



T.C.
KAHRAMANMARAŞ ST İMAM NİVERSİTESİ
FEN BİLİMLERİ ENSTİTS

NİVERSİTE RENCİLERİNİN YZ
İFADELERİNDEN DUYGU DURUMLARININ
YAPAY ZEKA TEKNİKLERİYLE TESPİTİ

MUHAMMED SAME ABDULLAH

DOKTORA TEZİ
ELEKTRİK ELEKTRONİK MHENDİSLİĖİ ANABİLİM DALI

KAHRAMANMARAŞ 2023

T.C.
KAHRAMANMARAŞ SÜTÇÜ İMAM ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ÜNİVERSİTE ÖĞRENCİLERİNİN YÜZ
İFADELERİNDEN DUYGU DURUMLARININ
YAPAY ZEKA TEKNİKLERİYLE TESPİTİ

MUHAMMED ÜSAME ABDULLAH

DOKTORA TEZİ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

KAHRAMANMARAŞ 2023

TEZ BİLDİRİMİ

Tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, ayrıca tez yazım kurallarına uygun olarak hazırlanan bu çalışmada, alıntı yapılan her türlü kaynağa eksiksiz atıf yapıldığını bildiririm

Muhammed Üsame ABDULLAH



Bu çalışma Kahramanmaraş Sütçü İmam Üniversitesi BAP Projesi(2022/4-25M). Çalışmada kullanılan veri seti için Etik Kurul İzni Kahramanmaraş Sütçü İmam Üniversitesi Fen ve Mühendislik Bilimleri Etik Kurul Kararı (karar no. 2019/07, tarih ve 15.06.2021) alınmıştır.

Not: Bu tezde kullanılan özgün ve başka kaynaktan yapılan bildirişlerin, çizelge, şekil ve fotoğrafların kaynak gösterilmeden kullanımı, 5846 sayılı Fikir ve Sanat Eserleri Kanunundaki hükümlere tabidir.

**ÜNİVERSİTE ÖĞRENCİLERİNİN YÜZ İFADELERİNDEN DUYGU
DURUMLARININ YAPAY ZEKÂ TEKNİKLERİYLE TESPİTİ
(DOKTORA TEZİ)**

MUHAMMED ÜSAME ABDULLAH

ÖZET

Yüz ifadeleri, bir kişinin bir işteki performansını etkileyebilecek psikolojik durumunu temsil edebilecek çok çeşitli duygu ve durumları iletebilir. Aynı zamanda, bu insanlar arasındaki sözsüz iletişim için etkili bir stratejidir. Son yıllarda, otomatik yüz ifadesi (FER) tanıma sistemleri oluşturmaya yönelik çalışmalar önemli ölçüde artmıştır. Bu modeller, insan-bilgisayar etkileşimi, sağlık ve güvenlik dâhil olmak üzere çeşitli amaçlar için kullanılabilir. Bu tez çalışmada, yüz yüze veya çeşitli uzaktan eğitim programları kapsamında, üniversite öğrencileri için derslerde yüz tanıma sistemlerine odaklanan bir çalışma sunulmaktadır. Bu çalışmada, FER için derin öğrenme ve görüntü işleme teknikleri kullanılmaktadır. Bu bağlamda, çalışmada k-katlı çapraz doğrulama kullanan transfer öğrenme tekniklerinden ve transfer öğrenme teknikleri yardımıyla da ECOC-MSVM, KNN ve karar ağacı sınıflandırıcılarından faydalanılmaktadır.

Bu noktada, sekiz yüz ifadesini (öfke, tiksinti, korku, mutluluk, tarafsızlık, üzüntü, dikkat ve sürpriz) içeren üniversite öğrencilerinin yüz görüntülerinden oluşan veri setimiz kullanılarak bu çalışma kapsamında deneyler yapılmıştır. Sonuçlar, her iki yöntemin de FER'de iyi ölçüde doğruluk elde edebildiğini göstermektedir. Bununla birlikte, çeşitli önceden eğitilmiş CNN'lerden toplanan çoklu özelliklerin birleştirilmesi ve bu birleştirilmiş özelliklerin, sırasıyla ECOC-MSVM, KNN ve karar ağacı sınıflandırıcılarına uygulanması sonucunda, eğitim-test süresi ile F1-Skor değeri açısından elde edilen sonuçların daha başarılı olduğu gözlemlenmiştir.

Bu noktada, sekiz yüz ifadesini (öfke, tiksinti, korku, mutluluk, tarafsızlık, üzüntü, dikkat ve sürpriz) içeren üniversite öğrencilerinin yüz görüntülerinden oluşan veri setimiz kullanılarak bu çalışma kapsamında deneyler yapılmıştır. Sonuçlar, her iki yöntemin de FER'de iyi ölçüde doğruluk elde edebildiğini göstermektedir. Bununla birlikte, çeşitli önceden eğitilmiş CNN'lerden toplanan çoklu özelliklerin birleştirilmesi ve bu birleştirilmiş özelliklerin, sırasıyla ECOC-MSVM, KNN ve karar ağacı sınıflandırıcılarına uygulanması sonucunda, eğitim-test süresi ile F1-Skor değeri açısından elde edilen sonuçların daha başarılı olduğu gözlemlenmiştir.

Çalışmada, en iyi sonuçlar AlexNet+GoogleNet+ResNet18+ResNet50+TL+KNN kullanılarak elde edilmiştir. Bu durumda, F1-Skoru açısından %99 değerinde bir performansa erişilmiştir ve bu işlemde eğitim-test süresi 35 dakikadır. Ayrıca, AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2+TL+KNN, AlexNet+GoogleNet+ResNet18+ResNet50+TL+ECOC-MSVM ve AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2+TL+ECOC-MSVM kullanıldığında da benzer F1-Skor değerleri elde edilmiştir. Bununla birlikte, eğitim ve test zaman dilimleri, dakika cinsinden sırasıyla 35, 40, 40 ve 44 olmak üzere farklılık göstermektedir.

Anahtar Kelimeler: Yüz ifadeleri, CNN, Derin öğrenme, Transfer öğrenme, ECOC-MSVM, KNN, DT, K-katlı-çapraz doğrulama, Özellik çıkartma

Kahramanmaraş Sütçü İmam Üniversitesi

Fen Bilimleri Enstitüsü

Elektrik Elektronik Mühendisliği Ana Bilim Dalı, Ağustos/2023

Danışman: Prof. Dr. Ahmet ALKAN

Sayfa sayısı: 123

DETECTION OF EMOTIONAL STATES FROM FACIAL EXPRESSIONS OF UNIVERSITY STUDENTS WITH ARTIFICIAL INTELLIGENCE TECHNIQUES

(Ph. D. THESIS)

MUHAMMED ÜSAME ABDULLAH

ABSTRACT

Facial expressions are an effective strategy for non-verbal communication between humans that can convey a wide range of feelings and sensations that may represent a person's psychological state, which may affect his performance at work. In recent years, there has been an increase in interest in building automatic facial expression (FER) recognition systems. These systems can be used for a variety of purposes, including human-computer interaction and health and safety.

In this thesis, a study is presented on facial recognition systems for university students during lectures, whether these lectures are face-to-face or remote lectures through various distance education programs. This study is based on the use of deep learning techniques and image processing for FER. Two methods were investigated: CNNs with transfer learning with k-fold validation, and CNNs with transfer learning using ECOC-MSVM, KNN, or a decision tree.

Experiments were conducted on our own dataset of university students' facial images, which includes eight facial expressions (anger, disgust, fear, happiness, neutrality, sadness, attention, and surprise). The results showed that both methods were able to achieve excellent accuracy in FER. However, Merging multiple features collected from several pre-trained CNNs with transfer learning and applying these combined features to one of the machine learning methods listed following (ECOC-MSVM, KNN, or decision tree) methods used as classifiers outperformed CNNs with transfer learning utilizing the k-fold cross-validation approach in terms of training and testing time, as well as F1-Score value.

The best results were obtained by merging the features extracted when using AlexNet+GoogleNet+ResNet18+ResNet50 +TL+KNN, which recorded 99% in terms of the F1-score and took 35 minutes for training and testing, and it comes with the same value of the F1-score when using AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2 +TL+KNN or AlexNet+GoogleNet+ResNet18+ResNet50+ TL+ECOC-MSVM, and AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2 +TL+ECOC-MSVM. However, the training and testing timeframes differed between them, with the numbers in minutes being 35, 40, 40, and 44, respectively.

Keywords: Facial expressions, CNN, Deep learning, Transfer learning, ECOC-MSVM, KNN, DT, K-fold-cross validation, Feature extraction.

Kahramanmaraş Sütçü Imam University

Graduate School of Natural and Applied Sciences

Department of Electrical Electronics Engineering, August/2023

Supervisor: Prof. Dr. Ahmet ALKAN

Page number: 123

TEŞEKKÜR

Tezimi yöneten ve bu uzun süreçte bana bilimsel, pratik ve psikolojik desteğini esirgemeyen değerli hocam **Prof. Dr. Ahmet ALKAN**'a teşekkür ederim. Bu süreçte kendisinden çok şey öğrendim.

Tez izleme süreçlerinde tezimin ilerleyişinde öneri ve desteklerinden dolayı **Dr. Öğr. Üyesi Erdal KILIÇ** ve **Dr. Öğr. Üyesi Mücahid GÜNAY**' a, jüri üyesi olarak tezimi inceleyerek değerli vakitlerini ayıran **Prof. Dr. Taner TUNCER** ve **Doç. Dr. Halil EROL**' a ve yardımlarından dolayı **Arş. Gör. Kubilay Muhammed SÜNNETCİ** ve **Arş. Gör. Mahmut Nedim EKERSULAR**' a teşekkürü borç bilirim.

Son olarak; bana her zaman destek veren, çalışmalarım sırasında sabır gösteren eşime, anne ve babama şükranlarımı sunarım.

Muhammed Üsame ABDULLAH

İÇİNDEKİLER

ÖZET	i
TEŞEKKÜR.....	iii
İÇİNDEKİLER	iv
SİMGELER VE KISALTMALAR DİZİNİ	vii
ŞEKİLLER DİZİNİ	viii
ÇİZELGELER DİZİNİ	xi
EK ŞEKİLLER DİZİNİ	xiii
1. GİRİŞ	1
1.1. Amaç	2
1.2. Tezin Organizasyonu	3
2. ÖNCEKİ ÇALIŞMALAR	5
3. MATERYAL VE METOT	20
3.1. Veri Setinin Toplanması	20
3.1.1. Veri kümesinin önemi	21
3.1.2. Veri setinin elde edilmesi.....	21
3.2. Derin Öğrenme	23
3.2.1. CNN ağları	24
3.2.2. CNN ağlarının katmanları	29
3.2.2.1. Evrişimsel katman.....	29
3.2.2.2. Padding ve stride	30
3.2.2.3. Havuzlama katmanı.....	31
3.2.2.4. Tam bağlı katman (FC)	31
3.2.2.5. Dropout katmanı.....	32
3.2.2.6. Doğrusal olmayan katmanlar	33
3.2.2.6.1 ReLU fonksiyonu (Rectified Linear Unit-Doğrultulmuş Doğrusal Birim)	
.....	34

3.2.2.6.2 Sigmoid fonksiyonu	34
3.2.2.6.3 Tanh (Tanjant Hiperbolik) fonksiyonu	35
3.2.2.6.4 Softmax fonksiyonu	36
3.2.3. Bazı evrişimli sinir ağı mimarileri	38
3.2.3.1. AlexNet mimarisi	38
3.2.3.2. GoogleNet	39
3.2.3.3. MobileNet V2.....	40
3.2.3.4. ResNet18	40
3.2.3.5. ResNet50	41
3.2.3.6. VGG16	42
3.3. Transfer Learning Algoritması	42
3.4. Makine Öğreniminde Bazı Algoritmalar	44
3.4.1. SVM & MSVM sınıflandırıcıları	44
3.4.2. ECOC-SVM sınıflandırıcı.....	44
3.4.3. K en yakın komşular (KNN) sınıflandırıcı.....	45
3.4.4. Karar Ağacı (DT) sınıflandırıcı.....	46
3.5. Makine Öğreniminde Öğrenme Türleri	47
3.5.1. Gözetimli öğrenme	48
3.5.1.1. Regresyon.....	49
3.5.1.2. Sınıflandırma.....	49
3.5.2. Denetimsiz Öğrenme.....	49
3.5.2.1. Kümeleme	50
3.5.2.2. Birliktelik	50
3.5.3. Reinforcement learning	50
3.6. Kullanılan Bilgisayar Sistemi	50
3.7. K-Kat Çapraz Doğrulama	51
4. BULGULAR VE TARTIŞMA	53

4.1. Transfer Öğrenme ve K-Kat Çapraz Doğrulama Kullanılarak Elde Edilen Sonuçlar	53
4.1.1. İlk kısım için performans ölçümleri	57
4.2. Transfer Öğrenme ve ECOC-MSVM, KNN ve DT Gibi Makine Öğrenmesi Algoritmalarıyla Birlikte Önceden Eğitilmiş Evrişimli Sinir Ağları Kullanılarak Elde Edilen Sonuçlar.....	70
4.2.1. AlexNet tabanlı ECOC-MSVM, KNN ve DT	72
4.2.2. GoogleNet tabanlı ECOC-MSVM, KNN ve DT	76
4.2.3. ResNet18 tabanlı ECOC-MSVM, KNN ve DT	80
4.2.4. ResNet50 tabanlı ECOC-MSVM, KNN ve DT	83
4.2.5. MobileNetV2 tabanlı ECOC-MSVM, KNN ve DT	87
4.2.6. Çoklu CNN ağlarının kombinasyonu ve birleştirme özellikleri.....	90
4.2.6.1. AlexNet ve GoogleNet.....	91
4.2.6.2. AlexNet, GoogleNet ve ResNet18	94
4.2.6.3. AlexNet, GoogleNet, ResNet18 ve ResNet50	98
4.2.6.4 AlexNet, GoogleNet, ResNet18, ResNet50 ve MobileNetV2	102
5. SONUÇ VE ÖNERİLER.....	108
KAYNAKLAR	110
ÖZGEÇMİŞ	122

SİMGELER VE KISALTMALAR DİZİNİ

CNN	: Convolutional Neural Network (Evrişimli Sinir Ağı)
DL	: Deep Learning (Derin Öğrenme)
TL	: Transfer Learning (Transfer öğrenme)
KCV	: K-fold cross validation (k-kat çarpaz doğrulama)
ROC	: Receiver Operating Characteristic
DVM	: Destek vektör makineleri
KNN	: k-Nearest Neighbors (k-En Yakın Komşular)
ECOC-SVM	: Ensemble of error-Correcting Output Codes- Support Vector Machine (Hata Düzeltici Çıkış Kodları- Destek Vektör Makineleri)
DT	: Decision Tree (Karar Ağacı)
AUC	: Area under curve
ML	: Machine learning (Makine öğrenme)
SNS	: Sensitivity (Duyarlılık)
PRC	: Precision (Kesinlik)
SPC	: Specificity (Özgüllük)
REC	: Sensitivity (Duyarlılık)
ACC	: Accuracy (Doğruluk)
F1-Skoru	: F1-Score
TPR	: True positive rate (Doğru pozitif oranı)
FPR	: False positive rate (Yanlış pozitif oranı)
GPU	: Graphics Processing Unit (Grafik İşlemci Ünitesi)
ReLU	: Rectified Linear Unit (Doğrultulmuş Doğrusal Birim)
Tanh	: Tangent Hyperbolic function (Tanjant Hiperbolik fonksiyonu)

ŞEKİLLER DİZİNİ

Şekil 3.1 Aynı öğrenci için veri setinden sekiz yüz ifadesi örneği.....	20
Şekil 3.2 Çeşitli öğrenciler için veri setinden sekiz yüz ifadesi örneği.....	21
Şekil 3.3 A Veri toplama için geliştirilen araç kutusunun örnek ekran görüntüleri.....	22
Şekil 3.3 B Veri toplama için geliştirilen araç kutusunun örnek ekran görüntüleri	23
Şekil 3.4 Derin öğrenmenin yapay zeka haritasındaki yeri	24
Şekil 3.5 CNN'in klasik mimarileri	25
Şekil 3.6 AlexNet CNN ağı için aritmetik mimari	26
Şekil 3.7 Evrimsel Katman	30
Şekil 3.8 Evrimsel Katmanda Padding ve Stride	30
Şekil 3.9 Havuzlama katmanı	31
Şekil 3.10 Tam bağlantılı katman.....	32
Şekil 3.11 Dropout uygulanmamış standart sinir ağı ve dropout uygulandıktan sonra.....	33
Şekil 3.12 ReLU fonksiyonu	34
Şekil 3.13 Sigmoid fonksiyonu	35
Şekil 3.14 Tanh fonksiyonu.....	35
Şekil 3.15 Softmax fonksiyonu	37
Şekil 3.16 CNN ağı için son katmanda softmax fonksiyonunun çalışması.....	37
Şekil 3.17 AlexNet mimarisi	39
Şekil 3.18 GoogleNet mimarisi	39
Şekil 3.19 MobileNet V2 mimarisi	40
Şekil 3.20 ResNet 18 mimarisi	41
Şekil 3.21 ResNet 50 mimarisi	41
Şekil 3.22 VGG16 mimarisi	42
Şekil 3.23 Transfer learning iş akışı	43
Şekil 3.24 Sıfırdan oluşturulan ağ (solda) ile transfer öğrenimi ile oluşturulan ağın (sağda) karşılaştırılması	43

Şekil 3.25 Veri noktaları arasındaki Öklid mesafesi	46
Şekil 3.26 Basit karar ağacı modelinin mimarisi.....	47
Şekil 3.27 Makine öğreniminde öğrenme türleri ve makine öğrenimi algoritmaları	48
Şekil 3.28 GPU yolu.....	51
Şekil 3.29 k-kat çapraz doğrulama	52
Şekil 4.1 Bu bölümde kullanılan algoritmanın akış şeması.....	55
Şekil 4.2 AlexNet+TL+KFCV için konfüzyon matrisi	58
Şekil 4.3 GoogleNet+TL+KFCV için konfüzyon matrisi	58
Şekil 4.4 ResNet18 +TL+KFCV için konfüzyon matrisi.....	59
Şekil 4.5 ResNet50+TL+KFCV için konfüzyon matrisi.....	59
Şekil 4.6 MobileNet +TL+KFCV için konfüzyon matrisi	60
Şekil 4.7 VGG 16 +TL+KFCV için konfüzyon matrisi	60
Şekil 4.8 Bu çalışmada kullanılan altı derin öğrenme yöntemi için performans ölçümlerinin ortalamalarının karşılaştırılması.....	64
Şekil 4.9 Altı CNN modelin sınıflandırmadaki ROC-AUC grafikleri	66
Şekil 4.10 AlexNet tabanlı ECOC-MSVM, KNN ve DT modelleri	73
Şekil 4.11 AlexNet tabanlı ECOC-MSVM, KNN ve DT için konfüzyon matrisleri	74
Şekil 4.12 AlexNet tabanlı makine öğrenimi yöntemleri için ROC eğrileri	76
Şekil 4.13 GoogleNet tabanlı ECOC-MSVM, KNN ve DT modelleri	77
Şekil 4.14 GoogleNet tabanlı ECOC-MSVM, KNN ve DT için konfüzyon matrisleri	78
Şekil 4.15 GoogleNet tabanlı makine öğrenimi yöntemleri için ROC eğrileri	80
Şekil 4.16 ResNet18 tabanlı ECOC-MSVM, KNN ve DT modelleri	81
Şekil 4.17 ResNet18 tabanlı ECOC-MSVM, KNN ve DT için konfüzyon matrisleri.....	81
Şekil 4.18 ResNet18 tabanlı makine öğrenimi yöntemleri için ROC Eğrileri	83
Şekil 4.19 ResNet50 tabanlı ECOC-MSVM, KNN ve DT modelleri.....	84
Şekil 4.20 ResNet50 tabanlı ECOC-MSVM, KNN ve DT için konfüzyon matrisleri.....	85
Şekil 4.21 ResNet50 tabanlı makine öğrenimi yöntemleri için ROC Eğrileri	87

Şekil 4.22 MobileNetV2 tabanlı ECOC-MSVM, KNN ve DT modelleri	88
Şekil 4.23 MobileNetV2 tabanlı ECOC-MSVM, KNN ve DT için konfüzyon matrisleri .	88
Şekil 4.24 MobileNetV2 tabanlı makine öğrenimi yöntemleri için ROC Eğrileri.....	90
Şekil 4.25 ECOC-MSVM, KNN ve DT ile AlexNet+GoogleNet mimarisi	91
Şekil 4.26 AlexNet+GoogleNet tabanlı ECOC-MSVM, KNN ve DT için konfüzyon matrisleri.....	92
Şekil 4.27 AlexNet+GoogleNet tabanlı makine öğrenimi yöntemleri için ROC Eğrileri....	94
Şekil 4.28 ECOC-MSVM, KNN ve DT ile AlexNet+GoogleNet+ResNet18 mimarisi	95
Şekil 4.29 AlexNet+GoogleNet+ResNet18 tabanlı ECOC-MSVM, KNN ve DT için konfüzyon matrisleri	96
Şekil 4.30 AlexNet+GoogleNet+ResNet18 tabanlı makine öğrenimi yöntemleri için ROC Eğrileri.....	98
Şekil 4.31 ECOC-MSVM, KNN ve DT ile AlexNet+GoogleNet+ResNet18+ResNet50 mimarisi.....	99
Şekil 4.32 AlexNet+GoogleNet+ResNet18+ResNet50 tabanlı ECOC-MSVM, KNN ve DT için konfüzyon matrisleri	100
Şekil 4.33 AlexNet+GoogleNet+ResNet18+ResNet50 tabanlı makine öğrenimi yöntemleri için ROC Eğrileri	102
Şekil 4.33 ECOC-MSVM, KNN ve DT ile AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2 mimarisi.....	103
Şekil 4.34 AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2 tabanlı ECOC-MSVM, KNN ve DT için konfüzyon matrisleri	104
Şekil 4.35 AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2 tabanlı makine öğrenimi yöntemleri için ROC Eğrileri.....	106

ÇİZELGELER DİZİNİ

Çizelge 3.1 AlexNet ağındaki her katman için aktivasyon boyutu ve parametre sayısı	29
Çizelge 4.1 GPU kullanılarak çeşitli CNN ağları için eğitim ve test aşamalarında geçen süre	56
Çizelge 4.2 Çeşitli CNN ağları için maksimum parametre sayısı	57
Çizelge 4.3 AlexNet için performans ölçümlerinin değerleri.....	62
Çizelge 4.4 GoogleNet için performans ölçümlerinin değerleri.....	62
Çizelge 4.5 ResNet18 için performans ölçümlerinin değerleri	62
Çizelge 4.6 ResNet50 için performans ölçümlerinin değerleri	63
Çizelge 4.7 MobileNetV2 için performans ölçümlerinin değerleri.....	63
Çizelge 4.8 VGG16 için performans ölçümlerinin değerleri	63
Çizelge 4.9 Mevcut çalışmanın yüz ifadesi tanıma için çeşitli derin öğrenme yöntemlerinin kullanıldığı diğer çalışmalarla karşılaştırılması	68
Çizelge 4.10 GPU kullanılarak çeşitli CNN ağlarına ve birçok makine öğrenmesi algoritmasına ait eğitim ve test aşamalarının zamanları	71
Çizelge 4.11 CNN ağlarının özellik çıkarılan katmanları	72
Çizelge 4.12 AlexNet+ECOC-MSVM için performans ölçümlerinin değerleri	74
Çizelge 4.13 AlexNet+KNN için performans ölçümlerinin değerleri.....	75
Çizelge 4.14 AlexNet+DT için performans ölçümlerinin değerleri.....	75
Çizelge 4.15 GoogleNet+ECOC-MSVM için performans ölçümlerinin değerleri	79
Çizelge 4.16 GoogleNet+KNN için performans ölçümlerinin değerleri.....	79
Çizelge 4.17 GoogleNet+DT için performans ölçümlerinin değerleri	79
Çizelge 4.18 ResNet18+ECOC-MSVM için performans ölçümlerinin değerleri.....	82
Çizelge 4.19 ResNet18+KNN için performans ölçümlerinin değerleri	82
Çizelge 4.20 ResNet18+DT için performans ölçümlerinin değerleri.....	82
Çizelge 4.21 ResNet50+ECOC-MSVM için performans ölçümlerinin değerleri.....	85
Çizelge 4.22 ResNet50+KNN için performans ölçümlerinin değerleri	86
Çizelge 4.23 ResNet50+DT için performans ölçümlerinin değerleri.....	86

Çizelge 4.24 MobileNetV2+ECOC-MSVM için performans ölçümlerinin değerleri	89
Çizelge 4.25 MobileNetV2+KNN için performans ölçümlerinin değerleri.....	89
Çizelge 4.26 MobileNetV2+DT için performans ölçümlerinin değerleri	89
Çizelge 4.27 AlexNet+GoogleNet+ECOC-MSVM için performans ölçümlerinin değerleri	92
Çizelge 4.28 AlexNet+GoogleNet+KNN için performans ölçümlerinin değerleri.....	93
Çizelge 4.29 AlexNet+GoogleNet+DT için performans ölçümlerinin değerleri	93
Çizelge 4.30 AlexNet+GoogleNet+ResNet18+ECOC-MSVM için performans ölçümlerinin değerleri	96
Çizelge 4.31 AlexNet+GoogleNet+ResNet18+KNN için performans ölçümlerinin değerleri	97
Çizelge 4.32 AlexNet+GoogleNet+ResNet18+DT için performans ölçümlerinin değerleri.....	97
Çizelge 4.33 AlexNet+GoogleNet+ResNet18+ResNet50+ECOC-MSVM için performans ölçümlerinin değerleri	100
Çizelge 4.34 AlexNet+GoogleNet+ResNet18+ResNet50+KNN için performans ölçümlerinin değerleri	101
Çizelge 4.35 AlexNet+GoogleNet+ResNet18+ResNet50+DT için performans ölçümlerinin değerleri.....	101
Çizelge 4.36 AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2+ECOC-MSVM için performans ölçümlerinin değerleri	104
Çizelge 4.37 AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2+KNN için performans ölçümlerinin değerleri	105
Çizelge 4.38 AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2+DT için performans ölçümlerinin değerleri	105

EK ŐEKİLLER DİZİNİ

Ek Őekil 1 Etik kurul onay belgesi.	121
--	-----



1. GİRİŞ

Yüz ifadeleri, bir kişinin belirli bir konudaki memnuniyetini veya memnuniyetsizliğini gösteren en önemli kriterlerden birisidir. Aynı zamanda kişinin davranışını ve sonraki eylemlerini ifade edebilir ya da kişinin deneyimlediği psikolojik durumu yansıtabilir. Ayrıca, kişinin sahip olabileceği belirli bir hastalığın varlığı hakkında da bir izlenim verebilir.

