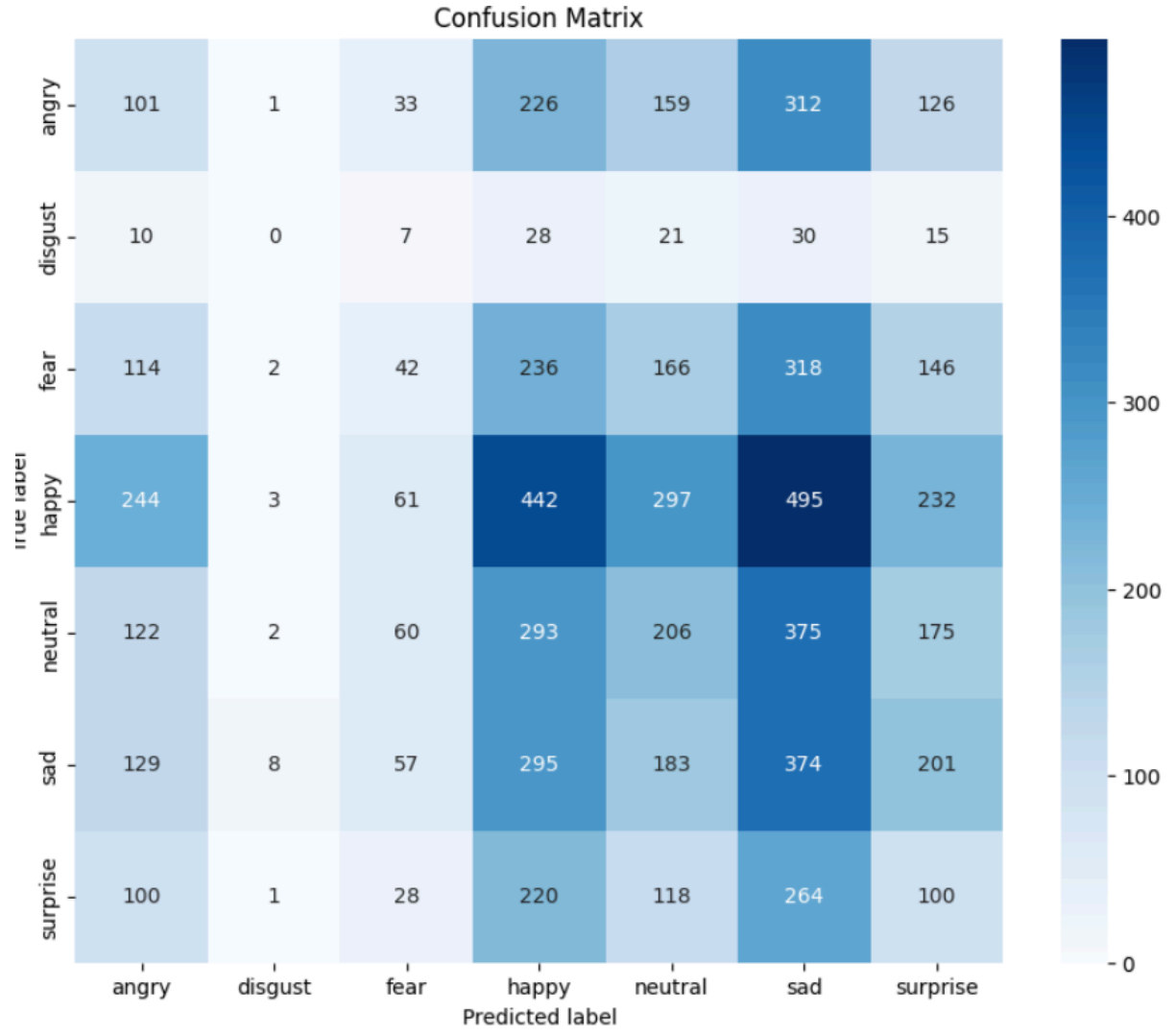


Şekil 3: Önerilen Modelin Gösterimi

Şekil 3 te verilen modelde CNN için kullanılan parametreler: unit 32'den 512'ye kadar nöron, 3x3 kernel, 2 pool dropout 0.25, relu aktivasyon, padding same kullanılmıştır.