Bu nedenle, insan yüz ifadeleri için otomatik tanıma sistemleri, eğitim, psikoloji, güvenlik ve tıp dâhil olmak üzere çeşitli alanlarda kritik öneme sahiptir.

Bu teknolojiler, eğitimcilere öğrencilerin dersi anlayıp anlamadıkları veya ders boyunca başka sorunlar olup olmadığı noktasında yardımcı olabilir. Bu sistemler ayrıca psikolojik ve tıbbi alanlarda bir kişinin psikolojik veya fizyolojik durumunu belirlemeye de yardımcı olabilir. Ayrıca, bu modeller güvenlik alanında belirli bir düşmanca eylemde bulunma olasılığı yüksek olan kişilerin tespit edilmesi hususunda da faydalı olabilir.

Sonuç olarak, bu tür sistemler kolejler ve okullardaki sınıfların yanı sıra havaalanları ve benzerleri gibi yüksek düzeyde güvenlik gerektiren yerlere kurulabilir.

Bu çalışmanın üniversite öğrencilerinin yüz ifadelerini dersler sırasında tanımaya odaklanması nedeniyle, yüz duyguları ile öğrencilerin psikolojik durumlarını ilişkilendirmenin değerini araştıran birkaç eğitim çalışması burada listelenecektir.

Sathik ve Jonathan, sanal öğrenme ortamlarında yüz duygularının öğrencilerin kavrayışı üzerindeki etkisini araştırmaktadır. Çalışmanın amacı, duygusal durumlarla bağlantılı farklı yüz hareketlerini ve bu duyguların öğrencilerin kavrayışıyla nasıl ilişkili olduğunu ortaya çıkarmaktır. Veriler, sanal sınıflardaki öğrencilerin en yaygın sözel olmayan iletişim kanalı olarak yüz ifadelerini kullandıklarını, bu ifadelerin duygularıyla yüksek oranda bağlantılı olduğunu ve bilgilerinin değerlendirilmesine yardımcı olduğunu göstermektedir. Bulgulara göre, sanal sınıflardaki öğretmenler, öğrencilerin kavrayışını değerlendirmek için yüz ifadelerini kullanabilirler. Bu bağlamda, sanal sınıflarda duyguları anlamak için yüz ifadeleri önem arz etmektedir (Sathik ve Jonathan, 2013).

Öğretim becerilerini değerlendirmede yüz ifadesi tanımanın (FER) önemi Utami, Hartanto, & Soesanti'nin makalesinde vurgulanmaktadır. FER'de öğretme becerilerini ölçmenin önündeki düşük kaliteli fotoğraflar, değişen aydınlatma koşulları gibi belirli engeller bu çalışmada vurgulanır. Bu sorunların üstesinden gelmek için bu makale, öğretme ve

öğrenmede FER ile ilgili önceki araştırmaları analiz eder ve özel algoritmalar ve veri kümeleri için gerekliliği vurgular. Ayrıca çalışma, otomatik öğretim sistemlerinde (ATS) FER için CNN ve değiştirilmiş CNN algoritmalarının kullanımını tartışmaktadır. Bu bağlamda, algoritma iyileştirmelerini test etmek için popüler bir seçim olarak CK+ veri seti kullanılmaktadır. Ayrıca, yazarlar ATS'deki birçok yüz ifadesi biçimini araştırmak ve öğretim durumlarında yaşanan zorluklara yanıt sağlamak için daha fazla araştırmaya ihtiyaç olduğunu vurgulamaktadır (Utami ve ark., 2019).

Pise, Vadapalli ve Sanders tarafından yapılan çalışma, yüz ifadesi tanımayı kullanarak, bir e-öğrenme ortamında öğrenme etkisini değerlendirmek için bir metodoloji önermektedir. Yazarlar, yüz ifadelerini ayırt eden ve öğrenme sonuçlarıyla eşleştiren derin bir öğrenme modeli sunmaktadır. Bu model, gerçek zamanlı geri bildirim vererek hem öğrenciler hem de öğretmenler için çevrimiçi öğrenme deneyimini geliştirmeye çalışır. 20 katılımcı içeren bu çalışma, yüzdeki duygular ile öğrenme etkisi arasında bir bağlantı olduğunu savunmakta ve sistemin geri bildiriminin olumlu olduğunu ifade etmektedir. Ayrıca bu çalışma eğitim açısından yüz ifadesi tanıma teknolojilerini kullanmanın potansiyel avantajlarını ve dezavantajlarını da araştırmaktadır. Sonuç olarak, bu teknolojinin e-öğrenme etkinliğini arttırdığı ve daha kişiselleştirilmiş öğrenme deneyimleri oluşturma potansiyelinin olduğu ifade edilebilir (Pise ve ark., 2022).

Yazarlar, bu çalışmada eğitimde duygusal kabulün önemine ve bunun öğrenciler ve öğretmenler için olası faydalarına odaklanmaktadır. Algılama, ayıklama ve sınıflandırma aşamaları dâhil olmak üzere bir yüz ifadesi tanıma sisteminin temel çerçevesi bu çalışma kapsamında tanımlanmaktadır. Bu çalışma, küresel, yerel ve hibrit yaklaşımlar gibi çeşitli yüz tanıma sistemlerini araştırmaktadır. Ayrıca, eğitim uygulamaları açısından farklı yaklaşımlarda burada karşılaştırılmaktadır. Aktif görünüm modelleri, sinir ağları ve evrişimli sinir ağları gibi e-öğrenme sistemlerinde yüz ifadesi tanımlamaya yönelik çeşitli yöntemler kullanılarak çalışma dizayn edilmiştir (Bouhlal ve ark., 2020).

1.1. Amaç

Eğitim alanında derin öğrenme, yüz ifadelerini tanıyarak öğrenciler ve öğretmenler için öğrenme deneyimini geliştirmek amacıyla kullanılmaktadır. Yüz ifadelerini otomatik olarak tanıma becerisine sahip olmak, eğitimcilerin, öğrencilerin konuyla nasıl ilgilendikleri hususunda gerçek zamanlı bir geri bildirim almalarına ve öğretimlerini buna göre

değiřtirmelerine olanak tanır. Böylece, daha yüksek test puanları gibi öğrenci dostu çıktıları elde edilebilir. Eğitimciler, dersi anlamakta güçlük çeken öğrencileri belirlemek için yüz ifadesi tanıma özelliğini kullanabilir. Dolayısıyla öğrencilerin başarılı olmalarına yardımcı olmak için gerekli müdahaleler sağlanabilir. Eğitimciler, bir ders boyunca öğrenci katılımını izlemek için de yüz ifadesi tanıma modellerini kullanabilir. Bu husus, öğrencilerin dikkat etmelerini ve takip etmelerini sağlamada onlara yardımcı olabilir. Otizm spektrum bozukluğu veya dikkat eksikliği hiperaktivite bozukluğu gibi öğrenme bozuklukları olan öğrenciler, yüz ifadesi tanıma kullanılarak belirlenebilir. Böylece gençlerin başarılı olması için gerekli erken müdahaleyi almalarına yardımcı olunabilir.

Ek olarak, yüz ifadesi tanıma, öğrenciler ve öğretmenler arasındaki iletişimi geliřtirmeye de yardımcı olabilir. Bu noktada, bu konunun daha olumlu ve üretken bir öğrenme ortamı yaratma potansiyeli vardır.

Genel olarak, eğitim alanında derin öğrenme kullanılarak yüz ifadelerini tanıma arařtırmaları, öğretme ve öğrenme biçimimizde devrim yaratma potansiyeline sahiptir. Eğitimciler, öğrencilerin nasıl hissettiklerini daha iyi anlayarak onlara başarılı olmaları için ihtiyaç duydukları desteęi sağlayabilir.

Bu tez, önerilen yöntemler ve Kahramanmarař Sütçü İmam Üniversitesi öğrencilerinin yüz ifadelerini içeren verileri kullanan bir çalışma olması açısından özgündür. Aynı zamanda, daha önce hiç kullanılmamış olan sekiz farklı yüz ifadesinin (öfke, dikkat, ięrenme, korku, mutluluk, tarafsızlık, üzüntü ve řaşkınlık) kullanması bakımından da yenilikçidir. Bu yüz ifadeleri, öğrencilerin dersler sırasındaki durumlarını doğru bir şekilde değerlendirebilmek için Abdullah ve Alkan (2022) tarafından yayınlanan arařtırmada kullanılmıştır.

Bu tezin, makine öğrenmesi ve derin öğrenme algoritmaları yardımıyla yüz ifadelerini otomatik olarak belirleyebilmesi hususunda literatüre katkı sağlayacağına inanmaktayız.

1.2. Tezin Organizasyonu

Bu tez çalışması aşağıda anlatıldığı gibi organize edilmiştir: Tezin ilk bölümünde çalışmaya genel bir bakış verilmeye çalışılacak ve çalışma hakkında bazı temel bilgiler sunulacaktır. Burada amacı vurgulanmıştır.

İkinci bölümde çalışma ile ilgili yapılan önceki arařtırmalar, okuyucunun konunun önemini ve literatürdeki yerini anlayabilmesi için özetlenmiştir.

Üçüncü bölüm bu tezde kullanılan veriler hakkında detaylı açıklamanın yapıldığı, çalışmada kullanılan yöntemler hakkında detaylıca teorik bilgilerin yer verildiği bölümdür.

Dördüncü bölüm yapılan çalışmaların anlatıldığı kısımdır. Tez kapsamında Kahramanmaraş Sütçü İmam Üniversitesinden gerekli Etik Kurul izni alarak oluşturduğumuz veri seti kullanılmıştır.

Beşinci ve son bölümde ise yapılan çalışmalarda elde edilen sonuçlar yorumlanmıştır. Çalışmanın literatüre katkısı açıklanmış ve gelecekte yapılabilecek çalışmalar hakkında yorumlar yapılmıştır.



2. ÖNCEKİ ÇALIŞMALAR

Yazarlar (Shan Li ve diğerleri), veri ön işleme, öznelik öğrenme, sınıflandırma ve performans değerlendirmesi gibi yüz ifadesi tanımanın birçok alanına genel bakış sağlamıştır. Çalışmada, statik görüntüler, dinamik görüntü dizileri ve çok modlu duyarlılık analizi gibi farklı bir yüz ifadesi tanıma (Facial Expression Recognition, FER) bileşenine odaklanılmaktadır. Çeşitli yüz hizalama dedektörlerine genel bir bakış sağlamakla birlikte verimlilikler ve performanslar karşılaştırılmıştır. Ayrıca verileri geliştirmek için GAN'lar (Generative Adversarial Networks) kullanılmaktadır. Yazarlar, aydınlatma ve konumdaki dalgalanmalar gibi birkaç FER problemini tartışmakta olup, ön eğitim ve ince ayar metodolojilerinin yanı sıra çeşitli veri kümelerinde performans değerlendirmesi de dâhil olmak üzere özellik öğrenme ve sınıflandırma için çeşitli derin öğrenme mimarilerinden faydalanmaktadır (Li ve Deng, 2022).

Yazarlar (Pritam Sarkar ve diğerleri), elektrokardiyogramlar (EKG'ler) kullanarak duygu tanıma için kendi kendini denetleyen derin çok görevli bir öğrenme çerçevesini önermektedirler. Burada, yöntemler iki aşamaya ayrılmaktadır: EKG temsillerini öğrenme ve duyguları sınıflandırmayı öğrenme. Bir sinyal dönüşümü tanıma ağı, ilk adımda EKG temsillerini öğrenir. Ağ, etiketlenmemiş EKG verilerinden üst düzey soyut temsilleri öğrenir. EKG verileri altı farklı sinyal dönüşümüne tabi tutulur ve dönüşüm tanıma bir görev olarak gerçekleştirilir. Modeli bu görevler üzerinde eğitmek, ağın farklı veri kümeleri ve duygu kategorileri arasında iyi bir şekilde genelleşen uzay-zamansal temsilleri öğrenmesini sağlar. Kendi kendini denetleyen, ağın ağırlıkları, ikinci aşamada bir duygu tanıma ağına iletilir. Bu bağlamda, halka açık dört veri kümesinde duygu tanıma yapmak için EKG kullanılmıştır: AMIGOS, DREAMER, WESAD ve SWELL (Sarkar ve Etemad, 2022).

Zedong Li ve diğerleri, uyarlanabilir bir bulanık semantik çerçeve kullanarak yüz ifadelerini tanımak için bir yöntem önermiştir. Yöntem, yüz özelliklerinden anlamsal kavramların çıkarılmasını ve her bir ifade için kavramların değiştirilmesini içermektedir. Anlamsal yakınlık derecesi daha sonra göze çarpan anlamsal kavramları tanımlamak için kullanılır. Burada, yeni örnekler için sınıf etiketlerini tahmin etmek amacıyla bulanık çıkarım sistemleri kullanılmaktadır. Bulgulara göre ağız, gülümseme ve nötr ifadeleri ayırt etmede önemli bir bölgeyken, göz ve ağız bölgeleri öfke, tikslenme, mutluluk ve şaşırma ifadelerini tanımada ön plana çıkmaktadır (Li ve ark., 2021).

Daksh Maheshwari ve arkadaşları, çok kanallı EEG sinyallerini kullanarak duygu sınıflandırması için ritme özgü derin evrişimli sinir ağı (CNN) önermektedirler. Ayrıca, burada DEAP, DASPS ve DREAMER veritabanları gibi halka açık veri kümelerinden faydalanılmıştır (Maheshwari ve ark., 2021).

Qionghao Huang ve arkadaşları, yüz ifadelerini tanımaya yönelik, grid-wise ve görsel dönüştürücü tekniklerinden yararlanan bir yöntem önermektedir. Yöntemleri, yüz görüntülerinden öznitelikler çıkarmak için önceden eğitilmiş bir CNN ve ifade tanıma ile ilgili bölgelere odaklanmak için grid-wise kullanmaktadır. Ayrıca, yazarlar yöntemlerini CK+, Oulu-CASIA ve RAF-DB (Huang ve ark., 2021) dâhil olmak üzere çeşitli kıyaslama veri kümelerinde test etmişlerdir.

Zhengning Wang ve diğerleri, yüz ifadelerini doğru bir şekilde tanımak için bir dizi yönlendirilmiş mekanizma önermiştir. Bir öznitelik çıkarma ağı, bu yöntemde gözler, burun ve ağız gibi yüzün farklı bölgelerine odaklanmak üzere eğitilmiş çoklu yönelimli mekanizmalarıyla birleştirilmektedir. Bu mekanizmalar, modelin ayrıntıları ve yüz ifadelerindeki küçük değişimleri daha doğru bir şekilde yakalamasını sağlamaktadır. Bu bağlamda, araştırmacılar, CK+, JAFFE ve AffectNet (Wang ve ark., 2021) dahil olmak üzere çeşitli kıyaslama veri kümeleri üzerinde testler yapmıştır.

Yazarlar (Chen ve ark.), RMT-Net mimarisini eş zamanlı yüz işareti ve ifade tanımlaması için lokal bir derin öğrenme mimarisi olarak sunmuşlardır. Bir ön uç ağ, bir yer işareti ve bir yüz ifadesi tanıma dalı ağı oluşturmaktadır. Burada yer işareti, bir ısı haritası regresyon yaklaşımı tahmininden faydalanarak yüzdeki yer işareti konumlarını kullanırken, FER ise, bir softmax sınıflandırıcı kullanarak yüz ölçümünü tahmin etmektedir. Performansı artırmak amacıyla iki kısım yatay ve dikey bağlantılarla birbirine bağlanmıştır. Önerilen yöntem, iki bölümünde her bir özelliğinden faydalanmak için kalıntı öğrenmeyi kullanır. Ağlarını yaygın olarak kullanılan dört veri kümesi üzerinde test etmişlerdir: 300-W, AFLW, AffectNet ve RAF (Chen ve ark., 2021).

Haythem Ghazouani, her bir ifade sınıfı çifti için en ayırt edici özellikleri seçmek ve birleştirmek amacıyla genetik programlamayı kullanmıştır. Bu noktada, yüz ifadelerini tanınması için geliştirilmiş bir yöntem olan GP-FER önerilmektedir. Geometrik ve doku özelliklerini birleştirmek için bir ikili sınıflandırıcı kullanılır. Bu yöntem, üç geometrik ve görünüş özelliği (özellik seçimi, özellik birleştirme ve sınıflandırma) yardımıyla test edilir. Bu çalışmada herhangi bir özelliğin genelleştirilebildiği de ifade edilmektedir. Ayrıca,

önerilen yöntem, DISFA, DISFA+, CK+ ve MUG veri setlerinde doğrulanmıştır (Ghazouani, 2021).

Asit Barman ve diğerleri, mesafe ve şekil imza özelliklerinden yararlanan bir yüz ifadesi tanıma sistemi önermiştir. Sistem, Aktif Görünüm Modeli'ni (AAM) kullanarak yüz görüntülerindeki yer işaretlerini algılamaktadır. Sonrasında bu yer işaretleriyle bir ızgara oluşturarak, bir mesafe imzası elde etmek için ızgaradaki her bir yer işareti noktası çifti arasındaki normalleştirilmiş Öklid mesafelerini hesaplar ve ızgara içindeki üçgenleri tanımlar. Mesafe ve şekil imzaları ayrıca moment, çarpıklık, yığılma ve entropi gibi istatistiksel özellikleri hesaplamak için kullanılır. Çıkarılan özellikler daha sonra ifade kategorizasyonu için kullanılır. Sistemleri JAFFE, CK+, MMI ve MUG (Barman ve Dutta, 2021) dâhil olmak üzere çeşitli veri tabanlarında test edilmiştir.

Yazarlar (Xiao Sun ve diğerleri), yüzeysel ve derin özelliklerin yanı sıra dikkat mekanizmasını da içeren dinamik sıralı bir yüz ifadesi algılama sistemi önermişlerdir. Sistemde üç model vardır: ASModel (Attention Shallow Model, ASModel), ADModel (Attention Deep Model, ADModel) ve MSDModel (Multi-attention Shallow and Deep Model, MSDModel). Burada, FACS'ı (Facial Action Coding System, FACS) karakterize etmek için; ASModel, yüzdeki yer işaretlerinin göreceli konumuna ve yüzün ilgili alanının dokusal özelliklerine dayalı olarak özellikleri çıkarır. Çalışmada LSTM (Bidirectional LSTM, BiLSTM) kullanımı yoluyla verimliliğin artırılması hedeflenmektedir. Ayrıca, ADModel, sıralı yüz görüntülerinden derin özellikler çıkarmak ve AlexNet yapısını iyileştirmek için tasarlanmıştır. MSDModel, bir ağırlık-dikkat mekanizması kullanarak ASModel ve ADModel'in çıktılarını birleştirir. Sonrasında ise, bir kayıp fonksiyonu ve türev hesaplaması kullanılarak eğitilir. Yaklaşımları, halka açık üç veri tabanında test edilmiştir: CK+, MMI ve OuluCASIA (Sun ve ark., 2021).

Yazarlar (Türker Tuncer ve ark.) EEG sinyallerine dayalı bir duygu tanıma sistemi önermektedirler. Önerilen yöntem, FFP (Fractal Firat Pattern, FFP) ve TQWT (Tunable Q-factor Wavelet Transform, TQWT) sinyal ayrıştırma yaklaşımlarıyla çok düzeyli özellikler yaratan bir fraktal model özellik çıkarma stratejisini kullanır. Bu yaklaşım ayrıca, özellik seçimi için geliştirilmiş yinelemeli bir seçici ve kategorizasyon için sığ sınıflandırıcıları kullanmaktadır. Teknikleri GAMEEMO veri seti üzerinde test edilmiştir (Tuncer ve ark., 2021).

Araştırmacılar (Jinzhaoh Zhou ve diğerleri), FMN (Frequency Multiplication Network, FMN) olarak adlandırılan yüz ifadesi tanıma (FER) için yeni bir derin öğrenme mimarisi önermektedirler. FMN, HOG (Histogram of Oriented Gradients, HOG) ve ULBP (Uniform Local Binary Pattern, ULBP) tanımlayıcılarının yanı sıra Ayrık Kosinüs Dönüşümü (Discrete Cosine Transform, DCT) kombinasyonu bu çalışmada kullanılır. Otomatik filtreleme, ortalama havuzlama ve yığın normalleştirme için FMN'ye bir çoğaltma katmanı dâhil edilmiştir. Teknikleri, herkese açık üç veri kümesi kullanılarak test edilmiştir: CK+, Oulu-CASIA ve MMI (Zhou ve ark., 2021).

Tingting Liu ve diğerleri, kızılötesi teknolojiyi sınıf bağlamında kullanarak yüz duygularını tanımlamak için yeni bir teknik sundular. Önerilen teknik, çok kişili durumlarda kullanılabilir ve bir Cauchy dağıtım tabanlı etiket öğrenme ağı kullanılarak IR ifade fotoğraflarından yüz özelliklerini öğrenmektedir. Yaklaşımları, GoogleNet omurga yapısı, kovaryans havuzlama katmanı ve KL kaybından oluşan CDLLNet mimarisini kullanmaktadır. Sınıfta sözlü olmayan davranışları anlama önerilen CDLLNet yaklaşımı kullanılarak gerçekleştirilebilir. Burada, önerilen ağları değerlendirmek için iki yüz ifadesi veri seti, Oulu-CASIA ve CK+ kullanılmıştır (Liu ve ark., 2021).

Yazarlar (Lee ve diğerleri), gömülü cihazlarda gerçek zamanlı yüz duygu sınıflandırması için derin bir evrişimli sinir ağı mimarisi olan EmotionNet Nano'yu önermişlerdir. Ayrıca, mimarinin oluşturulması için, insan tarafından belirlenen kısıtlamalar altında ağlar oluşturan üretken makineleri öğrenmek için üretken sentezi birleştiren bir insan-makine işbirliği stratejisi kullanılmıştır. Mimari, yüksek derecede makro mimari ve mikro mimari heterojenliğe, seçici uzun menzilli bağlantıya sahiptir ve model ifadesini korurken mimari ve hesaplama karmaşıklığı arasında güçlü bir denge sağlamaktadır. EmotionNet Nano mimarisi, en azından bir kat daha küçük olmakla birlikte, son teknoloji ağ topolojileriyle karşılaştırılabilir doğruluk sağlamaktadır. Yüksek hızı ve enerji ekonomisi ile ağ, gömülü senaryolar için idealdir. Ağı eğitmek ve test etmek için CK+ veri seti kullanılmıştır (Lee ve ark., 2021).

Yazarlar (Ab Wahab ve diğerleri), EfficientNet-Lite ve hibrit CNN-KNN modellerini birleştirerek bir yüz ifadesi algılama sistemini nasıl oluşturduklarını bu çalışmada sunmaktadırlar. Sistem, gerçek zamanlı yüz ifadesi tanınması için kullanılabilen, düşük maliyetli ve taşınabilir bir cihaz olan Raspberry Pi üzerinde çalışacak şekilde oluşturulmuştur. Yazarlar, sistemin performansını değerlendirmek için FER-2013 veri setini

kullanarak deneyler yapmışlardır. Özellikleri çıkarmak için CNN'yi ve yüz ifadelerinin fotoğraflarını sınıflandırmak için de KNN'yi kullanmışlardır (Ab Wahab ve ark., 2021).

Yazarlar (Shervin Minaee ve diğerleri) yüz ifadelerini tanımak için bir yöntem sunmuşlardır. Bu yöntem, iki bölümden oluşan dikkat evrişimli ağ (Attentional Convolutional Network, ACN) modelini kullanır: bir özellik çıkarma ağı ve bir dikkat ağı. Özellik çıkarma ağı, bir giriş görüntüsünden özellikleri çıkaran geleneksel bir evrişimli sinir ağıdır (CNN). Dikkat ağı daha sonra farklı duyguları tanımak için yüzün en alakalı kısımlarını vurgulayan bir dikkat haritası oluşturmak için bu özellikleri kullanır. Yaklaşımları, yaygın olarak kullanılan dört yüz ifadesi tanıma veri seti kullanılarak test edilmiştir: FER2013, CK+, JAFFE ve FERG (Minaee ve ark., 2021).

Hardik Uppal ve arkadaşları, RGB-D (kırmızı, yeşil, mavi ve derinlik) sensörlerini kullanarak yüz temsili öğrenimi için bir dikkat mekanizması önermektedir. Önerilen derinlik yönlendirmeli dikkat yöntemi, bir CNN tarafından çıkarılan derinlik özellikleri kullanarak, derin ağın RGB görüntüsündeki yüzün kişiye özel daha belirgin bilgiler içeren bölümlerine odaklanmasına rehberlik etmektedir. RGB-D tabanlı yüz tanıma teknikleri, RGB görüntüleri için bir dikkat haritası oluşturmak üzere hem derinlikten hem de RGB modalitelerinden görsel özellik yerleştirmeleri kullanır. Yaklaşımları dört genel veri kümesi üzerinde test edilmiştir: Lock3DFace, Curtin Faces, IIIT-D RGB-D ve KaspAROV (Uppal ve ark., 2021).

Torki Altameem ve diğerleri, sağlık hizmetlerinde, yani ruh sağlığı ve ağrı yönetimi alanlarında hasta bakımını ve sonuçlarını iyileştirmek için yüz ifadesi tanıma teknolojisini kullanan bir sistem sunmuştur. Yazarlara göre, yüz ifadesi tanıma, bir hastanın duygusal durumu ve ağrı miktarı hakkında hayati bilgiler sağlayarak sağlık hizmeti sağlayıcılarının daha iyi tedavi kararları vermesine olanak sağlayabilir. Hastalardaki yüz ifadelerini güvenilir bir şekilde algılamak ve yorumlamak için önerilen sistem, insan-makine etkileşimi ve çok modlu görselleştirme analizinin bir kombinasyonunu kullanır. Önerilen yöntem, girdi temsiliinde yüz duygularından ve dokularından yararlanır. Ayrıca, önerilen yöntem, doku sınıflandırması, korelasyon ve depolanan bilgileri kullanarak yüz görselleştirmenin tespiti için üç katmanlı bir evrişimli sinir ağı (CNN) kullanır. Üç kademeli operasyonlar, analitik karmaşıklık ve yanlış algılamayı azaltmakla bağlantılıdır (Altameem ve Altameem, 2020).

Yazarlar (Xiao Sun ve diğerleri), etkili bir yüze özgü evrişimli sinir ağını (CNN) eğitmek için ilgi bölgesini (ROI) belirlemeye odaklanan yüzdeki duygu tespiti amacıyla derin bir öğrenme çerçevesi sunmaktadır. ROI bölgeleri arasındaki bağlantılardan yararlanan derin

mimari, tahmin edilen hedeflerin öngörülebilirliğini artırır. Yazarlar ayrıca yüz ifadesi ile ilgili bölgeleri kullanan CNN'yi eğitmek için yüz görüntüsü, ön işleme ve ROI ayırmayı tartışmaktadır. Yapay yüz veri artırma yaklaşımı, derin CNN eğitiminin sağlamlığını artırmayı amaçlamaktadır. Önerilen yöntem, herkese açık veri kümeleri CK+, JAFFE ve FER-2013'ün yanı sıra kendi derledikleri bir Wild veritabanı (Sun ve ark., 2020) kullanılarak test edilmiştir.

Yazarlar, DSAE (Deep Sparse Autoencoders, DSAE) kullanan, yüz ifadelerini tanımak için yeni bir yöntem olan ve hem yüz geometrisi hem de görünüm bilgisini içeren yüksek boyutlu bir özelliği çalışmalarında sunmaktadırlar. Burada, DSAE tabanlı teknik, CK+ veri tabanındaki çeşitli yüz ifadelerini ayırt etmek için kullanılmaktadır. Yazarlara göre, önerdikleri teknik, temel bileşen analizi (PCA) kullanarak yüksek boyutlu yüz işareti verilerini sıkıştırmaktadır. Yazarlar ayrıca üst üste binmiş gürültü giderici otokodlayıcılar temel alınarak yeni özellik çıkarma ve sınıflandırma tekniklerini de sunmaktadırlar (Zeng ve ark., 2018).

Yazarlar, evrişimli bir sinir ağını (CNN) görsel bir dikkat mekanizmasıyla birleştirerek yüz ifadelerini tanımak için bir yöntem önermektedirler. Ağ, yüzlerden yerel evrişimsel özellikleri ayıklamakla birlikte, ilgilenilen bölgeleri otomatik olarak saptamak için yerleşik bir dikkat modeli kullanmaktadır. Sonrasında, duygusal etiketi anlamak için bu bölgelerdeki yerel özellikler toplanmaktadır. Önerilen yöntem üç veri kümesinde test edilmiştir: RaFD-FRONT, RaFD-POSE ve SFEW (Sun ve ark., 2018).

Araştırmacılar, derin öğrenme ve kompakt çerçeve tabanlı bir yüz ifadesi tanımlama sistemi önermektedir. Önerilen sistem, yüz fotoğraflarından ayırt edici özellikleri öğrenmek için bir evrişimli sinir ağı (CNN) mimarisine dayanmaktadır. Zamansal bilgiden yararlanmak için mimari, GRU (Gated Recurrent Units, GRU) kullanmaktadır. GRU'lar, bir video klipteki ardışık kareler arasındaki zamansal ilişkileri simüle etmek için kullanılır. Buna ek olarak, araştırma, hibrit veri kaynaklarıyla derin sinir ağlarını eğitirken aşırı öğrenmeyi azaltmak için bir artırma stratejisi sunmaktadır. Önerilen teknikler, Extended Cohn-Kanade ve Oulu-CASIA veritabanları gibi halka açık veri kümelerindeki performansı arttırmaktadır (Kuo ve ark., 2018).

Yazarlar, derin sinir ağı (DNN) ve koşullu rastgele alan (CRF) modülünden oluşan iki parçalı bir ağ kullanan ve filmlerdeki yüz ifadelerini tanıyabilen bir yöntem sunmaktadır. DNN, yüz görüntüleri içindeki uzamsal ilişkiyi yakalamak amacıyla evrişimli katmanları ve

Inception-ResNet modüllerini kullanırken, CRF modülü görüntü çerçeveleri arasındaki zamansal ilişkiyi yakalamaktadır. Önerilen ağları analiz etmek için halka açık üç veri tabanı (CK+, MMI ve FERA) kullanılmıştır (Hasani ve Mahoor, 2017).

Yazarlar, standart el yapımı özellikleri (SIFT) ve evrişimli sinir ağlarından çıkarılan derin öğrenme özellikleriyle harmanlayan bir yüz ifadesi tanıma (FER) yaklaşımı sunmaktadır. Birleştirilen özellikler daha sonra destek vektör makineleri (SVM) kullanılarak sınıflandırılır. Yazarlar, halka açık CK+ veri tabanlarını kullanarak önerilen stratejilerini doğrulamışlardır. Ayrıca, önerilen FER tekniğine ek olarak, insanın görsel algısını taklit eden bir multimedya kalite modeli de sunmuşlardır. Temsili bölgeleri seçmek üzere çoklu görüntüleme aktif öğrenme yöntemi kullanılmaktadır. Aynı zamanda, yüksek kaliteli eğitim fotoğraflarının insan algısına dayalı olarak görüntü kalitesini tahmin etmek için bir olasılıksal model kullanılır. Yazarlar, modellerin görüntü kalitesinde sığ olanlardan daha iyi performans gösterdiğini de iddia etmektedirler (Wang ve ark., 2019).

Yazarlar (Chen ve diğerleri), otomatik yüz ifadesi tanıma (FER) için iki aşamalı bir çerçeve önerdiler. Çalışmanın ilk aşaması, nötr ve tamamen anlamlı çerçeveleri yüz ifadesi dizilerinden ayırmak için ikili bir CNN kullanılmasıdır. Nötr ve tamamen anlamlı çerçeveler arasındaki fark bilgisi, uçtan uca bir DCNN (Difference Convolution Neural Network, DCNN) kullanarak altı temel yüz ifadesini sınıflandırmak için ikinci adımda kullanılmaktadır. Önerilen teknik, CK+ ve BU-4DFE olmak üzere iki kıyaslama veri seti üzerinde test edilmiştir. Önerilen yöntem, öğrencilerin bir E-öğrenme ortamındaki duygusal durumlarını incelemek için de kullanılmaktadır (Chen ve ark., 2019).

Yazarlar (Rajendra ve diğerleri), yüz ifadelerini tanımak için denetimsiz ve denetimli öğrenmeyi birleştirmekle birlikte, HOG ve 2D-DWT özelliklerinden bu çalışma kapsamında yararlanmaktadır. Sistemin doğruluğunu ve hesaplama verimliliğini artırmak için önerilen yaklaşım, DBN (Deep Belief Network, DBN) olarak adlandırılan dört katmanlı bir sinir ağı ve azaltılmış yüz özelliklerini kullanır. Bu teknik üç genel veri kümesi için kullanılmıştır: CK+, RaFD ve MMI Veri tabanı (Rajendra ve ark., 2019).

Yazarlar (Farkhod Makhmudkhujaev ve diğerleri), yüz ifadesi tanımlaması için LPDP (Local Prominent Directional Pattern, LPDP) olarak adlandırılan kenar tabanlı bir tanımlayıcı önermektedirler. Bu bağlamda, yerel şekle göre önemli kenarlar elde etmek ve kenar kodlamasını sağlamak için bir pikselin yerel komşuluğunu analiz etmektedirler. Yöntemleri, özelliksiz bölgeleri kodlama sürecinden çıkarmak amacıyla uyarlamalı bir

eşikleme mekanizması ve gelişmiş kategorizasyonu sağlayan ilgili özellikleri seçmek amacıyla da bir özellik seçme mekanizmasını kullanmaktadır. Önerilen LPDP, CK+, MMI, BU-3DFE ve ISED veri setlerine uygulanmıştır. Ayrıca bu çalışmada, önerilen yöntem için gürültü ve konumsal dalgalanmaların varlığındaki kararlılığı da incelenmektedir (Makhmudkhujaev ve ark.,2019).

Yazarlar (Renda ve diğerleri), evrişimli sinir ağlarını (CNN) kullanarak, yüz ifadesi tanımlama için toplu öğrenme uygulamasını araştırmaktadırlar. Eğitim verilerini değiştirme ve ağırlık başlatma gibi temel ağ toplulukları içinde çeşitlilik üretmenin farklı yolları burada test edilmiştir. Ayrıca, temel sınıflandırıcı çıktılarını birleştirmek için çeşitli füzyon stratejilerini incelemekle birlikte çeşitli OWA (Ordered Weighted Averaging, OWA) operatörleri ile de deneyler yapmışlardır. Amaç, FER durumlarında daha verimli topluluk oluşturma süreçlerinin sağlanmasıdır. Çalışmada, el yapımı bilgiler kullanılmadan duygu tanıma için evrişimli sinir ağları (CNN), topluluklarının performansını incelemektedir. Dört farklı topluluk tasarım metodolojisi çalışma kapsamında araştırılmıştır. FER2013 veri seti, yöntemlerinde eğitim ve test için kullanılırken, CK+ ve SFEW veri setleri ön eğitim için kullanılmaktadır (Renda ve ark., 2019).

Yazarlar (Yingsheng Ye ve diğerleri), yüz ifadelerini tanımak için bölge tabanlı Evrişimli Füzyon Ağı (RCFN) adı verilen bir yöntem önermektedirler. Tüm yüz yerine, yaklaşımları yüzün üç kritik bölgesine (alın, göz ve ağız) odaklanır. Yazarlar, bu kritik bölgeleri bir kas hareketi modeli ve yüz işaretlerini kullanarak tanımlar. Daha sonra farklı boyutlar ve yerleşimlerin neden olduğu girişimi azaltmak için normalleştirilir. RCFN, her bir kritik bölge içindeki karşılık gelen kalıpları öğrenmek ve en düşükten en yüksek soyutlama derecesine kadar üç seviyeli özellikleri çıkarmak için üç ayrı alt ağ kullanır. Ağ optimizasyonu ve daha etkili bir kayıp fonksiyonu oluşturmak için, popüler SCE'ye (Softmax Cross Entropy, SCE) CPL (Constrained Punitive Loss, CPL) eklenir. Yöntemleri üç genel veri kümesinde test edilmiştir: KDEF, CK+ ve Oulu-CASIA (Ye ve ark., 2019).

Yazarlar (Hamid Sadeghi ve diğerleri) histogram verilerini kullanarak yüzdeki duygu tanıma için bir öğrenme yöntemi önermektedir. Bu yöntemde ki-kare mesafesi ayarlanır ve yüz alt bölgelerinde yerel olarak yapılan metrik öğrenme ile konveks maliyet fonksiyonu kullanılır. Histogram veri sınıflandırması, özellikle yüz ifadesi tanımlaması için, bu metrik öğrenme yaklaşımı v2 mesafesini kullanır. Aşırı öğrenmeyi en aza indirmek için bu çalışma bir düzenleyici kullanır ve doğruluğu artırmak için yüzün alt bölgelerine yerel olarak uygulanır.

Yöntemleri, CK+, SFEW ve RAF-DB dâhil olmak üzere çeşitli veri kümeleri kullanılarak test edilmiştir (Sadeghi ve Raie, 2019).

Yazarlar (Weicheng Xie ve diğerleri), derin öğrenme tekniklerini kullanarak yüz ifadelerini tanımak için bir yöntem önermektedirler. Öğrenilen özelliklerin genelleme kapasitesini geliştirmek amacıyla bu yöntem, derin metrik öğrenmeye özellik seyrekliği ekler. Önerilen öznitelik seyrekliği, ağıın son iki tam bağlantılı (FC) katmanına dâhil edilir ve FC özniteliklerinin seyreklik kayıpları formüle edilir. Kayıp fonksiyonu, softmax ve özellik seyrekliği kayıplarının bir kombinasyonudur. Eğitim ve değerlendirme için VGG ve ResNet ağlarından yararlanılmıştır. Yöntemleri, halka açık dört veri seti kullanılarak test edilmiştir: FER2013, CK+, Oulu-CASIA ve MMI (Xie ve ark., 2019).

Radzi ve diğerleri, kullanıcı girişi olarak Grafik Kullanıcı Arayüzü (GUI) ile MATLAB tabanlı bir Evrişimli Sinir Ağını (CNN) kullanarak bir yüz tanıma sistemi geliştirdiler. Önerilen CNN, sinir ağı eğitim süresini azaltma yeteneğine sahiptir. AT & T veritabanından 40 ve JAFFE veritabanından 10 denekten oluşan görüntüleri kullandılar (Radzi ve ark.,2016).

Mayya ve arkadaşları, tek bir görüntüden bir bireyin yüz ifadelerini tanımaya odaklanan bir Derin Evrişimli Sinir Ağını (DCNN) önerdiler. Genel amaçlı grafik işleme birimi (GPGPU) kullanılarak özellik çıkarma süresi önemli ölçüde azaltılmıştır. CK+ Veri Kümesi (Japon Kadın Yüz İfadesi) ve JAFFE Veri Kümesi'nden; öfke, iğrenme, korku, mutluluk, nötr, üzüntü ve sürpriz gibi birçok yüz ifadesi için görüntüler kullandılar (Mayya ve ark., 2016).

Abir Fathallah ve arkadaşları, otomatik yüz ifadesi tanıma için evrişimli sinir ağlarına (ESA) dayalı Derin Sinir Ağlarını (DNN) kullandılar. Sonuçları iyileştirmek amacıyla, mimarilerinde Görsel Geometri Grubu modelini (VGG) kullanarak ince ayar yaptılar. Sonrasında, mimarilerini değerlendirmek için, birçok büyük halka açık veritabanını (CK+, MUG ve RAFD) kullanarak modelleri test ettiler. Elde edilen sonuçlar, CNN yaklaşımının, yüz ifadesi analizinde ve birçok genel veritabanında yüz ifadesi tanıma çok etkili olduğunu göstermiştir (Fathallah ve ark.,2017).

Sang ve arkadaşları, yüzlerde bulunan anlamsal bilgileri, özellik tanımlayıcılarını elle tasarlamadan otomatik ve doğru bir şekilde yorumlayabilen etkili derin Evrişimsel Sinir Ağlarını (CNN'ler) önermek için, derin öğrenmedeki son gelişmelerden faydalandılar. Kaggle yüz ifadesi tanıma yarışmasından sağlanan FER2013 veri kümesini kullandılar. Veri kümesi 48x48 boyutunda 35.887 gri görüntüden oluşmaktadır. Kaggle bu görüntüleri

28.709 eğitim görüntüsü, 3589 genel test görüntüsü ve 3589 özel test görüntüsü olarak ayırmıştır (Sang ve ark.,2017).

Arya ve arkadaşları, yüz tanıma sisteminin tipik yapılarının elde edilmesi, yüz özelliklerinin çıkarılması ve yüzlerin tanınması olmak üzere üç ana adımdan oluşan derin öğrenme ve Doğrusal Ayrıştırma Analizine (LDA) dayalı bir yüz tanıma yöntemi önermişlerdir. İlk aşamada, yüz görüntüleri birden fazla yüz parçasına bölünür. Bu adım görüntülerin ön işlenmesi olarak adlandırılır. İkinci olarak, görüntüler LDA algoritmasını uygulamak üzere özellik çıkarımı için kullanılır. Son olarak, sinir ağından çıkarılan yüz özelliklerini ve sınıflarını kullanarak modelin eğitilmesi ve sonrasında eğitilen modelin yüzleri tanınması için kullanılması amaçlanmıştır (Arya ve ark., 2018).

Sharma ve arkadaşlarının önerdiği sistemde yüzün ve gözlerin boyutuna dayalı yüz algılama için derin öğrenme kullanıldı. Çalışmada yüzün tanımlanması için open-cv kullandı. Çünkü open-cv görüntü işlemede yardımcı olan bir sınıflandırıcı olarak kabul edilmektedir. Bu işlemlerinin sonraki aşamalarında, Python programından faydalanılarak önceki tüm işlemler için derin bir mimari kullanılmıştır (Sharma ve ark.,2017).

Li ve Lam, yüz ifadelerini tanımak için derin sinir ağı kullanan bir teknik geliştirmişlerdir. Yüz ifadelerini çıkarmak için Gabor filtresi kullanılmış ve daha sonra PCA uygulanmıştır. Ardından önceki aşamadan çıkarılan özellikler derin bir sinir ağında işlenmiştir (Li ve Lam,2015).

Hussein ve arkadaşları, araştırmalarında nispeten düşük karmaşıklığa sahip yüz ifadelerinin geometrik özelliklerini kullanmıştır. VDMFP, BINED ve MMI olmak üzere farklı yüz görüntülerinden oluşan üç veri tabanı çalışma kapsamında kullanılmıştır. Bu veri tabanları, farklı görüntüleme koşullarında farklı yüz ifadelerinin görüntülerini içermektedir. Bu görüntüler merkezi bir nokta etrafında döndürülmüş ve sonrasında bu görüntüler işlenmiştir. Bu noktada, derin öğrenme kullanılarak özellikler çıkarılmıştır (Husseinve ark.,2018).

Prasada ve diğerleri, yüz ifadeleri ve konuşma özelliklerini birleştirmek için bir yöntem önermişlerdir. Burada konuşmayı tanımak için MFCC (Mel Frequency Cepstral Coefficients) ve yüz duygularını tanımak için MSER (Maximally Stable Extremal Regions) kullanılmıştır. Duyguları sınıflandırmak için ise SVM sınıflandırıcı ve cinsiyet sınıflandırması için AdaBoost kullandılar. Bu çalışmada Hint Yüz Veritabanı ve Berlin Konuşma Veritabanından faydalanılmıştır (Prasada ve ark., 2019).

Tarnowski ve diğerkleri, yüz ifadelerini kullanarak duygu tanıma için KDEF veri setini kullandılar ve burada yüz ifadelerine dayalı yedi duygusal durum (nötr, neşe, üzüntü, sürpriz, öfke, korku, iğrenme) bulunmaktadır. Ayrıca, K-NN ve MLP sinir ağı sınıflandırıcıları kullanılarak bu veri seti sınıflandırıldı (Tarnowski ve ark., 2017).