Şekil 4: CNN modelinin acc-loss grafikleri

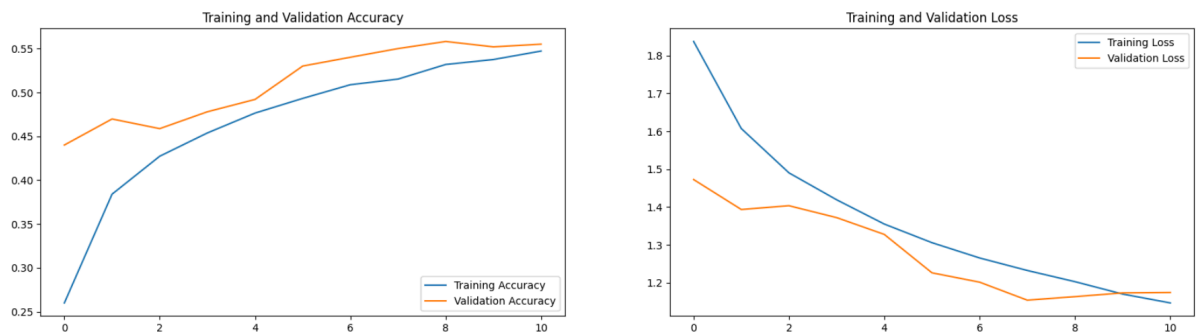


Şekil 5: CNN modelinin confusion matrisi

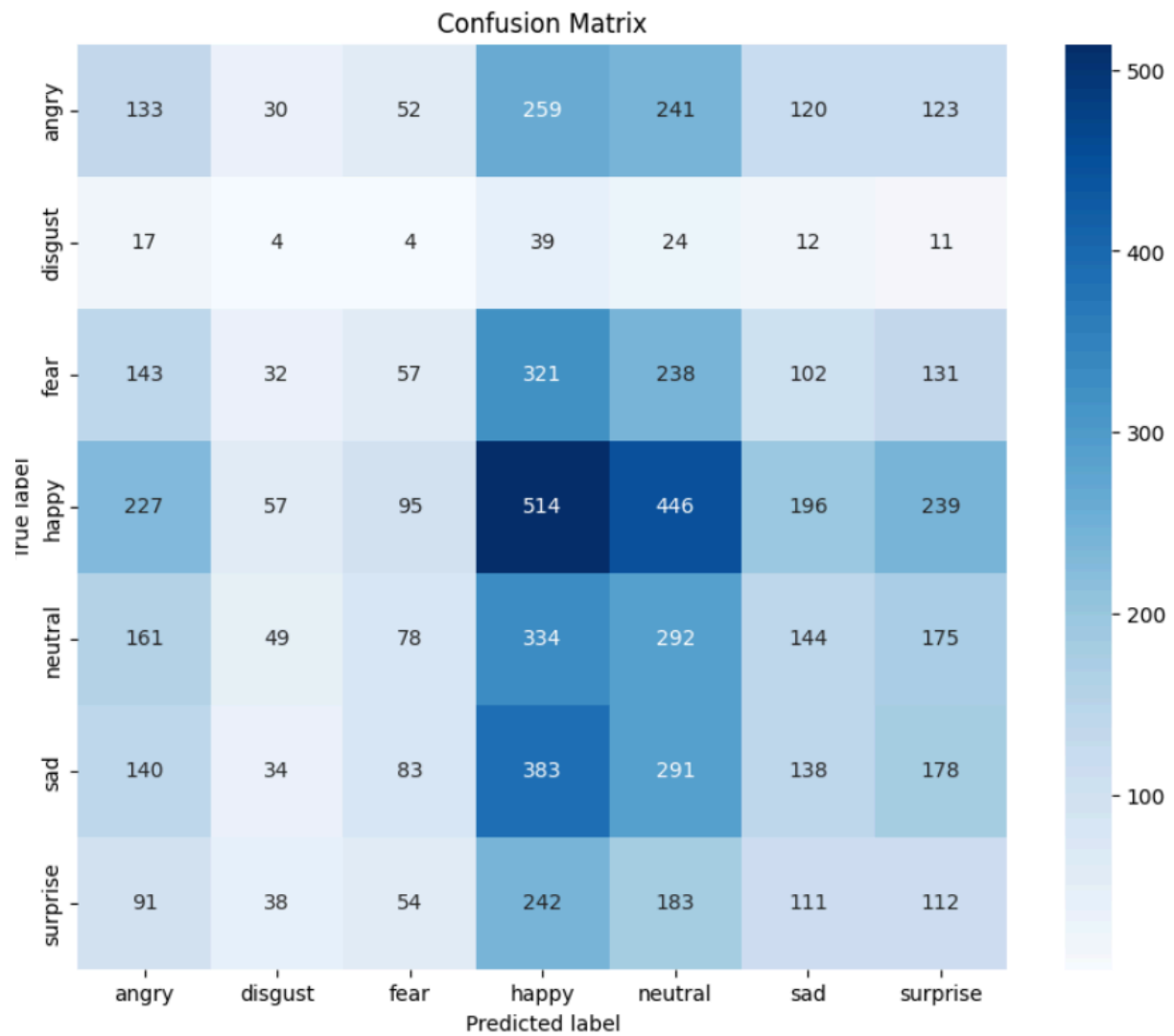
| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| angry | 0.54 | 0.57 | 0.55 | 958 |
| disgust | 0.57 | 0.61 | 0.59 | 111 |
| fear | 0.43 | 0.37 | 0.40 | 1024 |
| happy | 0.89 | 0.79 | 0.84 | 1774 |
| neutral | 0.53 | 0.63 | 0.58 | 1233 |
| sad | 0.55 | 0.39 | 0.46 | 1247 |
| surprise | 0.56 | 0.84 | 0.67 | 831 |
| accuracy | | | 0.61 | 7178 |
| macro avg | 0.58 | 0.60 | 0.58 | 7178 |
| weighted avg | 0.61 | 0.61 | 0.60 | 7178 |