Qayyum ve diğerkleri, çalışmalarında JAFFE, CK+ ve MS-Kinect veri setlerini kullanmıştır. Yüz ifadesi tanıma amacıyla özellikleri çıkarmak için sabit dalgacık dönüşümü (SWT) kullanılmıştır. Sabit dalgacık dönüşümünü düzgün bir şekilde uygulayabilmek için yatay ve dikey alt bantların bir kombinasyonu oluşturulmuştur. Burada alt bantlar, yüz ifadelerinin çoğu için kas hareketi bilgilerini içermektedir (Qayyum ve ark., 2017).

CK+ veri kümesinin kullanıldığı bu çalışmada, yüz ifadesi tanıma için yönelimli gradyanların histogramlarına (HOG) dayalı bir yöntem önerilmiştir. Burada HOG bir tanımlayıcı olarak ele alınmaktadır ve HOG özellikleri elde edilmektedir. Sonrasında bu özellikler bir grup Destek Vektör Makineleri (SVM'ler) ile sınıflandırılmıştır. Burada SVM ayırt edici bir sınıflandırıcı olarak nitelendirilebilir (Carcagni ve ark., 2015).

Waqas ve Choe, birçok yüz görüntüsünden insan yüzü görüntülerini seçmek için CNN kullandılar. Bu çalışma için on duygudan (gülümseme, dil, kızgın, ağlama, öpücük, sırtıma, iğrenme, şaşırma, göz kırpması, üzgün) oluşan iyi görüntüler seçilmiş ve yüz ifadesi tanımlanmıştır. Ayrıca, yüzlerin tespit edilmesi için open-cv kullanılmıştır (Waqas ve Choe, 2018).

Leo vd., robot ile otizm spektrum bozukluğu (OSB) olan çocuklar arasındaki etkileşimi hesaplamak için bir robot kullanan ve robota taklit için yüz ifadesinin tanınmasını sağlayan bir sistem önermişlerdir. Daha sonra robotun, duygu taklit yeteneğini değerlendirmek ve robotun isteği ile çocuğun eylemi arasında geçen süreyi ölçmek için ifadeyi taklit etmesi istenmiştir. Bu noktada, yüzler kırılıp kaydedilmiştir. Kullanılan SVM'e girdi olarak sağlanan bir veri vektörü oluşturmak için HOG tanımlayıcısı kullanılmıştır (Leo ve ark., 2015).

Lopes ve ark., az sayıda veri kullanarak yüz ifadesi tanıma için basit bir çözüm önerdiler. Bu yöntemde, konvolüsyonel sinir ağı (CNN) ve belirli görüntü ön işleme adımlarının (rotasyon düzeltme, kırpması, aşağı örnekleme ve yoğunluk normalizasyonu) kombinasyonu kullanıldı. Ön işleme adımları open-cv ve CNN Kütüphanesi (CAFÉ) kullanılarak yapıldı. Ayrıca, eğitim ve test aşamaları için halka açık veritabanları (CK +, JAFFE ve BU-3DFE) kullanıldı (Lopes ve ark., 2017).

Sajjad vd., yüz ifadesi ve şüpheli aktivite tanıma için bir yöntem önerdiler. Bu amaçla, Raspberry Pi kullanılmış olup, video Raspberry Pi kamera tarafından elde edilmiştir. Yüzler Viola Jones algoritması kullanılarak algılanırken, yüz bölgesinde ön işleme için Gabor filtresi kullanılmıştır. Özelliklerin çıkarılması için Yönlendirilmiş Hızlı ve Döndürülmüş Brief (ORB) algoritması kullanıldı. SVM'in yüz ifadelerini (öfke, iğrenme, korku, mutlu, nötr, üzgün ve sürpriz) sınıflandırdığı sahnenin arkasındaki duygunun tahmin edildiği görülmüştür. Bu tahminlerle belirli bir durumun şüpheli olup olmadığı ve dolayısıyla ortaya çıkmadan önce önlenebileceği bu çalışma kapsamında araştırılmıştır. Test ve eğitim aşamaları için halka açık veri tabanı (CK +, JAFFE, MMI) kullanılmıştır (Sajjad ve ark., 2019).

Chen ve ark., yüz ifadelerini tanımak, öğrenme verimliliği ve hesaplama karmaşıklığı sorunlarını ele almak için, softmax regresyon otoenkoder ağını (SRDSAN) önerdiler. Burada, derin yedek otoenkoder ağı (DSAN), üst düzey özellikleri çıkarmak ve yüz duygu özelliklerini öğrenmek için kullanıldı. Ayrıca yüz duyguları, softmax regresyon (SR) kullanılarak sınıflandırıldı ve veritabanı olarak (CK +, JAFFE) kullanıldı. Önerilen yöntem insan duygularını analiz etmek ve anlamak amacıyla robotlara entegre edildi (Chen ve ark., 2018).

Surace vd., vahşi doğada duygu tanıma için derin CNN'ler ve Bayes sınıflandırıcıları arasındaki kombinasyona dayalı bir yöntem önerdiler. Önerilen algoritma Python'da çalıştırılmış olup CNN'lerin eğitimi için TensorFlow Kütüphanesi kullanılmıştır. Ayrıca görüntüler üzerinde ön işleme adımları için open-cv kullanılmıştır. Burada veritabanı olarak Emotion Recognition in the Wild 2017 (EmotiW17) kullanılmıştır (Surace ve ark., 2017).

Wu ve ark., ağırlık uyarlamalı konvolüsyon sinir ağına (WACNN) dayalı bir yüz ifadesi tanıma yöntemi önerdiler. Burada yüz ifadesi görüntüleri için ön işleme, geometri normalizasyonu ve gri normalizasyonu ile yapılmıştır. Bu adımdan sonra PCA kullanılarak düşük seviyeli ifade özellikleri çıkarılmıştır. Bu adımdan sonra ağırlık uyarlamalı konvolüsyon sinir ağı (WACNN) özellikleri öğrenmek ve bunları tanımak için kullanılmıştır. Böylece yüz ifadesi bilgisi elde edilebilmiştir. Çalışmada veritabanları olarak CK +, JAFFE ve SFEW2.0 kullanılmıştır (Wu ve ark., 2019).

Sadeghi ve arkadaşları Ki-kare mesafeli derin histogram metrik öğrenmeye dayalı Evrişimsel Sinir Ağında (CNN) yüz ifadesi tanıma için bir sistem önerdiler. Çalışmada, CK+, MMI, SFEW ve RAF-DB genel veri kümeleri kullanılarak yedi yüz ifadesi

sınıflandırıldı. Sonuçlardan, önerilen yöntemin doğruluğun %98,47'e eşit olduğu görüldü (Sadeghi ve ark., 2022).

Zhang ve arkadaşları yüz ifadelerini tanımak için üç alt modüle dayalı bir yöntem geliştirdiler. İlk olarak, İfade Özellik Çıkarıcı modülü ifade özellikleri çıkarılmış ve kabaca ifadeyle ilgili bölgeleri gösteren bir dikkat maskesi oluşturulmuştur. Sonrasında, İfade Maskesi Arıtıcı (EMR) modülü, ifadeyle ilgili bölgeler arasındaki ilişkiyi modelleyerek ve maskelenmiş özellikleri üreterek maskeyi daha hassas hale getirmiştir. Böylece, İfade Örüntü Haritası Oluşturucu (EPMG) modülü maskelenmiş özellikleri üretilmiştir. Önerilen yöntemde, dört genel veri kümesi CK +, Oulu-CASIA, RAF-DB ve AffectNet üzerinde 6 yüz ifadesi kullanılmıştır (Zhang ve ark., 2022).

Ruan ve diğerleri, yüz ifadesi tanıma için dört grup halka açık veri kümesini (Fer2013, JAFFE, KDEF ve Raf-DB) tek bir büyük veri kümesi elde etmek için birleştirdiler. Veri kümelerine maske takmayı da içeren üç senaryo uygulanmıştır (burun ve ağız gizlenmiştir). Güneş gözlüğü takmak gözleri gizlemek için kullanılırken, şapka takmakta gözleri ve burnu gizlemek için kullanılmaktadır. Gizleme işleminden sonra veri kümeleri ResNet50 ağında eğitilmiş ve test edilmiştir. Bu çalışmada yüz ifadeleri olarak öfke, iğrenme, korku, mutluluk, normal, üzüntü ve şaşkınlık sınıfları kullanılmıştır (Ruan ve ark.,2022).

Mahmoudi ve arkadaşları, CNN tabanlı yüz ifadesi tanıma için bir yöntem önerdiler. Çalışmada klasik evrişimsel katman, klasik doğrusal evrişim çekirdeğinden daha karmaşık kalıpları öğrenebilen ve böylece daha ayırt edici olabilen Taylor serisi çekirdekli evrişim katmanı ile değiştirilmiştir. Bu çalışmada, üç genel veri kümesi RAF-DB, ExpW ve FER2013 kullanılmıştır ve önerilen yöntem yedi yüz ifadesini tanıyabilmektedir (Mahmoudi ve ark., 2022).

Ge ve diğerleri, yüz ifadesi tanıma için derin öğrenmeye dayalı bir yöntem önerdiler ve AlexNet-Emotion ağını kullandılar. Tam bağlı ikinci katmanda değişiklik yaptılar ve model tarafından öğrenilen ifade özelliklerini daha ayırt edici ve kompakt hale getirebildiler. Burada, karar katmanında hesaplanan softmax kaybı, CNN parametrelerinin ince ayarından sorumludur (Ge ve ark.,2022).

Maithri ve diğerleri çok sayıda makaleyi değerlendirmiş ve insan destekli ER için önerilen en son modelleri incelemişlerdir. Otomatikleştirilmiş ER için EEG, yüz, konuşma ve çok modlu sinyaller, çeşitli makine ve derin öğrenme teknikleri kullanılmıştır. Kontrollü bir

ortamda yüksek performanslı ölçümlere izin veren derin öğrenme teknikleri kullanan otomatikleştirilmiş ER'de önemli bir artış kaydedilmiştir (Maithri ve ark.,2022).

Xin Ma ve diğerleri, yüz ifadelerini tanımak için RCANet'i önerdiler. RCANet ağ mimarisi, sınıflar arası benzerlik ve sınıf içi varyasyon problemlerini çözmek için kullanılmıştır. Özellik temsil yeteneğini geliştirmek için uzamsal ve kanal boyutları arasındaki ilişkiler araştırılmış ve toplanmıştır. Alıcı alanları artırmak ve çok ölçekli bağlamı yakalamak için basamaklı bir bağlam çıkarma yöntemi tanıtılmıştır. Bağlam modülü, kademeli bir şekilde birbirine bağlanan birkaç genişlemiş birimden oluşur. Yerel bağlam öğrenimi için, bağlam modüllerinde genişletilmiş evrişim kullanılmıştır. Hem RAF-DB hem de AffectNet genel veri kümeleri kullanılmıştır. Veri kümelerinde sırasıyla %90 ve %65,65 doğruluk elde edilmiştir (Ma ve ark.,2022).

Albraikan ve diğerleri, yüz ifadesi tanıma için ResNet50 tabanlı yüz algılamalı Mask RCNN, Adam optimize edici tabanlı hiperparametre optimizasyonu, DenseNet121 tabanlı özellik çıkarımı ve FER2013 genel veri kümesini kullanan ağırlıklı kernel aşırı öğrenme makinesi (WKELM) tabanlı sınıflandırmayı içeren üç aşamalı bir süreçten oluşan IFER DTFM modelini önerdiler (Albraikan ve ark.,2022).

Guo ve diğerleri, yüz ifadesi tanıma amacıyla genel (genel yüz öğretileri) ve yerel yüz ayrıntılarını (ağız, burun ve gözler) harmanlayan yüz ifadesi tanıma için bir sistem olan Çok Bölge Dikkat Dönüşüm Çerçevesini önermiştir. Sistemleri CNN ağlarına dayanmakta olup RAF-DB, FERPlus ve CAER-S gibi genel veri kümelerini kullanmışlardır (Guo ve ark.,2022).

Allaert ve diğerleri, yüz ifadesi tanıma bağlamında yüz hareketini kodlamak için optik akışı ve özellik çıkarımı için CNN ağlarını ve bir sınıflandırıcı olarak SVM'yi birleştiren ve yüz ifadelerini tanımak için bir sistem önermiştir. Veri kümeleri SNaP-2DFE şeklindedir (Allaert ve ark.,2022).

Yahui Nan ve diğerleri, mobileNet'e dayalı yüz ifadesi tanıma için bir model önermiştir. Temel MobileNet'i değiştirmiş olup modelin yüz ifadelerinin ince taneli özelliklerini çıkarma yeteneğini geliştirmek için her evrişim katmanından sonra iki dikkat katmanı eklenmiştir. Ayrıca aşırı öğrenmeyi önlemek için dropout teknolojisi eklenmiştir. Son katmanda, model parametrelerini optimize etmek için merkez kaybı ve softmax kaybı birleştirilmiştir. Bu sistemler FERPlus ve RAF genel veri kümelerini kullanır (Nan ve ark.,2022).

Suchitra Saxena ve diğerkleri, yüz ifadelerini mutlu, öfkeli, tiksinti, sürpriz ve tarafsız bir durum gibi farklı ifadeler olarak tanımak için CNN ağına dayalı bir yöntem önerdiler. Bu modeller, olumlu ve olumsuz duygular gibi duygu yoğunluğu düzeylerini CMU Multi-PIE veri setini kullanarak tanıyabilmektedir (Saxena ve ark., 2022).

Silvio Barra ve arkadaşları, hastaların yüz ifadelerini uzaktan izlemek için, giriş çerçevesinden görüntü ön işleme sırasında tanımlanan yüz işaret noktalarını LBP ve HoG gibi birkaç farklı doku tanımlayıcısı kullanılarak yüz bölgesinin algılandığı bir IoT sistemini önermiştir. Yüz görüntülerinden çıkarılan özellik vektörleri, çok sınıflı bir doğrusal SVM sınıflandırıcısına tabi tutulmuş ve genel veri kümeleri olarak KDEF ve GENKI-4K kullanılmıştır (Barra ve ark., 2022).

Zhu ve arkadaşları, suçlara karışan çocukların psikolojik özelliklerini araştırmak için bir yüz ifadesi tanıma modelini geliştirdiler. Bu modeller birkaç sinir ağını (VGGNet, AlexNet ve LeNet-5) ve optimize edilmiş bir Merkezi Yerel İkili Model (CLBP) algoritmasını birleştirmektedir. Burada, JAFFE ve CK+ gibi genel veri kümeleri kullanılmıştır (Zhu ve ark.,2022) .

3. MATERYAL VE METOT

3.1. Veri Setinin Toplanması

Bu çalışmada kullanılan yüz görüntüleri, üniversitenin etik kurulundan gerekli izinler alındıktan sonra Kahramanmaraş Sütçü İmam Üniversitesi Mühendislik Mimarlık Fakültesi öğrencilerinden temin edilmiştir. 70'i erkek ve 13'ü kadın olmak üzere 83 öğrencinin yüz resimleri kullanılarak bir veri tabanı oluşturulmuştur. Erkek/kız oranı kabaca fakültenin genel öğrenci grubunu temsil etmektedir.

Her öğrenciden 24 fotoğraf alınmak suretiyle 83 öğrencinin görüntüleri çalışmamız için yeterli veri setini oluşturmaya yetmiştir. Aynı zamanda bu görseller artırılarak veri setindeki yüz görüntülerinin sayısı 4 katına çıkarılmıştır. Bu noktada, Türkiye ve Orta Doğu benzer yüz ifadelerine sahip olabilir. Sonuç olarak, bu çalışmanın bulguları Orta Doğu'daki geniş bir nüfusa uygulanabilir.

Bu çalışmada öğrencilerin sekiz yüz ifadesi (öfke, tiksinti, korku, mutluluk, tarafsızlık, üzüntü, dikkat ve sürpriz) kullanılmıştır. Veritabanımız şu anda yukarıda gösterilen sekiz kategoriye eşit olarak dağılmış 7968 fotoğraf içermektedir ve 996 fotoğraf her bir sınıfta bulunmaktadır.

Şekil 3.1 ve 3.2 veri tabanımızdan örnek görüntüleri göstermektedir. Öğrencilerin yüzlerinin görüntüleri, belirli alanlarda aydınlatma eksikliği veya diğer alanlarda aydınlatma varlığı gibi gerçek dünyadaki aydınlatma koşullarını simüle eden çeşitli ortamlarda alınmıştır.



Şekil 3.1 Aynı öğrenci için veri setinden sekiz yüz ifadesi örneği



Şekil 3.2 Çeşitli öğrenciler için veri setinden sekiz yüz ifadesi örneği

3.1.1. Veri kümesinin önemi

Veri setimizin aşağıdaki nedenlerden dolayı avantajlıdır:

1. Yüz ifadesi tanıma: Aydınlatma ortamlarındaki farklılıklar nedeniyle, yüz ifadesi tanımda gerçek aydınlatma dikkate alınmıştır.
2. Veri kümemiz genel yüz bilgilerini işler.
3. Veri setimiz sekiz yüz ifadesi içermektedir. Bu çalışma üniversite öğrencileri üzerinde yapılması nedeniyle, öğretmene öğrencilerin derse katılım ve anlayış düzeyi hakkında bir ilk izlenim sağlaması açısından çok önemlidir. Araştırma üniversite öğrencileri üzerinde yapıldığı için en önemli yüz ifadesinin dikkat olduğu düşünülebilir.

3.1.2. Veri setinin elde edilmesi

COVID-19 koşulları nedeniyle ve herkesin sağlığını korumak için yüz yüze veri toplama yöntemini uzaktan olarak değiştirdik. Bu nedenle cep telefonu veya bilgisayar aracılığıyla veri toplamak için bir araç tasarlanmıştır. Bu araç Kobo araç kutusuna bağlıdır. Ayrıca, üç dilde (Türkçe, Arapça ve İngilizce) desteklenen bir anket tasarlandı. Bir diğer taraftan öğrenciler için örnekler koyulmuştur. Her örnek, ekteki örneğe benzer bir fotoğraf çekmek için gereken yüz ifadesi durumunu ifade etmektedir. Her yüz ifadesi için üç resim gerekmektedir. Öğrencinin her örneğin yanındaki kutucuğa tıkladığında iki seçeneği vardır; öğrenci resmi doğrudan çekebilir veya mevcut bir dosyadan kendi resmini yükleyebilir.

Bu şekilde, 169 kız ve erkek öğrenciye ait fotoğraf toplayabildik. Bu görüntülerin toplama işlemini tamamladıktan sonra, çekilen görüntülerin ekteki fotoğraflara ne ölçüde uyduğunu (belirlediğimiz kurallara göre) tespit edebilmek için bir değerlendirme süreci yürüttük. Buna göre elde edilen veriler ve görüntüler için bir temizleme işlemi gerçekleştirdik. Bu sayede veri tabanımızın açıklamasında da gösterildiği gibi 83 erkek ve kız öğrencilerin görüntülerinden yararlanmanın mümkün olduğu sonucuna vardık.

Aşağıdaki bağlantı, yukarıda belirtilen üç dilde desteklenen öğrencilere dağıtılan anketi göstermektedir.

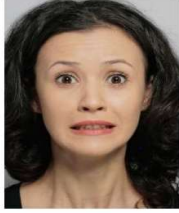
<https://ee.humanitarianresponse.info/x/G7hGyqzD>

Aşağıdaki şekiller, gerekli sekiz yüz ifadesini gösteren ankete eşlik eden örnekleri göstermektedir.

Aşağıdaki Şekil 3.3 A ve 3.3 B, gerekli sekiz yüz ifadesini gösteren ankete eşlik eden örnekleri göstermektedir.

» KORKU DURUMU

Lütfen korku durumunu ifade eden bir fotoğrafınızı çekin (lütfen bu durumla hakkında uç fotoğraf çekin)



* İlk resim korku ile ilgili

Click here to upload file. (< 10MB)

* İkinci resim korku ile ilgili

Click here to upload file. (< 10MB)

* Üçüncü resim korku ile ilgili

Click here to upload file. (< 10MB)

» ÖFKE DURUMU

Lütfen öfke halinizi ifade eden bir fotoğrafınızı çekin (lütfen bu durumla ilgili uç fotoğraf çekin)



* İlk resim öfke durumuyla ilgili

Click here to upload file. (< 10MB)

* İkinci resim öfke durumuyla ilgili

Click here to upload file. (< 10MB)

* Üçüncü resim öfke durumuyla ilgili

Click here to upload file. (< 10MB)

» TARSIZLIK VEYA NÖTR DURUMU

Lütfen tarafsızlık veya Nötr durumu ifade eden bir fotoğrafınızı çekin (lütfen bu durumla uç fotoğrafınızı çekin)



* İlk resim tarafsızlık veya Nötr durumu hakkındadır

Click here to upload file. (< 10MB)

* İkinci resim tarafsızlık durumu veya Nötr durumla ilgilidir.

Click here to upload file. (< 10MB)

* Üçüncü resim, tarafsızlık veya Nötr durumuyla ilgilidir.

Click here to upload file. (< 10MB)

» TİKSİNTİ DURUMU

Lütfen tiksinti halinizi ifade eden bir fotoğrafınızı çekin (lütfen bu durumla ilgili uç fotoğraf çekin)



* İlk resim tiksinti durumuyla ilgili

Click here to upload file. (< 10MB)

* İkinci resim tiksinti durumuyla ilgili

Click here to upload file. (< 10MB)

* Üçüncü resim tiksinti durumuyla ilgili

Click here to upload file. (< 10MB)

Şekil 3.3 A Veri toplama için geliştirilen araç kutusunun örnek ekran görüntüleri

» DİKKAT DURUMU

Lütfen dikkat durumunu ifade eden bir fotoğrafınızı çekin (yoğunlaşma hali) (Lütfen bu durumun üç fotoğrafını çekin)



*İlk resim dikkat durumuyla ilgili (yoğunlaşma hali)

Click here to upload file. (< 10MB)

*İkinci resim dikkat durumuyla ilgili (yoğunlaşma hali)

Click here to upload file. (< 10MB)

*Üçüncü resim dikkat durumuyla ilgilidir (yoğunlaşma hali)

Click here to upload file. (< 10MB)

» ÜZÜNTÜ DURUMU

Lütfen üzüntü halinizi ifade eden bir fotoğrafınızı çekin (lütfen bu durumla ilgili üç fotoğraf çekin)



*İlk resim üzüntü haliyle ilgili

Click here to upload file. (< 10MB)

*İkinci resim üzüntü haliyle ilgili

Click here to upload file. (< 10MB)

*Üçüncü resim üzüntü haliyle ilgili

Click here to upload file. (< 10MB)

» ŞAŞIRMA DURUMU

Lütfen Şaşırma halinizi ifade eden bir fotoğrafınızı çekin (Lütfen bu durum için üç fotoğraf çekiniz)



*İlk resim Şaşırma durumu hakkındadır

Click here to upload file. (< 10MB)

*İkinci resim Şaşırma durumu hakkındadır

Click here to upload file. (< 10MB)

*Üçüncü resim Şaşırma durumu hakkındadır

Click here to upload file. (< 10MB)

» MUTLULUK DURUMU

Lütfen mutluluk halinizi ifade eden bir fotoğrafınızı çekin (lütfen bu durumla ilgili üç fotoğraf çekin)



*İlk resim mutluluk hali hakkındadır

Click here to upload file. (< 10MB)

*İkinci resim mutluluk hali hakkındadır

Click here to upload file. (< 10MB)

*Üçüncü resim mutluluk hali hakkındadır

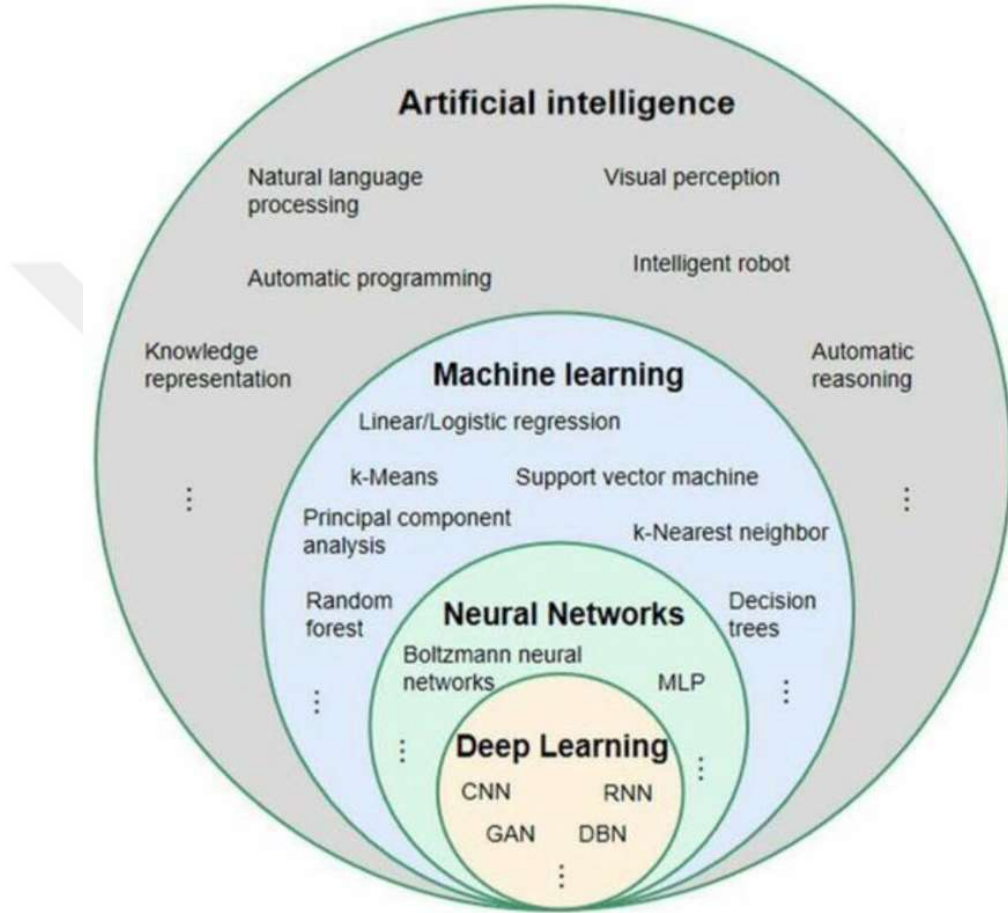
Click here to upload file. (< 10MB)

Şekil 3.3 B Veri toplama için geliştirilen araç kutusunun örnek ekran görüntüleri

3.2. Derin Öğrenme

Derin öğrenme, makine öğreniminin bir alt kümesidir ve aynı zamanda bilgisayarları insan beynine benzer bir şekilde verileri analiz etmek üzere eğiten bir yapay zekâ (AI) teknolojisidir. Derin öğrenme modelleri, doğru iç görüler ve tahminler üretmek için görüntü, metin, ses ve diğer verilerdeki karmaşık örüntüleri tanıma yeteneğine sahiptir. Derin öğrenme algoritmaları, fotoğrafları tanımlamak veya bir ses kaydını metne dönüştürmek gibi normalde insan zekâsı gerektiren görevleri otomatikleştirmek için kullanılabilir. Derin öğrenme algoritmaları insan beynini taklit etmek üzere tasarlanmış sinir ağlarıdır. Örneğin

bir insan beyni, bilgiyi öğrenmek ve işlemek için beraber çalışan milyonlarca bağlantılı nörondan oluşur. Yapay sinir ağları olarak da bilinen derin öğrenme sinir ağları, bilgisayar içinde iş birliği yapan çok sayıda yapay nöron katmanından oluşur. Yapay nöronlar ise matematiksel hesaplamalar yoluyla verileri işleyen düğüm adı verilen yazılım modülleridir. Karmaşık sorunların üstesinden gelmek için bu düğümleri kullanan derin öğrenme teknikleri yapay sinir ağları olarak bilinir (Şekil 3.4).

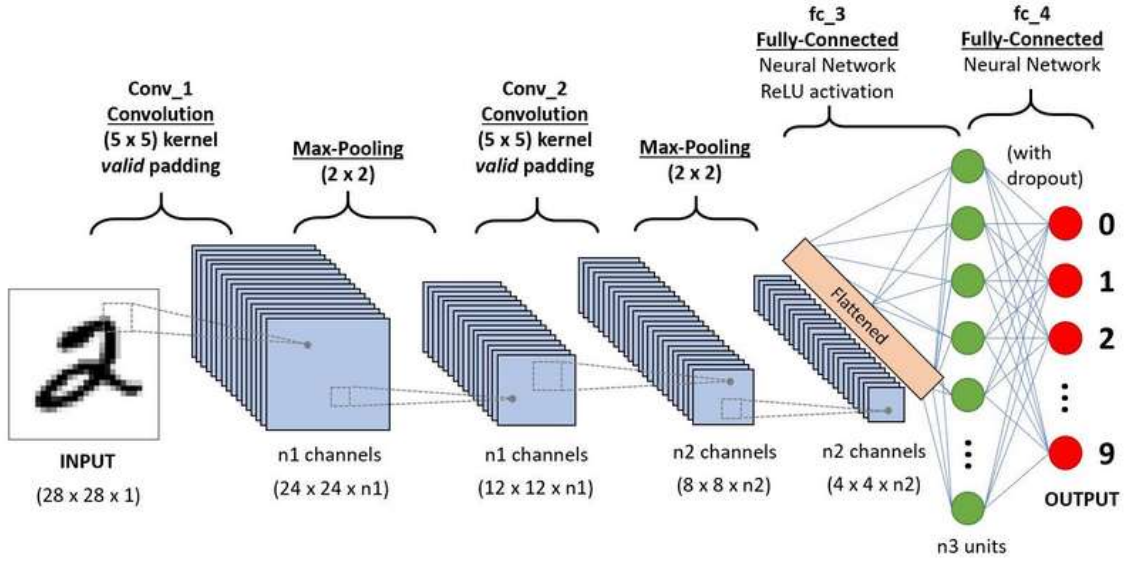


Şekil 3.4 Derin öğrenmenin yapay zekâ haritasındaki yeri (Li ve ark.,2021)

3.2.1.CNN ağları

Konvolüsyonel sinir ağı (CNN veya convnet), makine öğreniminin bir alt kümesidir. CNN, görüntü / nesne tanıma ve sınıflandırma için yaygın olarak kullanılan derin öğrenme algoritmalarından türetilen bir tür yapay sinir ağı ve ağ mimarisidir. Derin Öğrenme, bir CNN kullanarak görüntülerdeki nesneleri tanır. CNN, giriş görüntülerindeki çizgiler, gradyanlar, daireler ve hatta gözler ve yüzler gibi tasarımları tespit etmede çok iyidir. Bu özellik, evrimsel sinir ağlarını bilgisayarla görme için bu kadar etkili kılan şeydir.

CNN' in klasik mimarileri:



Şekil 3.5 CNN'in klasik mimarileri

Şekil 3.5'de gösterildiği gibi, geleneksel CNN mimarileri birçok konvolüsyon katmanı, havuzlama katmanı (Max veya ortalama havuzlama katmanı ile birlikte), tam bağlantılı katman ve son katman olarak sınıflandırıcıdan oluşur. Ancak her katmanın çıktısını nasıl hesaplayabilir ve ağı her katmanında öğrenmek için gereken parametrelerin miktarlarını nasıl hesaplayabilir? Bu soruyu cevaplamak için aşağıdaki örneği ele alabiliriz:

Şekil 3.6'de gösterildiği gibi, hesaplama için iki görevimiz vardır. İlk olarak her katman için çıktı boyutunu hesaplamak ve sonrasında her katman için parametre sayısını hesaplamak.

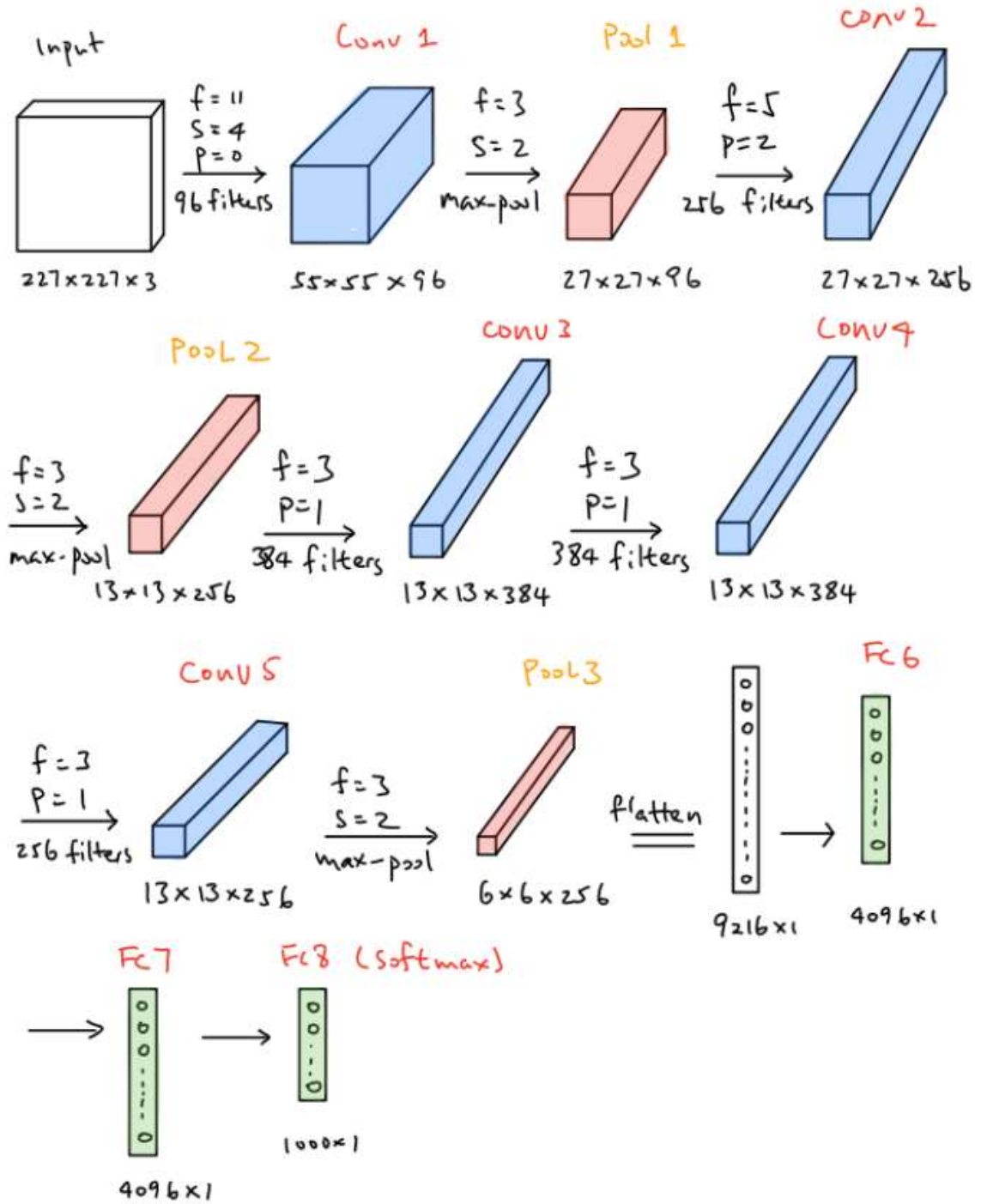
İlk görev: her katman için çıktı boyutunun hesaplanması:

Aşağıdaki denklemlere bakılırsa:

$$W_{out} = \frac{W_{in} - K + 2P}{S} + 1 \quad (3.1)$$

$$H_{out} = \frac{H_{in} - K + 2P}{S} + 1 \quad (3.2)$$

Burada W_{in} ve H_{in} , görüntünün boyutlarını, $K \times K$ çekirdek veya filtrenin boyutlarını (bazı kaynaklarda $F \times F$ olarak da belirtilmiştir), P , padding'i, S , stride'ı temsil eder.



Şekil 3.6 AlexNet CNN ağı için aritmetik mimari (Guan,2020)

Şekil 3.6'den ve önceki denklemlerden çıkış katmanlarının boyutunu hesaplayabiliriz. İlk konvolüsyon katmanı F ya da $K=11$, $S=4$, $P=0$, Filtre sayısı=96 olmak üzere çok sayıda parametreye sahiptir. Giriş görüntüsünün boyutu $227 \times 227 \times 3$ ' tür. Buradan, $W_{out} = H_{out} = \frac{227-11+0}{4} + 1 = 55$ olur ki, bu değer $W_{out} \times H_{out} \times \text{Filtre Sayısı}$ denkleminde yerine yazılırsa ilk konvolüsyon katmanının boyutunun $55 \times 55 \times 96$ olduğunu ortaya çıkar. İlk havuzlama

katmanı için aynı hesaplamalar yapıldığında, $W_{out}=H_{out}=\frac{55-3+}{2} + 1 = 27$ elde edilir ki, bu da ilk havuzlama katmanının boyutunun $27 \times 27 \times 96$ olduğunu gösterir. Yine aynı denklemler kullanılarak ikinci konvolüsyon katmanının boyutunun, $W_{out}=H_{out}=\frac{27-5+4}{1} + 1 = 27$ ve bunun sonucunda $27 \times 27 \times 256$ olduğu bilgisine ulaşılır. Tam bağlantılı katman, katmandaki nöron sayısına eşit uzunlukta bir vektör çıkışı üretir.

İkinci görev: her katman için parametre sayısının hesaplanması:

Her katman için parametre sayısını aşağıdaki denklemlerle hesaplayabiliriz:

1- Bir konvolüsyon katmanının parametre sayısı aşağıdaki denklemlerle (3.3) bulunabilir.

$$W_c = K^2 \times C \times N \text{ halbuki } B_c = N \text{ ve } P_c = W_c + B_c \quad (3.3)$$

W_c : Konvolüsyon katmanının ağırlıklarının sayısı.

B_c : Konvolüsyon katmanının önyargı (bias) sayısı veya filtre sayısı

P_c : Konvolüsyon katmanının parametre sayısı.

K : Konvolüsyon katmanında kullanılan çekirdeklerin boyutu (genişliği).

N : Çekirdek veya filtre sayısı

C : Giriş görüntüsünün kanal sayısı

Verilen denkleme göre, ilk konvolüsyon katmanı için parametre sayısı hesaplanabilir:

$$W_c = 11^2 \times 3 \times 96 = 34848, B_c = 96, \text{ dolayısıyla } P_c = 34944.$$

2- MaxPool Katmanı parametrelerinin sayısı: MaxPool katmanının ilişkili parametreleri yoktur. Havuz boyutu, stride uzunluğu ve padding, hiper parametrelerdir.

3- Tam bağlı (FC) katmanın parametre sayısı:

Bir CNN iki tür tam bağlantılı katmana sahiptir. İlk FC katmanı son konvolüsyon katmanına bağlanırken, sonraki FC katmanları diğer FC katmanlarına bağlanır.

Durum 1: Bir konvolüsyon katmanına bağlı Tam Bağlı (FC) katmanın parametre sayısı için denklemler (3.4), (3.5), &(3.7).

$$W_{cf} = O^2 \times N \times F \quad (3.4)$$

$$B_{cf} = F \quad (3.5)$$

$$P_{cf} = W_{cf} + B_{cf} = (O^2 \times N + 1) \times F \quad (3.6)$$

W_{cf} : Bir konvolüsyon katmanına bağlı olan bir FC katmanının ağırlık sayısı.

B_{cf} : Bir konvolüsyon katmanına bağlı olan bir FC Katmanının önyargı (bias) sayısı

O : Bir önceki konvolüsyon katmanının çıkış görüntüsünün boyutu (genişliği).

N : Önceki konvolüsyon katmanındaki çekirdek veya filtre sayısı.

F : FC katmanındaki nöron sayısı.

P_{cf} : Bir konvolüsyon katmanına bağlı olan bir FC katmanının parametre sayısı

FC-1 katmanı için; $O=6$, $N=256$, $F=4096$ parametreleri kullanılarak bu denklemde yerine yazılırsa,

$W_{cf} = 6^2 \times 256 \times 4096 = 37748736$, $B_{cf}=4096$, $P_{cf}=37752832$ değerleri elde edilir.

Durum 2: FC katmanına bağlı bir Tam Bağlantılı (FC) katmanın parametre sayısı için denklem.

W_{ff} : Bir FC katmanına bağlı olan bir FC katmanının ağırlık sayısı.

B_{ff} : Bir FC katmanına bağlı olan bir FC katmanının ön yargı (bias) sayısı.

P_{ff} : Bir FC katmanına bağlı olan bir FC katmanının parametre sayısı

F : FC katmanındaki nöron sayısı.

F_{-1} : Bir önceki FC katmanındaki nöron sayısı.

$$W_{ff} = F_{-1} \times F \quad (3.7)$$

$$B_{ff} = F \quad (3.8)$$

$$P_{ff} = W_{ff} + B_{ff} = (F_{-1} + 1) \times F \quad (3.9)$$

AlexNet' in son tam bağlı katmanı FC katmanına bağlıdır. Bu katman için $F_{-1} = 4096$ ve $F=1000$ parametreleri kullanılarak denklemde yerine yazılırsa,

$W_{ff}=4096 \times 1000=4096000$, $B_{ff}=1000$, böylece $P_{ff}=4097000$ değerleri elde edilir.

Her katman için boyut aktivasyonunu ve parametre sayısını göstermek için Çizelge 3.1 oluşturulabilir.

Çizelge 3.1 AlexNet ağındaki her katman için aktivasyon boyutu ve parametre sayısı

Katman adı	Aktivasyon Boyutu	Ağırlıklar	Filtre sayısı	Parametre Sayıları
Giriş Görüntüsü	227x227x3	0	0	0
CONV1(f=11,s=4,p=0,96 filters)	55x55x96	34,848	96	34,944
MaxPool-1 (f=3,s=2)	27x27x96	0	0	0
CONV2(f=5,s=1,p=2,256 filters)	27x27x256	614,400	256	614,656
MaxPool-2 (f=3,s=2)	13x13x256	0	0	0
Conv-3 (f=3,s=1,p=1,384 filters)	13x13x384	884,736	384	885,120
Conv-4 (f=3,s=1,p=1,384 filters)	13x13x384	1,327,104	384	1,327,488
CONV5(f=3,s=1,p=1,256 filters)	13x13x256	884,736	256	884,992
MaxPool-3 (f=3,s=2)	6x6x256	0	0	0
FC-6 (4096 neurons)	4096×1	37,748,736	4,096	37,752,832
FC-7(4096 neurons)	4096×1	16,777,216	4,096	16,781,312
FC-8 (1000 neurons)	1000×1	4,096,000	1,000	4,097,000
Çıkış	1000×1	0	0	0
Toplam				62,378,344

3.2.2. CNN ağlarının katmanları

3.2.2.1. Evrişimsel katman

Bu, giriş görüntülerinden farklı özellikleri çıkarmak için kullanılan ilk katmandır. Bu katman, giriş görüntüsü ile $M \times M$ boyutundaki bir filtre arasında matematiksel konvolüsyon işlemini gerçekleştirir. Filtre ile giriş görüntüsünün filtre boyutuna ($M \times M$) göre bölümleri arasındaki nokta çarpımı, filtrenin giriş görüntüsü boyunca kaydırılmasıyla hesaplanır. Giriş ve filtrenin nokta çarpımları serisinden elde edilen nihai çıktı özellik haritası, aktivasyon haritası veya bir konvolüsyonlu özellik olarak bilinir (bkz. Şekil 3.7).

Her konvolüsyon işleminden sonra, bir CNN özellik haritasına bir Doğrultulmuş Doğrusal Birim (ReLU) dönüşümü uygulanarak modele doğrusal olmayan bir özellik katar (Sayracı, 2021).

0	0	0	0	0	0	0
0	60	113	56	139	85	0
0	73	121	54	84	128	0
0	131	99	70	129	127	0
0	80	57	115	69	134	0
0	104	126	123	95	130	0
0	0	0	0	0	0	0

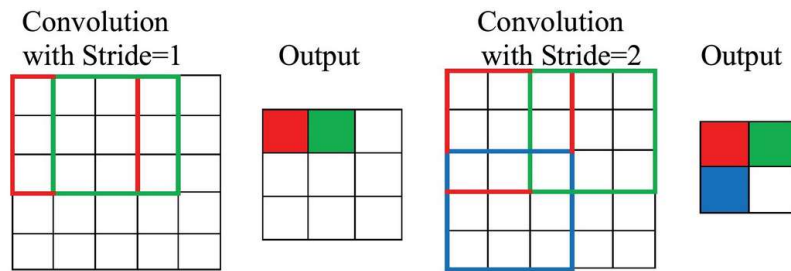
Kernel		
0	-1	0
-1	5	-1
0	-1	0

114	328	-26	470	158
53	266	-61	-30	344
403	116	-47	295	244
108	-135	256	-128	344
314	346	279	153	421

Şekil 3.7 Evrişimsel Katman

3.2.2.2. Padding ve stride

Padding ve stride, konvolüsyon tekniğinin nasıl gerçekleştirildiğini etkiler. Padding ve stride, giriş/çıkış vektörlerinin boyutlarını (yükseklik ve genişlik) genişletmek veya azaltmak için kullanılabilir. Padding, konvolüsyonel sınır ağlarında CNN çekirdeği işleme sırasında bir görüntüye kaç piksel eklendiğini belirtmek için kullanılan bir terimdir. Örneğin, bir CNN'deki padding sıfır olarak ayarlanırsa, her piksele eklenen değer 0 olacaktır. CNN tarafından işlenen bir görüntüye padding eklemek daha doğru görüntü analizi sağlarken, stride filtrenin giriş matrisi üzerinde nasıl konvolüsyon yaptığını, yani kaç piksel kaydırıldığını kontrol eder. Stride parametresi 1 olarak ayarlandığında, filtre her seferinde bir piksel boyunca hareket eder; stride parametresi 2 olarak ayarlandığında, filtre her seferinde iki piksel boyunca hareket eder. Stride değeri azaldığında çıktı daha küçük olur ve bunun tersi de geçerlidir (Pandey, 2021) (bkz. Şekil 3.8).