şekil 6: cnn classification report

vgg16 sonuçları



Şekil 7: vgg16 ile acc loss grafiği



şekil 8

| | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0 | 0.60 | 0.58 | 0.59 | 991 |
| 1 | 0.62 | 0.56 | 0.59 | 109 |
| 2 | 0.52 | 0.47 | 0.50 | 1024 |
| 3 | 0.89 | 0.88 | 0.88 | 1798 |
| 4 | 0.55 | 0.54 | 0.55 | 1216 |
| 5 | 0.76 | 0.79 | 0.78 | 800 |
| 6 | 0.60 | 0.68 | 0.64 | 1240 |
| accuracy | | | 0.67 | 7178 |
| macro avg | 0.65 | 0.64 | 0.65 | 7178 |
| weighted avg | 0.67 | 0.67 | 0.67 | 7178 |

şekil 9

5.. SONUÇ VE TARTIŞMASI

FER2013 veri seti özelinde oluşturan :TI ve CNN ile modelin oluşturulmasıdır

Çalışmada elde edilen modellerin literatürle Tablo 3'te verilmiştir.

Tablo 3: literatürle karşılaştırma

| REF. | mODEL | acc |
|--------------|----------|-------|
| [1] | CNN | 59.1 |
| [2] | CNN | 70 |
| [3] | CycleGAN | 64 |
| bu çalışmada | 6-CNN | 62.61 |
| | VGG16 | 67.10 |

6cnn ile alına sonuç state-of-art model olan 16 katmanlı VGG16 ya oldukça yakındır. Bu çalışma katman sayısı arttıkça model performansının artacağını gösteren bir çalışma olmuştur.

Geleceklerde VGG16 dışındaki diğer transfer learning yöntemleri kullanılarak alınacak performanslar karşılaştırılacaktır. Ayrıca detaylı optimizasyonu yapılacaktır. Diğer hedef ise SMOTE veya GAN ile balanced db oluşturularak sonuç alınacaktır. Ayrıca FER 2013 dışında diğer Db ler kullanılacaktır.

6.KAYNAKLAR

[1] Amal, V.S., Suresh, S., Deepa, G. (2022). Real-Time Emotion Recognition from Facial Expressions Using Convolutional Neural Network with Fer2013 Dataset. In: Karuppusamy

[2] P., Perikos, I., García Márquez, F.P. (eds) Ubiquitous Intelligent Systems. Smart Innovation, Systems and Technologies, vol 243. Springer, Singapore. https://doi.org/10.1007/978-981-16-3675-2_41

[3] Oguine, O. C., Oguine, K. J., Bisallah, H. I., & Ofuani, D. (2022). Hybrid facial expression recognition (FER2013) model for real-time emotion classification and prediction. arXiv preprint arXiv:2206.09509.

[4] Mazen, F. M. A., Nashat, A. A., & Seoud, R. A. A. A. (2021). Real time face expression recognition along with balanced FER2013 dataset using CycleGAN. International Journal of Advanced Computer Science and Applications, 12(6)

[5] <https://paperswithcode.com/dataset/fer2013>

- [6] A. M. Dawud, K. Yurtkan, and H. Oztoprak, "Application of Deep Learning in Neuroradiology: Brain Haemorrhage Classification Using Transfer Learning," *Comput Intell Neurosci*, vol. 2019, pp. 1–12, Jun. 2019, doi: 10.1155/2019/4629859.
- [7] Sengupta, A., Ye, Y., Wang, R., Liu, C., & Roy, K. (2019). Going deeper in spiking neural networks: VGG and residual architectures. *Frontiers in neuroscience*, 13, 95.
- [8] Başarslan, M.S., Kayaalp, F. MBi-GRUMCONV: A novel Multi Bi-GRU and Multi CNN-Based deep learning model for social media sentiment analysis. *J Cloud Comp* 12, 5 (2023). <https://doi.org/10.1186/s13677-022-00386-3>
- [9] M. S. Başarslan ve F. Kayaalp, "Sentiment Analysis on Social Media Reviews Datasets with Deep Learning Approach", *SAUC/IS*, c. 4, sy. 1, ss. 35–49, 2021, doi: 10.35377/saucis.04.01.833026.
- [10] M. S. Basarslan and F. Kayaalp, "A Hybrid Classification Example in the Diagnosis of Skin Disease with Cryotherapy and Immunotherapy Treatment," *2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, Ankara, Turkey, 2018, pp. 1-5, doi: 10.1109/ISMSIT.2018.8567247.