Şekil 3.8 Evrişimsel Katmanda Padding ve Stride

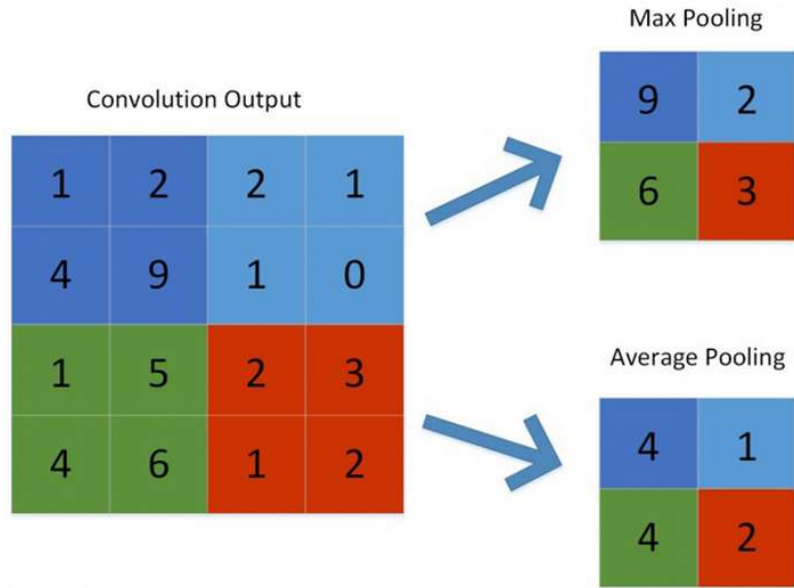
3.2.2.3. Havuzlama katmanı

Aşağı örnekleme olarak da bilinen havuzlama katmanları, girdideki parametre sayısını azaltmak için boyutsal azaltma yapar. Havuzlama, evrişimsel katmanlar gibi, tüm girdi boyunca bir filtreyi tarar, ancak bu filtrenin ağırlıkları yoktur. Bunun yerine çekirdek, alıcı alan verilerine bir toplama işlevi uygulayarak çıktı dizisini oluşturur. Havuzlama katmanı, çevredeki çıktıların bir özet istatistiğini hesaplayarak belirli noktadaki ağ çıktısının yerini alır. Bu, temsilin uzamsal boyutunu azaltmaya yardımcı olur, bu da gerekli hesaplama ve ağırlık miktarını azaltır. Havuzlama işlemi, temsilin her bir dilimi üzerinde ayrı ayrı gerçekleştirilir. Havuzlama iki türe ayrılır:

1-Maksimum havuzlama: Filtre, özellik haritasından geçerken çıkış dizisine göndermek için en yüksek değere sahip pikseli seçer.

2-Ortalama havuzlama: Filtre özellik haritasını dolaşırken, alıcı alan içindeki ortalama değeri hesaplar ve bunu çıkış dizisine gönderir.

Havuzlama, temsilin uzamsal boyutunu kademeli olarak azaltarak ağıdaki parametrelerin ve hesaplamaların sayısını azaltır ve aşırı uyumu önler. Havuzlama kullanılmazsa, çıktı girdiyle aynı çözünürlüğe sahip olur (bkz. Şekil 3.9).

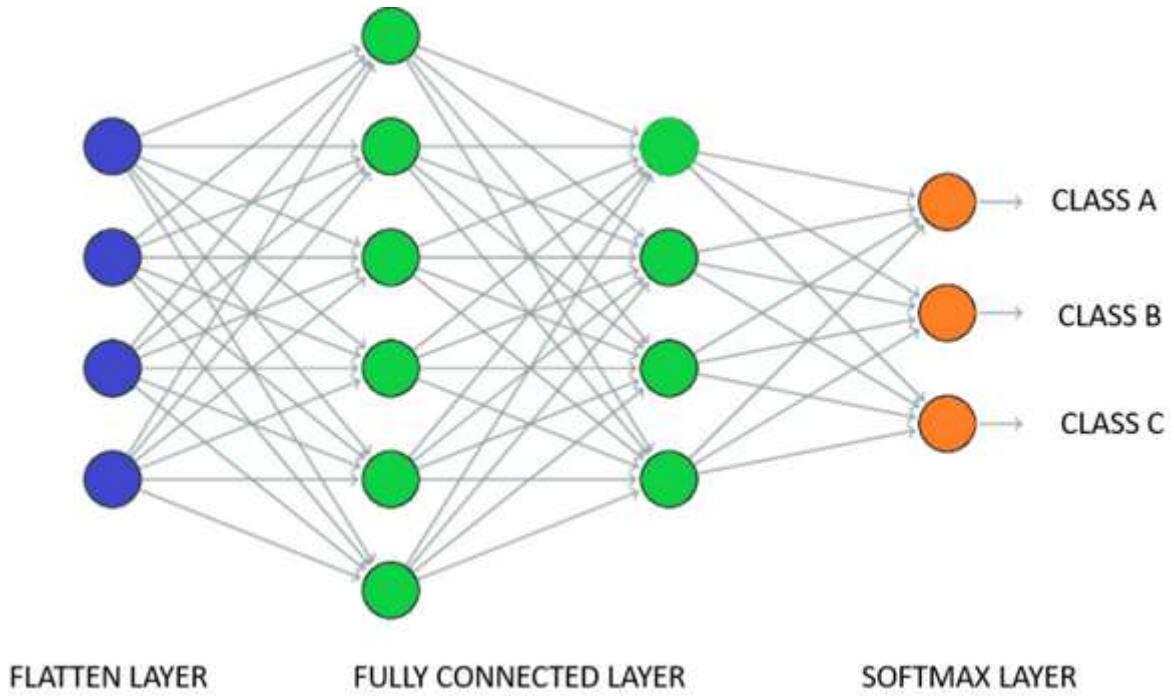


Şekil 3.9 Havuzlama katmanı (Kalita, 2022)

3.2.2.4. Tam bağlı katman (FC)

Konvolüsyonel sinir ağında, tam bağlantılı bir katman, bir önceki katmandaki her nöron ile mevcut katmandaki her nöron arasında bağlantıya sahip olan katmandır. Bu, tamamen bağlı

katmandaki her nöronun önceki katmandaki her nörona girdi aldığı ve tamamen bağlı katmandaki her nöronun ağırlıklı çıktısına katkıda bulunabileceği anlamına gelir. Bu katmanın nöronları, önceki ve sonraki katmanlardaki tüm nöronlara tamamen bağlıdır. Nöronların yanı sıra ağırlıkları ve önyargıları da içeren Tam Bağlı (FC) katman, farklı katmanlardaki nöronları bağlamak için kullanılır. Bu katman genellikle çıktı katmanından önce yerleştirilir ve bir CNN mimarisinin son birkaç katmanını oluşturur. Bu katman, önceki katmanlar ve bunların çeşitli filtreleri tarafından elde edilen özelliklere dayalı olarak kategorizasyon gerçekleştirir. Konvolüsyonel ve havuzlama katmanları ReLu fonksiyonlarını kullanma eğilimindeyken, FC katmanları genellikle girdileri uygun şekilde sınıflandırmak için bir softmax aktivasyon fonksiyonundan yararlanır (bkz. Şekil 3.10).

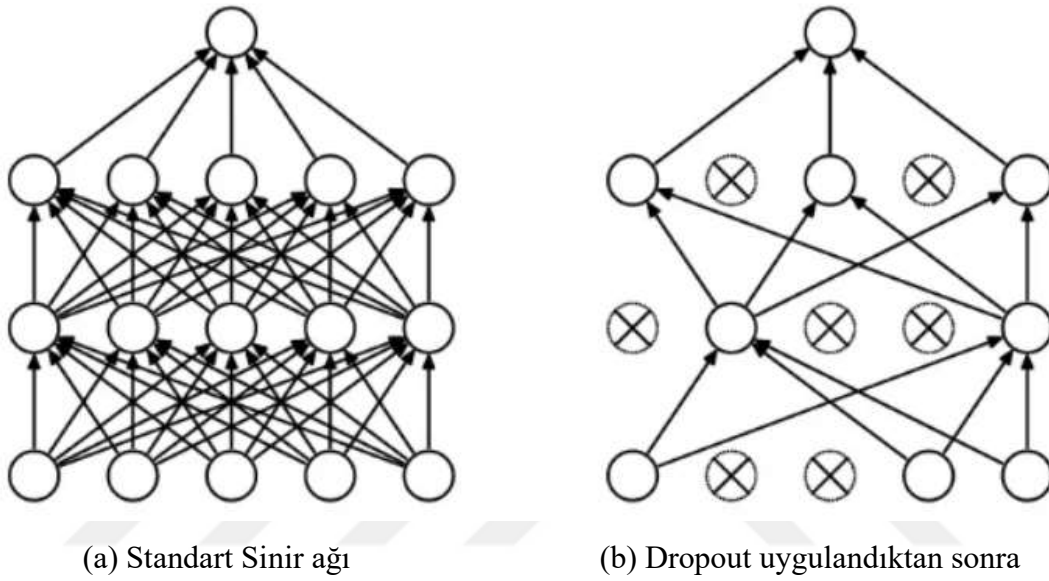


Şekil 3.10 Tam bağlantılı katman (Sayracı, 2021)

3.2.2.5. Dropout katmanı

CNN ayrıca bir Dropout katmanına sahiptir. Dropout katmanı, diğer tüm nöronların işlevselliğini korurken bazı nöronların bir sonraki katmana katkılarını ortadan kaldıran bir maske işlevi görür. Bir girdi vektörüne Dropout katmanı uygulandığında, bazı özellikleri ortadan kaldırılır; ancak bunu bir gizli katmana uyguladığımızda, bazı gizli nöronlar ortadan kaldırılır. Dropout katmanları CNN eğitiminde önemlidir çünkü eğitim verilerinin aşırı uyumunu önlerler. Eğer bunlar yoksa ilk eğitim örnekleri kümesi öğrenme üzerinde anormal derecede güçlü bir etkiye sahip olur. Sonuç olarak, yalnızca daha sonraki örneklerde veya gruplarda ortaya çıkan özellikler öğrenilemeyecektir. Tüm özellikler FC katmanına

bağlandığında eğitim veri setinde aşırı uyum yaygındır. Aşırı uyum, bir model eğitim verilerinde çok iyi performans gösterdiğinde, yeni verilere uygulandığında modelin performansı üzerinde olumsuz bir etkiye sahip olduğunda gerçekleşir. Bu sorunu ele almak için, eğitim süreci sırasında birkaç nöronun sinir ağından çıkarıldığı ve daha küçük bir modelle sonuçlanan bir dropout katmanı kullanıldığı görülür. Örneğin, gizli katmanlarda 500 nöron (düğüm) varsa ve 0,5 düşme olasılığına sahip bir dropout kullanılıyorsa, her iterasyonda (batch) 250 nöron rastgele silinecektir (bkz. Şekil 3.11 a&b).



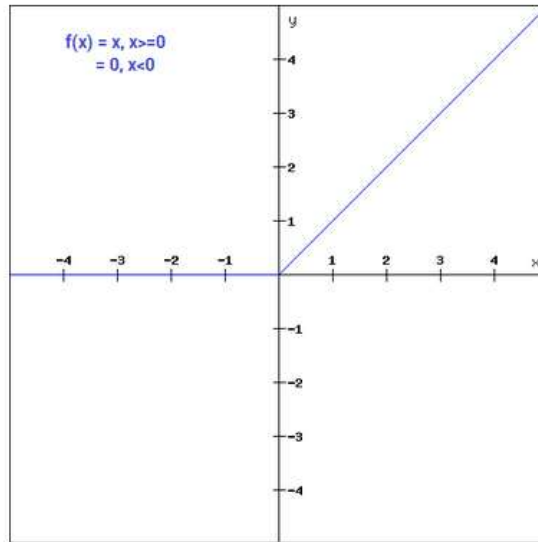
Şekil 3.11 Dropout uygulanmamış ve uygulanmış standart sinir ağı (Yang ve Yang, 2018).

3.2.2.6. Doğrusal olmayan katmanlar

Ağa doğrusal olmama özelliği ekler. Doğrusal olmayan katmanlar sıklıkla konvolüsyon katmanından hemen sonra eklenerek aktivasyon haritasına doğrusal olmama özelliği eklenir çünkü konvolüsyon doğrusal bir işlemdir ve görüntüler doğrusal olmaktan uzaktır. Aktivasyon fonksiyonu CNN modelinde kritik bir unsurdur. Ağ değişkenleri arasındaki her türlü sürekli ve karmaşık ilişkiyi öğrenmek ve tahmin etmek için kullanılır. Ağı sonunda modeldeki hangi bilgi bitlerinin öne çıkması ve hangilerinin çıkmaması gerektiğini belirler. ReLU, Softmax, TanH ve Sigmoid fonksiyonları dâhil olmak üzere düzenli olarak kullanılan çeşitli aktivasyon fonksiyonları vardır. Bu fonksiyonların her birinin özel bir kullanımı vardır. İkili sınıflandırma yapan bir CNN modeli için sigmoid ve softmax fonksiyonları tercih edilirken, çok sınıflı sınıflandırma için genellikle softmax kullanılır. CNN modelindeki aktivasyon fonksiyonları, bir nöronun etkinleştirilip etkinleştirilmeyeceğini belirler.

3.2.2.6.1 ReLU fonksiyonu (Rectified Linear Unit-Doğrultulmuş Doğrusal Birim)

Doğrusal olmayan bir aktivasyon fonksiyonu ReLU fonksiyonudur. Bu, özellikle Evrişimsel Sinir Ağlarında yaygın olarak kullanılan bir aktivasyon fonksiyonudur. Doğrultulmuş Doğrusal Birim, ReLU olarak kısaltılır. ReLU fonksiyonunu kullanmanın diğer aktivasyon fonksiyonlarına göre en önemli avantajı, tüm nöronları aynı anda uyarmamasıdır. Bu, nöronların yalnızca doğrusal dönüşüm sonucu 0'dan küçükse susturulacağı anlamına gelir. Doğrusal olmaması, hataları kolayca geri yayabileceği ve birçok nöron katmanını etkinleştirebileceği anlamına gelir. Daha basit matematiksel hesaplamalar gerektirdiği için ReLU, Tanh ve Sigmoid'e göre hesaplama açısından daha az maliyetlidir. Bir seferde yalnızca birkaç nöron aktiftir, bu da ağı seyrek ve hesaplama için verimli hale getirir. ReLU, $f(x) = \max(0, x)$ olarak tanımlanır (bkz. Şekil 3.12).

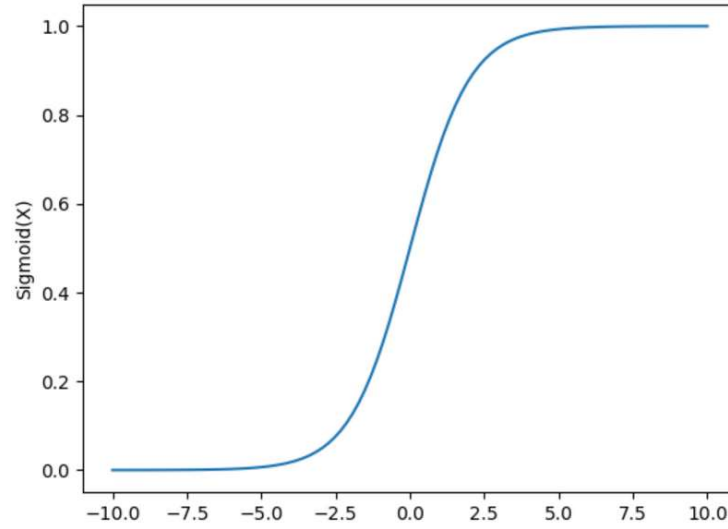


Şekil 3.12 ReLU fonksiyonu (Gupta, 2020)

3.2.2.6.2 Sigmoid fonksiyonu

Lojistik fonksiyon, sigmoid aktivasyon fonksiyonunun bir diğer adıdır. Lojistik regresyon sınıflandırma yaklaşımı da aynı fonksiyonu kullanır. Fonksiyon herhangi bir gerçek değeri girdi olarak kabul eder ve 0 ile 1 arasında değişen değerler döndürür. Girdi ne kadar büyükse (daha pozitif), çıktı 1.0'a o kadar yakın olurken, girdi ne kadar küçükse (daha negatif), çıktı 0.0'a o kadar yakın olur. Sigmoid fonksiyonlar tipik olarak sonucun 0 ya da 1 olduğu ikili sınıflandırmanın çıkış katmanında kullanılır. Sigmoid fonksiyonlar için değer yalnızca 0 ile 1 arasında olduğundan, değer 0,5'ten büyükse sonucun 1, aksi takdirde 0 olması beklenebilir (Sharma, 2022) (bkz. Şekil 3.13). Sigmoid aktivasyon fonksiyonu aşağıdaki gibi hesaplanır:

$$F(x) = 1/(1 + e^{-x}) \quad (3.10)$$

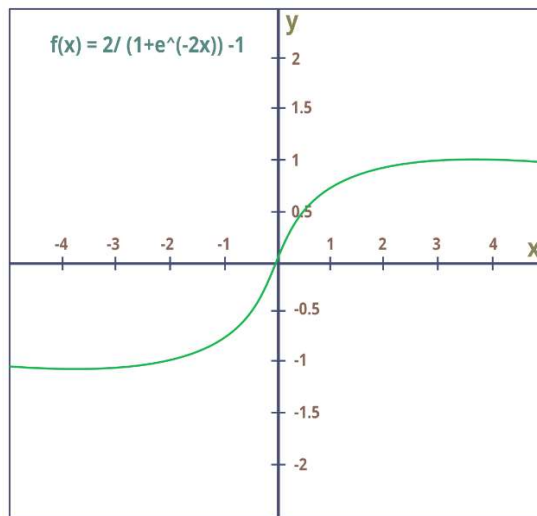


Şekil 3.13 Sigmoid fonksiyonu

3.2.2.6.3 Tanh (Tanjant Hiperbolik) fonksiyonu

Tanh fonksiyonları sigmoid fonksiyonlara oldukça benzer. Tek farkı orijin etrafında simetrik olmasıdır. Aslında geometrik olarak kaydırılmış bir sigmoid fonksiyonudur. Her ikisi de birbirleriyle ilişkilidir ve birbirinden çıkarılabilir; olası değerler -1 ile 1 arasında değişir. Fonksiyon herhangi bir gerçek değeri girdi olarak kabul eder ve -1 ile 1 arasında değişen değerler döndürür. Girdi ne kadar büyükse (daha pozitif), çıktı 1.0' a o kadar yakın olurken, girdi ne kadar küçükse (daha negatif), çıktı -1.0' a o kadar yakın olur. Sonuç olarak, sonraki seviyelerin girdileri her zaman aynı işaretle olmayacaktır (Gupta, 2020) (bkz. Şekil 3.14). Tanh aktivasyon fonksiyonu (3.11) bu şekilde hesaplanır:

$$\text{Tanh}(x) = 2\text{sigmoid}(2x) - 1 = (2/(1 + e^{-2x})) - 1 \quad (3.11)$$



Şekil 3.14 Tanh fonksiyonu

3.2.2.6.4 Softmax fonksiyonu

Softmax, sigmoidin daha genelleştirilmiş bir versiyonudur. Çok sınıflı sınıflandırmayı içeren problemlerde kullanılır. 0-1 aralığında değerler ürettiği için sınıflandırma modellerinde son katman olarak kullanılır. Softmax fonksiyonu sıklıkla sigmoidal bir kombinasyon olarak tanımlanır. Sigmoid fonksiyonunun 0 ile 1 arasında değerler ürettiğini biliyoruz; bu değerler bir veri noktasının belirli bir sınıfa ait olma olasılığı olarak yorumlanabilir. Sonuç olarak, sigmoid fonksiyonu genellikle ikili sınıflandırma konularında kullanılır. Çok sınıflı sınıflandırma sorunları için softmax fonksiyonu kullanılabilir. Bu fonksiyon, her bir veri noktası sınıfı için olasılığı döndürür (bkz. Şekil 3.15). Softmax aktivasyon fonksiyonunun matematiksel (Gülüm, 2023) (3.12)ifadesi şöyledir:

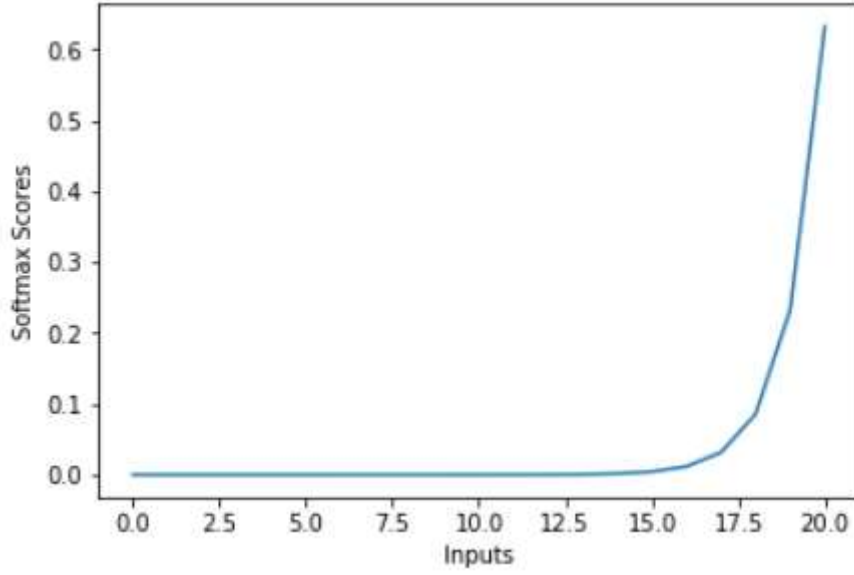
$$\text{Softmax}(Z_i) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (3.12)$$

Z_i : Softmax fonksiyonunun $(z_1, \dots, z_K)$ 'den oluşan girdi vektörü. z_i değerlerinin tümü softmax fonksiyonunun girdi vektörünün elemanlarıdır ve pozitif, sıfır veya negatif herhangi bir gerçek değer alabilirler. Örneğin bir sinir ağı $(-0.62, 8.12, 2.53)$ gibi bir vektör üretebilir.

e^{z_i} : Giriş vektörünün her bir üyesi geleneksel üstel fonksiyona tabi tutulur. Bu, sıfırdan büyük pozitif bir değer üretir; bu değer, giriş negatifse çok küçük, giriş büyükse çok büyüktür.

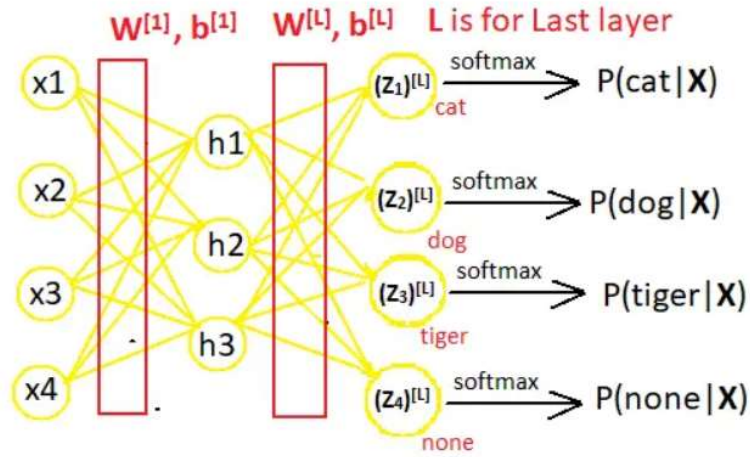
$\sum_{j=1}^K e^{z_j}$: Normalleştirme terimi formülün en altındaki terimdir. Fonksiyonun tüm çıktı değerlerinin toplamının 1 olmasını ve $(0, 1)$ aralığında olmasını sağlayarak geçerli bir olasılık dağılımı elde edilmesini sağlar.

K: Çok sınıflı sınıflandırıcıdaki sınıf sayısı.



Şekil 3.15 Softmax fonksiyonu

Softmax fonksiyonunun bir örneği kedi, kaplan ve köpek sınıflandırmasıdır (bkz. Şekil 3.16).



Şekil 3.16 CNN ağı için son katmanda softmax fonksiyonunun çalışması

Yukarıdaki Şekil 13'te belirtildiği gibi, normalde bir sinir ağının son katmanında bir softmax aktivasyon fonksiyonu kullanılır. Yukarıda gösterilen sinir ağı için;

$$Z^{[L]} = \begin{bmatrix} z_1^{[L]} \\ z_2^{[L]} \\ z_3^{[L]} \\ z_4^{[L]} \end{bmatrix} = W^L * H^{L-1} + b^L \text{ 'dir (3.13). Burada, } H^{L-1} = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} \quad (3.14) \quad (L-1) \text{ olarak bilinen}$$

$$\text{gizli katmandaki hesaplanan değerlerdir. } W^L = \begin{pmatrix} W_{11} & \cdots & W_{1n} \\ \vdots & \ddots & \vdots \\ W_{m1} & \cdots & W_{mn} \end{pmatrix} \quad (3.15) ; W^L, \text{ son}$$

katman için ağırlık matrisi; m, L-1 katmanındaki toplam düğüm sayısı ve n, çıkış katmanı

L'deki düğüm sayısıdır. Bu örnekte, m = 3, n = 4 ve $b^L = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ önyargı matrisidir.

Verilen X girişinin sınıflardan birine ait olma olasılığı şu şekilde hesaplanır (3.12) den:

$$P(\text{cat}|X) = \frac{e^{z_1^{[L]}}}{\sum e^{z^{[L]}}} , P(\text{dog}|X) = \frac{e^{z_2^{[L]}}}{\sum e^{z^{[L]}}} , P(\text{tiger}|X) = \frac{e^{z_3^{[L]}}}{\sum e^{z^{[L]}}} , P(\text{none}|X) = \frac{e^{z_4^{[L]}}}{\sum e^{z^{[L]}}}$$

Bir görüntü girildiğini ve aşağıdaki değerlerin elde edildiğini varsayarsak.

$$(3.13)'den \begin{aligned} Z_1^{[L]} &= 1.25 \\ Z_2^{[L]} &= 2.44 \\ Z_3^{[L]} &= 0.78 \\ Z_4^{[L]} &= 0.12 \end{aligned} \quad \text{Then, } e^{z_1^{[L]}} = e^{1.25} = 3.49, e^{z_2^{[L]}} = e^{2.44} = 11.47$$

$$e^{z_3^{[L]}} = e^{0.78} = 2.18, e^{z_4^{[L]}} = e^{0.12} = 1.27 \text{ ve } \sum e^{z^{[L]}} = 18.41$$

Sonuç olarak olasılık dağılımı şu şekilde hesaplanır:

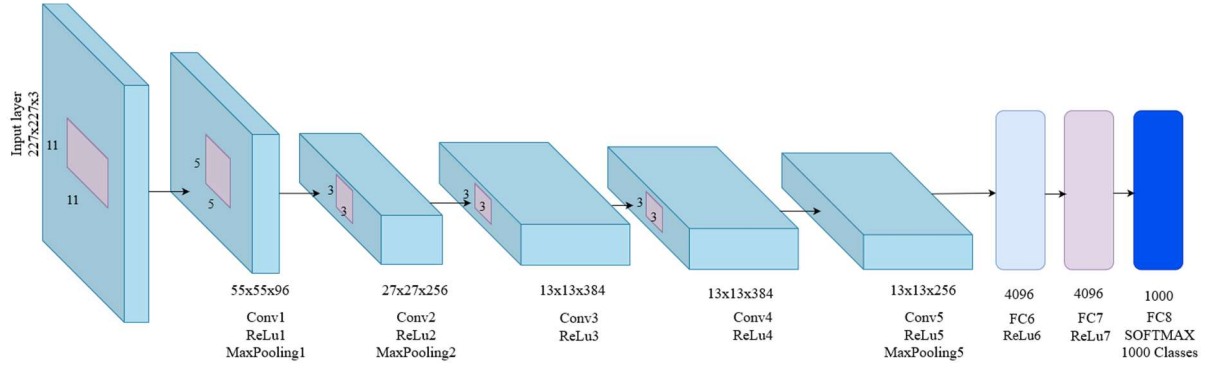
$$\begin{aligned} P(\text{cat}|X) &= \frac{e^{z_1^{[L]}}}{\sum e^{z^{[L]}}} = 0.189, & P(\text{dog}|X) &= \frac{e^{z_2^{[L]}}}{\sum e^{z^{[L]}}} = 0.623, & P(\text{tiger}|X) &= \frac{e^{z_3^{[L]}}}{\sum e^{z^{[L]}}} = 0.12 \\ P(\text{none}|X) &= \frac{e^{z_4^{[L]}}}{\sum e^{z^{[L]}}} = 0.069. \end{aligned}$$

Olasılık dağılımına dayanarak verilen görüntünün bir köpeğe ait olduğu sonucuna varmak mümkündür.

3.2.3. Bazı evrişimli sinir ağı mimarileri

3.2.3.1. AlexNet mimarisi

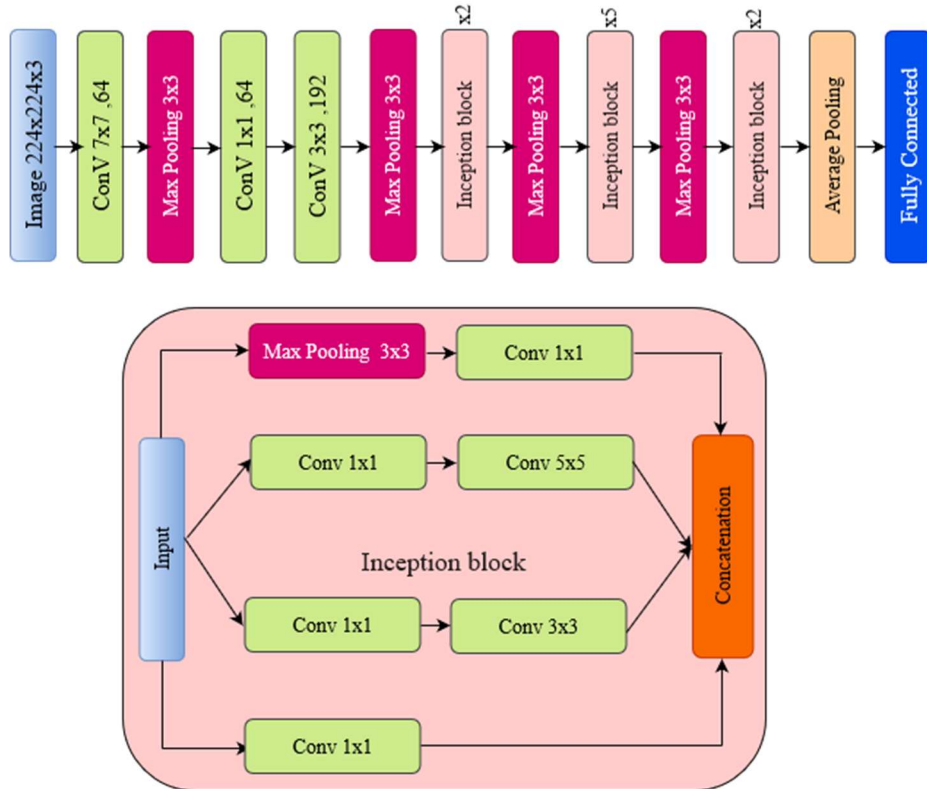
Alex Krizhevsky 2012 yılında AlexNet ağını oluşturmuş ve ImageNet veri kümesi üzerinde eğiterek 1,2 milyon yüksek çözünürlüklü fotoğrafı 1000 farklı sınıfa ayırmış ve bu sayede ImageNet Büyük Ölçekli Görsel Tanıma Yarışması'nda (ILSVRC) ödül almıştır (Krizhevsky, Sutskever, & Hinton, 2017) . AlexNet, beşi evrişimli ve üçü tam bağlantılı olmak üzere sekiz temel katmana sahip bir Derin Evrişimli Sinir Ağıdır (DCNN) (bkz. Şekil 3.17).



Şekil 3.17 AlexNet mimarisi

3.2.3.2. GoogleNet

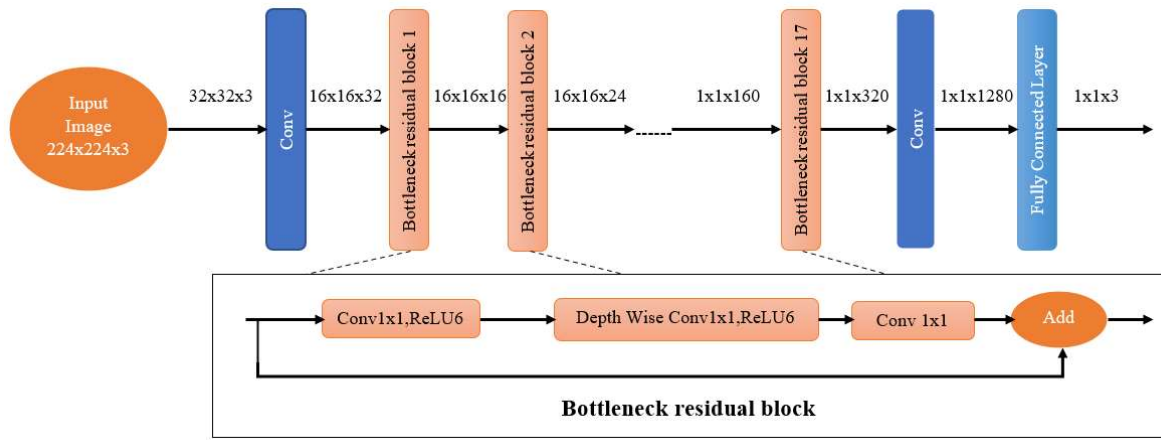
2014 yılında Google araştırmacıları, fotoğrafları 1000 öge kategorisine ayırabilen 18 katmanlı bir DCNN oluşturdu. Bu sayede ImageNet yarışması kazanılmıştır. GoogleNet'in temel evrimsel bloğu Inception bloğu olarak bilinir ve GoogleNet tahminlerini oluşturmak için küresel ortalama havuzlamanın yanı sıra toplam 9 inception bloğundan oluşan bir yığın kullanır. Inception blokları arasında maksimum havuzlama yapılarak boyutsallık azaltılır (Guo ve ark., 2017) (Bkz. Şekil 3.18).



Şekil 3.18 GoogleNet mimarisi

3.2.3.3. MobileNet V2

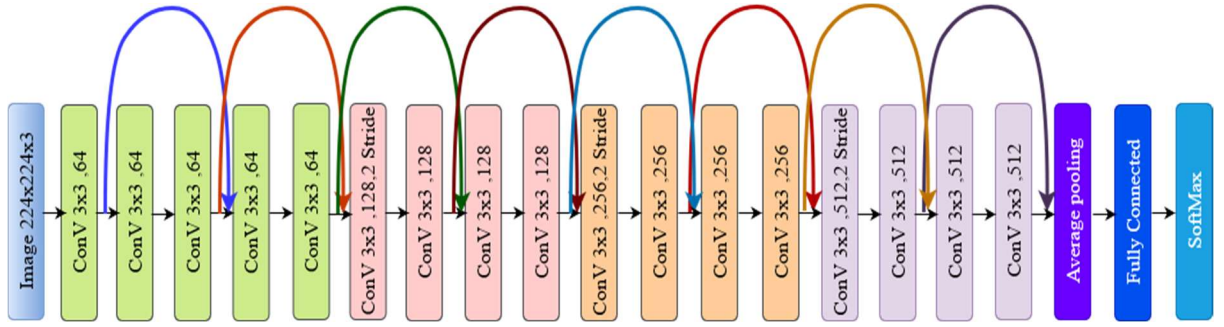
MobileNetV2, akıllı telefonlar ve diğer düşük güçlü cihazlar için geliştirilmiş 53 katmanlı bir DCNN mimarisidir. Bottleneck katmanları arasında ortaya çıkan artık bağlantılarla ters çevrilmiş bir artık yapı kullanır. Ara genişleme katmanı filtrelerinde derin hafif derinlemesine konvolüsyonlar kullanılır. MobileNetV2 tasarımı, 32 filtreli bir başlangıç tam konvolüsyonel katmanı ve ardından kalan 19 bottleneck katmanını içerir ve bu önceden eğitilmiş ağın fotoğrafları 1000 nesne grubuna kategorize etmesini sağlar. Ağ için görüntü giriş boyutu 224 x 224 pikseldir. MobileNetV2 mimarisi Şekil 3.19'da gösterilmiştir (Sandler ve ark., 2018).



Şekil 3.19 MobileNet V2 mimarisi

3.2.3.4. ResNet18

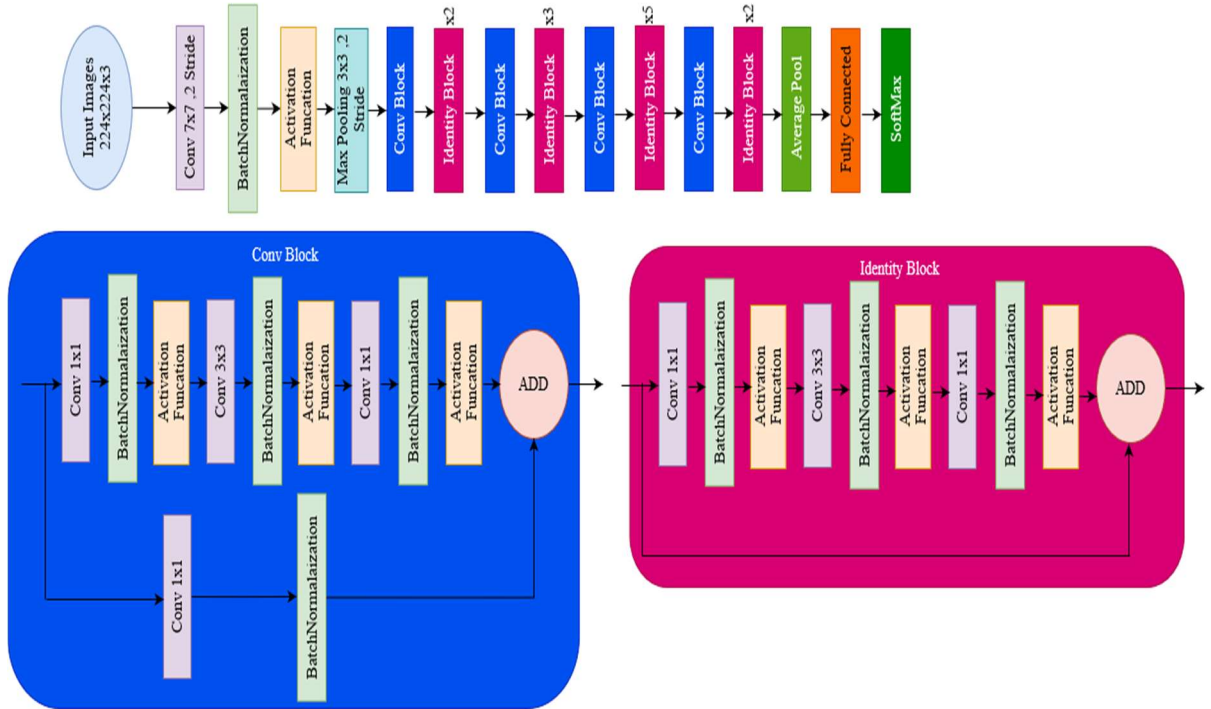
Derin artık ağların en temel modeli olan ResNet-18, 18 katman ve 224x224x3 giriş görüntü boyutu içerir. Bu önceden eğitilmiş ağ, bir milyondan fazla görüntü üzerinde eğitildikten sonra görüntüleri 1000 nesne kategorisine sınıflandırabilen bir CNN'dir. ILSVRC-2015 yarışmasının kazananları olan (He ve ark., 2016) tarafından tasarlanmıştır. ResNet18'de artık blok, bir katmanın çıktısı blokta daha derindeki başka bir katmana eklenecek şekilde yapılandırılmış bir katman yığındır. Doğrusal olmama ana yoldaki ilgili katmanın çıktısı ile birleştirilerek uygulanır. ResNet-18'in genel tasarımı Şekil 3.20'de gösterilmektedir.



Şekil 3.20 ResNet 18 mimarisi

3.2.3.5. ResNet50

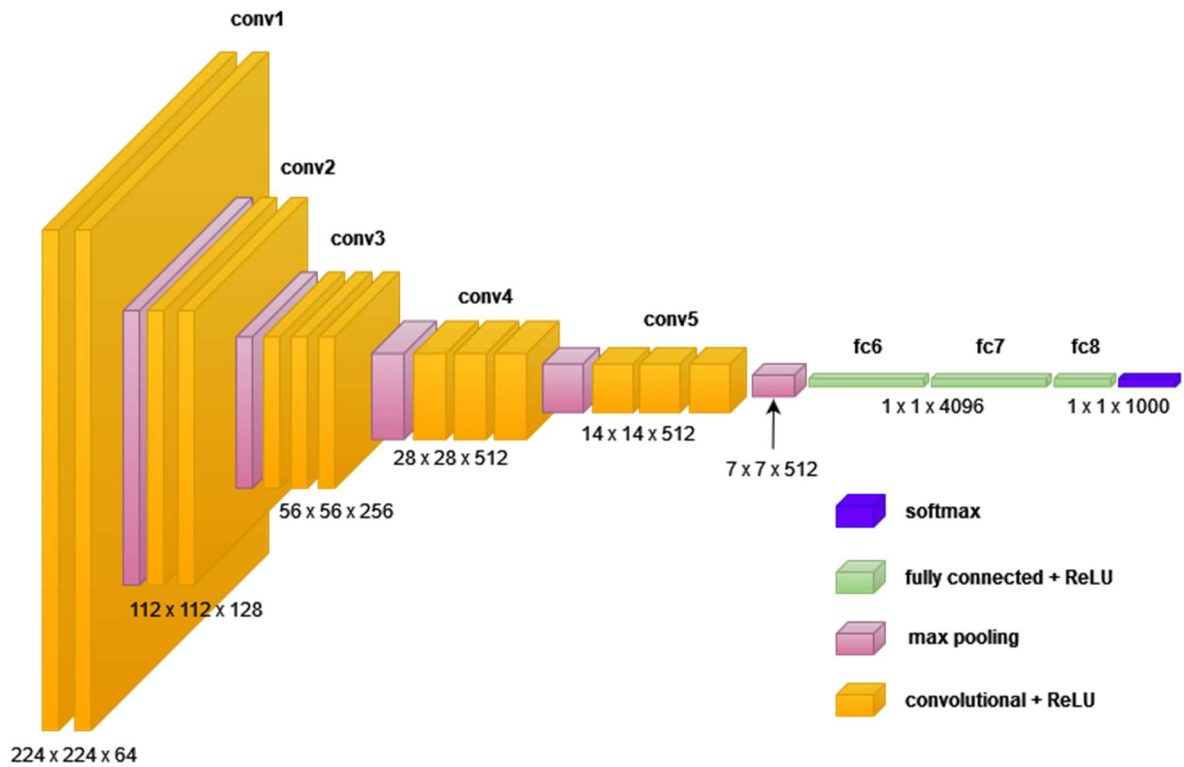
Bir ağ oluşturmak için artık blokları üst üste yığan bir tür derin artık ağı (ResNet) olan ResNet50, 48 Konvolüsyon katmanı, 1 MaxPool ve 1 Ortalama Havuz katmanından oluşur. Şekil 3.21'de iki atlama modülü görebiliriz; birincisi atlamalarda konvolüsyon katmanı olmayan bir kimlik bloğu, ikincisi ise atlamalarda konvolüsyon katmanı olan bir konvolüsyon bloğudur. Bu teknik darboğaz tasarımı olarak bilinir ve ağ performansını korurken parametre sayısını azaltır (He ve ark.,2016).



Şekil 3.21 ResNet 50 mimarisi

3.2.3.6. VGG16

Görsel Geometri Grubu (VGG) Derin Evrişimli Sinir Ağı, Simonyan ve Zisserman tarafından önerilmiştir. Modelleri ILSVRC-2014 yarışmasında ikinci olmuştur. Bu ağda 13 konvolüsyonel katman ve üç tam bağlantılı katman ile toplamda beş maksimum havuzlama katmanı bulunmaktadır. İlk blokta 64 filtre vardır ve bunlar 512 toplam filtre olana kadar ardışık bloklarda ikiye katlanır. Bu model, tam bağlantılı ve toplam 4096 nöron içeren iki gizli katman ve bir çıkış katmanı ile tamamlanmaktadır. Bu ağ, fotoğrafları 1000 nesne türüne göre sınıflandırabilir çünkü çıkış katmanında 1000 nöron vardır. VGG16'nın genel tasarımı Şekil 3.22'da gösterilmektedir (Simonyan ve Zisserman,2015).

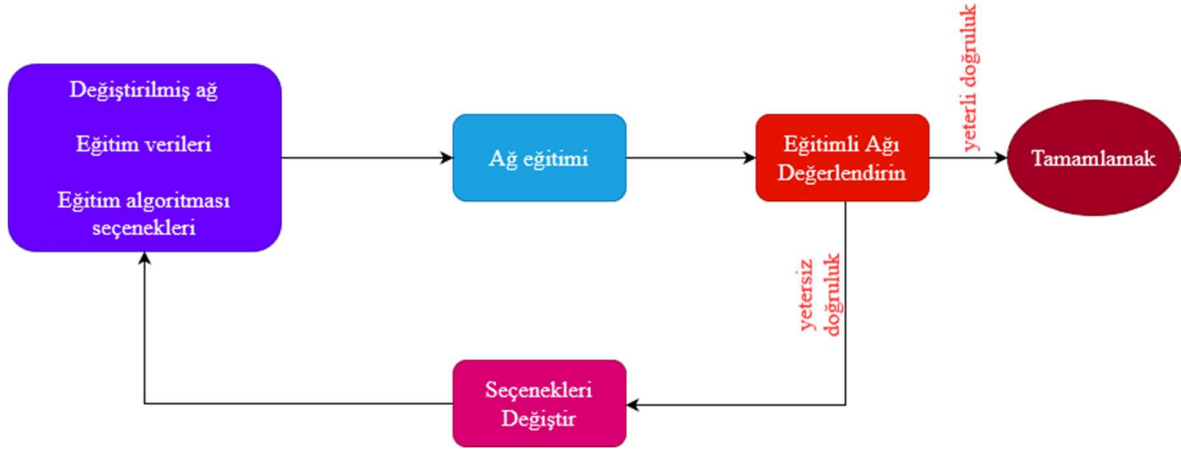


Şekil 3.22 VGG16 mimarisi

3.3. Transfer Learning Algoritması

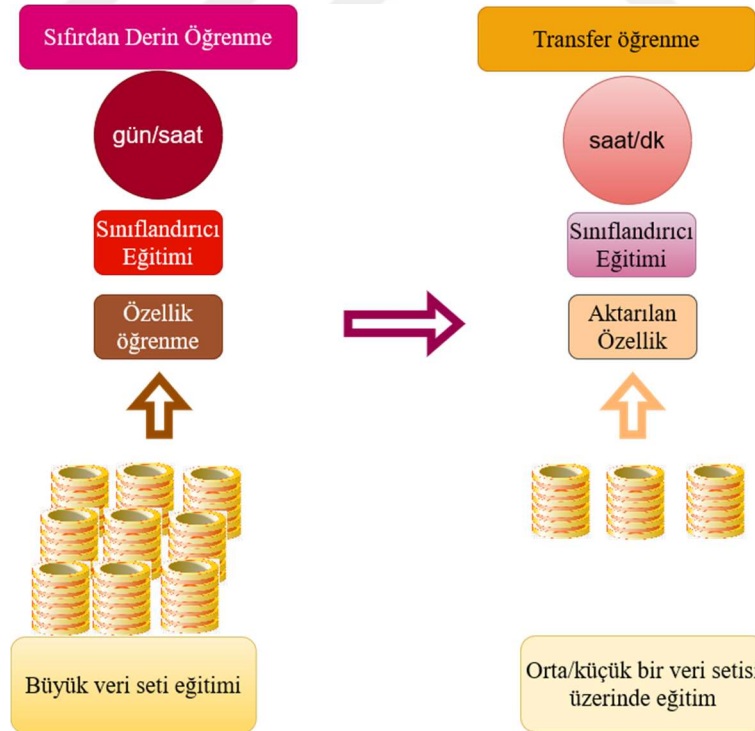
Mevcut çalışmada, Matlab 2021A'da görüntü işleme, derin öğrenme araç kutusu ve transfer öğrenme algoritması kullanılmakta, burada Evrişimsel Sinir Ağlarındaki son katmanlarda değişiklikler yapılmakta ve daha sonra bu ağlar yeni veriler üzerinde eğitilmektedir. Sınıflandırma doğruluğu yeterli ise istenen doğruluk elde edilir. Ancak, sınıflandırma

doğruluğu yetersizse, Evrişimsel Sinir Ağlarında değişiklikler yapılacak ve istenen yeterli doğruluk elde edilene kadar bu veriler üzerinde yeniden eğitim yapılacaktır (Şekil 3.23).



Şekil 3.23 Transfer learning iş akışı

Transfer öğrenme algoritması bu tür Derin Evrişimsel Sinir Ağlarında ve diğer Derin Evrişimsel Sinir Ağlarında kullanılabilir. Transfer öğrenme algoritmasının kullanılmasının nedeni, ağı eğitmek için küçük veya orta ölçekli verilerin kullanılabilmesi ve böylece eğitim süresinden tasarruf edilmesidir; sıfırdan eğitim ise ağı eğitmek için önemli miktarda veri gerektirir ve bu nedenle gerçekleştirilmesi uzun zaman alır (Bengio, 2012) (bkz. Şekil 3.24).



Şekil 3.24 Sıfırdan oluşturulan ağ (solda) ile transfer öğrenimi ile oluşturulan ağın (sağda) karşılaştırılması

3.4. Makine Öğreniminde Bazı Algoritmalar

3.4.1.SVM & MSVM sınıflandırıcıları

Destek Vektör Makinesi, sınıflandırma odaklı bir makine öğrenimi algoritmasıdır. Noktaları iki ayrı düzleme bölmek için denetimli öğrenmeyi kullanır. İki'den fazla hedef kategori olduğunda Çok Sınıflı DVM kullanılır. Buna karşın ECOC ve DVM kombinasyonu, sınıflandırma modelinin hata toleransını geliştirir.

3.4.2.ECOC-SVM sınıflandırıcı

ECOC-SVM (Hata Düzeltici Çıkış Kodları- Destek Vektör Makineleri), çok sınıflı sınıflandırma problemlerini ikili sınıflandırıcılarla çözen çok sınıflı bir sınıflandırma algoritmasıdır. Orijinal çok sınıflı problemi birden fazla ikili sınıflandırma problemine bölerek çalışır ve her bir ikili sınıflandırıcı bir çift sınıfı ayırt etmek için eğitilir. Bu ikili sınıflandırıcıların çıktıları daha sonra nihai çok sınıflı sınıflandırma sonucunu üretmek için bir hata düzeltme kodu kullanılarak birleştirilir.

ECOC-SVM'yi görüntü sınıflandırması için CNN ile bir sınıflandırıcı olarak kullanmak amacıyla aşağıdaki adımları takip edebilir:

- Giriş görüntüleri önceden işlenmeli ve eğitim, doğrulama ve test kümelerine ayrılmalıdır.
- Görüntülerden özellik temsillerini öğrenmek için eğitim seti üzerinde bir CNN modeli eğitilmelidir.
- Doğrulama ve test setlerindeki her görüntü için, CNN modelinin son tam bağlı katmanından özellik vektörleri çıkarılır.
- ECOC-SVM algoritmasını kullanarak eğitim setindeki özellik vektörleri üzerinde birden fazla ikili sınıflandırıcı eğitilir. Her bir ikili sınıflandırıcı iki sınıf arasında ayırım yapmak üzere eğitilmelidir.

Doğrulama ve test setlerinde, her görüntünün sınıfını tahmin etmek için eğitilmiş ikili sınıflandırıcıları kullanabilir. Hata düzeltme kodu, nihai çok sınıflı sınıflandırma sonucunu oluşturmak için ikili sınıflandırıcı çıktılarını birleştirmek için kullanılabilir. CNN'lerin güçlü özellik çıkarma yeteneklerini ECOC-SVM'nin verimli çok sınıflı sınıflandırma yetenekleriyle birleştirerek görüntü sınıflandırma görevlerinde yüksek doğruluk elde etmek mümkündür. Bununla birlikte, ECOC-SVM algoritmasının performansının hata düzeltme koduna ve ikili sınıflandırıcılara bağlı olduğunu belirtmek önemlidir. "Bire karşı hepsi" ve

"bire karşı biri" olmak üzere iki kodlama modeli vardır. "Bire karşı hepsi" ve "bire karşı biri" kodlama yöntemleri arasındaki karar, sınıf sayısı, veri kümesinin boyutu ve sınıflandırıcının karmaşıklığı gibi probleme özgü özelliklerden etkilenir.

Genel olarak, "bire karşı hepsi" kodlamasının uygulanması daha kolaydır çünkü k sınıf için yalnızca k ikili sınıflandırıcı gerektirir. Ayrıca, azınlık sınıflarına aşırı uyum riskini azalttığından, her sınıftaki örnek sayısında büyük bir eşitsizlik olduğunda daha uygundur.

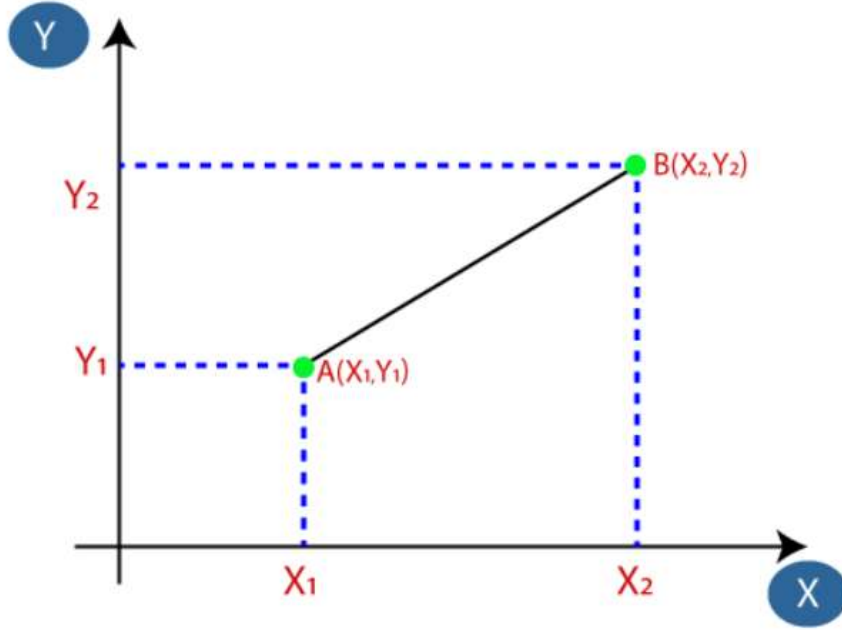
Öte yandan, k sınıf için "Bire-bir" kodlaması $k(k-1)/2$ ikili sınıflandırıcı gerektirir, bu da büyük k için hesaplama ve bellek açısından yoğun olabilir. Bununla birlikte, sınıfların yüksek oranda örtüştüğü durumlarda, her bir sınıf çiftini ayırt etmek için ikili sınıflandırıcıları eğittiği için "bire-bir" den daha doğru olabilir. "Bire karşı bir", yanlış pozitif ve yanlış negatif hataların sayısını azalttığı için orta sayıda sınıfa sahip problemlerde de "bire karşı hepsi"nden daha iyi performans gösterebilir. Uygulamada ECOC-SVM, veri kümesinin özelliklerine göre en uygun kodlama yöntemini otomatik olarak seçebilir (Lai ve ark.,2021).

3.4.3. K en yakın komşular (KNN) sınıflandırıcı

KNN, sınıflandırma ve regresyon analizi için kullanılan denetimli bir makine öğrenimi algoritmasıdır. Özellik uzayında yeni girdi veri noktasına en yakın k eğitim örneğini bularak ve ardından veri noktasını bu k en yakın komşunun çoğunluk sınıfına göre sınıflandırarak çalışır (Ab Wahab ve ark., 2021).

MATLAB'de KNN'ler, konvolüsyonel sinir ağları (CNN'ler) ile sınıflandırıcı olarak kullanılabilir. Temel fikir, CNN çıktısını bir özellik vektörü olarak kullanmak ve KNN algoritması ile sınıflandırmaktır. Bu yöntem, sınırlı miktarda etiketli veri mevcut olduğunda ve sınıflandırma doğruluğunu artırmak istediğimizde sıklıkla kullanılır. Sınıflandırma bağlamında k -En Yakın Komşu algoritması, temelde belirli bir 'görünmeyen' gözleme en çok benzeyen k örnek arasında bir çoğunluk oyu oluşturmak anlamına gelir. Benzerlik, iki veri noktası arasındaki bir mesafe metriği kullanılarak tanımlanır. Öklid mesafesi iyi bilinen bir seçenektir (bkz. Şekil 3.25). Düzlemdeki mesafeleri ölçerken sıklıkla Öklid mesafe formülünü kullanır. Öklid mesafe formülü, düzlemde (X, Y) ve (A, B) koordinatlarına sahip iki nokta arasındaki mesafeyi aşağıdaki denklemle (3.16) verir:

$$A \text{ ve } B \text{ arasındaki Öklid mesafesi} = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2} \quad (3.16)$$



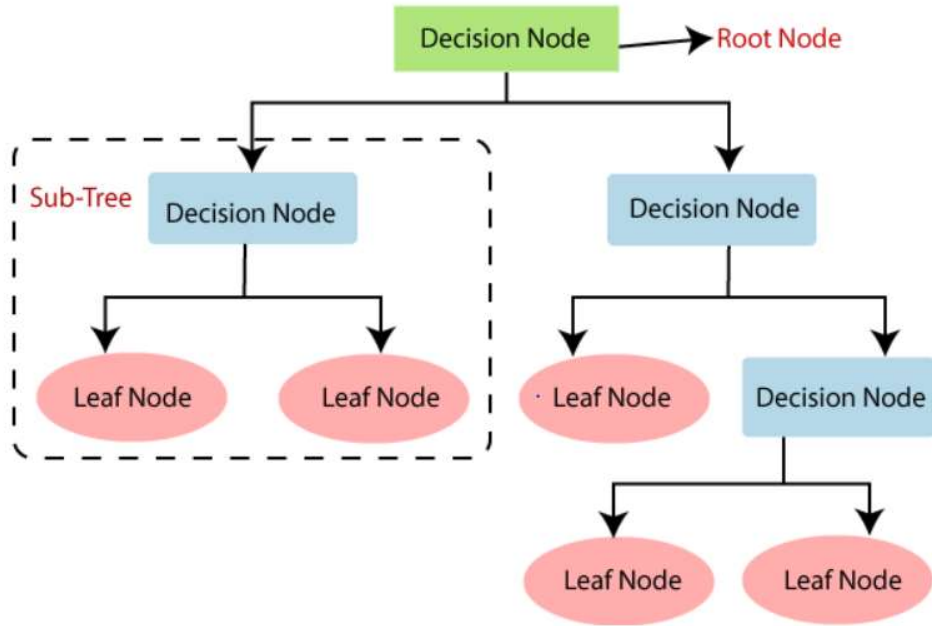
Şekil 3.25 Veri noktaları arasındaki Öklid mesafesi

k-NN, değiştirilecek yalnızca iki hiper parametresi olan çok basit bir algoritmadır: mesafe ölçeği ve 'k' değeri. Böylece, test setinde nihai modelimiz olarak kullanılacak olan en iyi validation doğruluğuna sahip modeli bulmak için modelimizi çeşitli 'k' değerleri için çalıştırabiliriz.

3.4.4. Karar Ağacı (DT) sınıflandırıcı

Karar ağacı, makine öğrenimi ve veri madenciliğinde yaygın olarak kullanılan bir denetimli öğrenme algoritmasıdır. Veri özelliklerinden basit karar kuralları çıkararak bir hedef değişkenin değerini tahmin etmek için bir model eğitmek üzere kullanılır. İç düğümler veri kümesi özelliklerini temsil ederken, dallar karar kurallarını ve her yaprak düğümü sonucunu temsil eder. Bir Karar ağacında iki düğüm vardır: Karar Düğümü ve Yaprak Düğümü. Karar düğümleri herhangi bir karar vermek için kullanılır ve birden fazla dala sahipken, Yaprak düğümleri bu kararların çıktısıdır ve başka dallara sahip değildir. Giriş özelliklerinin değerlerine bağlı olarak, veriler karar ağacı algoritmalarında özyinelemeli olarak bölünür ve her bir iç düğümün belirli bir özelliğe dayalı bir karar kuralını temsil ettiği ve her bir yaprak düğümünün hedef değişkenin tahmin edilen bir değerini temsil ettiği ağaç benzeri bir yapı ile sonuçlanır. Görüntü çok sınıflı sınıflandırma için bir sınıflandırıcı olarak kullanmak üzere, karar ağaçları evrimsel sinir ağları (CNN'ler) ile birleştirilebilir. Burada CNN'ler görüntülerden özellikler çıkarabilen ve bu özellikleri tahminler yapmak için kullanabilen

güçlü derin öğrenme algoritmalarıdır, oysa karar ağaçları da aynı şeyi yapabilir (bkz. Şekil 3.26) (Vikas Reddy ve Rama Parvathy, 2022).



Şekil 3.26 Basit karar ağacı modelinin mimarisi

3.5. Makine Öğreniminde Öğrenme Türleri

Öğrenme, tecrübenin beceri veya bilgiye dönüştürülmesidir. Öğrenme verilerinin doğasına ve öğrenen ile çevre arasındaki etkileşime dayanarak, bu bizi Makine Öğreniminin, sistemlerin büyük hacimli verilerden öğrenmesini ve belirli sorunları çözmesini sağlayan bir yapay zekâ uygulaması olduğu tanımına götürdü. Tecrübe yoluyla verimliliklerini otomatik olarak artıran bilgisayar algoritmaları kullanır (bkz. Şekil 3.27). Öğrenme, aşağıda listelendiği gibi genel olarak üç türe ayrılabilir.



Şekil 3.27 Makine öğreniminde öğrenme türleri ve makine öğrenimi algoritmaları (Kim ve Tagkopoulos, 2019)

3.5.1. Gözetimli öğrenme

Denetimli öğrenme, makine öğrenimi modellerinin etiketli veriler kullanılarak eğitildiği bir tür makine öğrenimidir. Denetimli öğrenmenin gerçek dünyadaki uygulamaları arasında yüz ve konuşma tanıma, ürün veya film önerileri ve satış tahminleri yer alır. Denetimli öğrenme iki türe ayrılır: regresyon ve sınıflandırma. Denetimli öğrenmedeki öğrenme verileri açıklamaları, etiketleri, hedefleri veya amaçlanan çıktıları içerir ve amaç, girdileri çıktılarıyla eşleştiren genel bir kural belirlemektir. Bu tür öğrenme verileri etiketli veriler olarak bilinir. Öğrenilen kural daha sonra bilinmeyen çıktıları etiketlemek için yeni verilere uygulanır. Denetimli öğrenme, etiketli örneklerle dayalı bir makine öğrenimi modeli geliştirmeyi gerektirir. Denetimli öğrenme, mevcut eğitim verilerinden bir fonksiyonun

öğrenilmesiyle ilgilidir. Bir öğrenme algoritması eğitim verilerini inceler ve yeni örnekleri eşlemek için türetilmiş bir işlev oluşturur. Örneğin, bir arsanın veya evin fiyatını boyut, konum gibi çeşitli faktörlere göre tahmin eden bir sistem tasarlamak istiyorsak, önce bir veritabanı oluşturmalı ve etiketlemeliyiz. Algoritmaya hangi özelliklerin bu fiyatlandırmaya karşılık geldiğini söylemeliyiz. Algoritma, bu verilere dayalı olarak girdi özelliklerinin değerlerini kullanarak gayrimenkulün fiyatını nasıl hesaplayacağını öğrenecektir. Benzer şekilde, denetimli öğrenmenin yaygın örnekleri arasında e-postaları spam veya spam değil olarak kategorize etme, web sayfalarını içeriklerine göre etiketleme ve ses tanıma yer alır. Denetimli öğrenme iki türe ayrılır: regresyon ve sınıflandırma (Fumo, 2017).

3.5.1.1. Regresyon

Regresyon: gayrimenkul fiyatları gibi sürekli değerli bir yanıtı eğitir ve tahmin eder. Regresyondaki hedef değişken sürekli bir değerdir. Regresyon, girdi değişkenlerine dayalı olarak hedef değişkenin değerini tahmin etmeye çalışır. Regresyon yöntemleri arasında doğrusal regresyon, polinom regresyon ve karar ağaçları bulunur.

3.5.1.2. Sınıflandırma

Sınıflandırmada hedef değişken bir kategori değeridir. Sınıflandırmanın amacı, girdi değişkenlerine dayalı olarak hedef değişkenin sınıfını veya kategorisini tahmin etmektir. Sınıflandırma teknikleri arasında lojistik regresyon, karar ağaçları, destek vektör makineleri ve sinir ağları yer alır. Sınıflandırma, uygun sınıf etiketini belirlemek için olumlu ve olumsuz duyular, erkek ve kadınları, iyi huylu ve kötü huylu kanserleri, güvenli ve güvensiz kredileri vb. analiz eder.

3.5.2. Denetimsiz Öğrenme

Denetimsiz öğrenme, etiketsiz verileri kullanarak makineleri eğiten bir tür makine öğrenmesidir. Etiketsiz veriler belirli bir çıktı değişkeninden yoksundur. Model verilerden öğrenir, verilerdeki örüntüleri ve özellikleri tanımlar ve sonuçları çıktı olarak verir. Bu durumda makinenin amacı, önceden herhangi bir veri eğitimi olmaksızın benzerliklere, örüntülere ve farklılıklara dayalı olarak sınıflandırılmamış verileri kategorize etmektir.

Denetimli öğrenmeden farklı olarak, herhangi bir öğretmen yoktur, bu da bilgisayarın eğitilmeyeceği anlamına gelir. Sonuç olarak makine, etiketlenmemiş verilerdeki gizli yapıyı kendi başına keşfetmekle sınırlıdır. Denetimsiz öğrenme algoritmaları, örüntüleri ve eğilimleri tespit etmek için son derece güçlü veri analizi araçlarıdır. En tipik olarak ilgili

girdileri mantıksal olarak gruplandırmak için kullanılırlar. Kmeans, Random Forests, Hiyerarşik kümeleme ve diğer denetimsiz öğrenme algoritmaları bunlara örnektir. Denetimsiz öğrenme algoritmaları iki türe ayrılır (Candan, 2022):

3.5.2.1. Kümeleme

Bir kümeleme problemi, müşterileri satın alma faaliyetlerine göre gruplandırmak gibi verilerdeki doğal gruplamaları ortaya çıkarmak istenilen bir problemdir. K-means kümeleme, DBSCAN kümeleme ve Agglomerative Hierarchy kümeleme bu tür tekniklere örnektir.

3.5.2.2. Birliktelik

Bir birliktelik kuralı öğrenme problemi, X satın alan kişiler aynı zamanda Y satın alma eğilimindedir gibi verilerinizin önemli parçalarını tanımlayan kuralları bulmak istediğiniz bir problemdir. Apriori, Eclat (Eşdeğerlik Sınıfı Dönüşümü) ve F-P (Sık Örüntü) Büyümesi bu tür tekniklere örnektir.

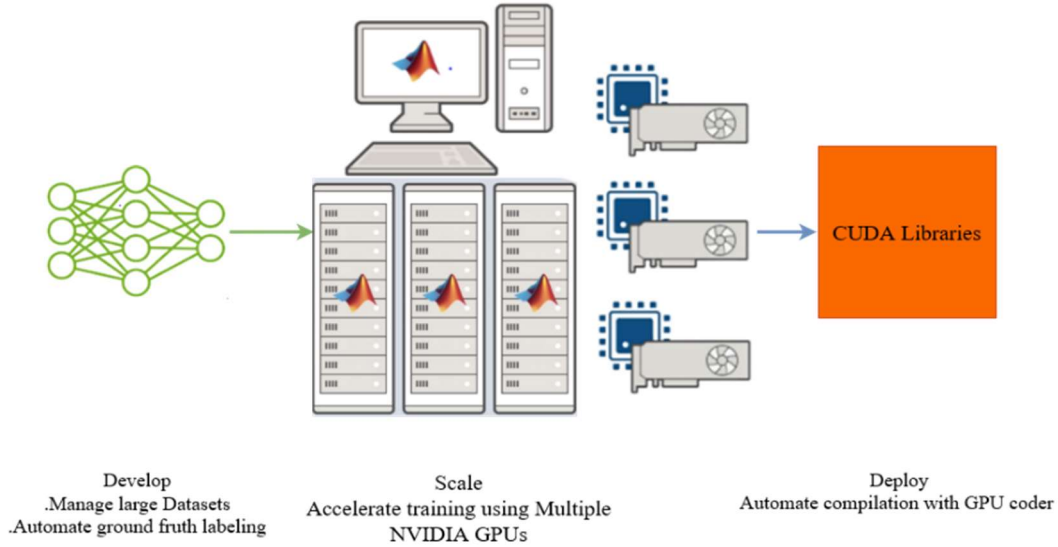
3.5.3. Reinforcement learning

Reinforcement learning, bir ajanın eylemler gerçekleştirerek ve bu eylemlerin sonuçlarını görerek belirli bir ortamda nasıl davranacağını öğrendiği geri bildirim tabanlı bir Makine Öğrenimi tekniğidir. Her olumlu faaliyet için ajan olumlu geri bildirim alır; her olumsuz eylem için ajan olumsuz geri bildirim veya ceza alır. İstenen sonucu elde etmek için, Reinforcement learning deneme yanılma prosedürlerini kullanır. Öğrenme verileri burada geri bildirim sağlayarak sistemin belirli bir hedefe ulaşmak için dinamik durumlara yanıt vermesine olanak tanır. Geri bildirim yanıtlarına dayanarak, sistem performansını değerlendirir ve buna göre tepki verir. Temsilci bir görevi tamamladıktan sonra ödüllendirilir. Modelleri eğitmek için, Reinforcement learning yaklaşımları herhangi bir dış izleme gerektirmez. Reinforcement learning algoritmaları genellikle oyun endüstrisinde oyunlar oluşturmak için kullanılır. Ayrıca robotlara insan benzeri görevleri yerine getirmeyi öğretmek için de kullanılır. Q-Learning, State Action Reward State action (SARSA) ve Deep Q Neural Network (DQN) bu tür tekniklere örnektir (Bhatt, 2019).

3.6. Kullanılan Bilgisayar Sistemi

Bu çalışmada i7/7th CPU, 16GB RAM ve 4GB GTX1050'ye sahip bir bilgisayar kullanılmıştır. GPU/CUDA özellikleri sayesinde eğitim ve test aşamaları boyunca işlem

hızlandırma için kullanılabilir. NVIDIA CUDA® Deep Neural Network (cuDNN) kütüphanesi, derin sinir ağı ilkellerinin GPU ile hızlandırılmış bir kütüphanesi olmakla birlikte, ön plan, arka plan, konvolüsyon, havuzlama, normalizasyon ve aktivasyon katmanları gibi standart işlemlerin oldukça ince ayarlı uygulamalarını da içerir (Şekil 3.28).

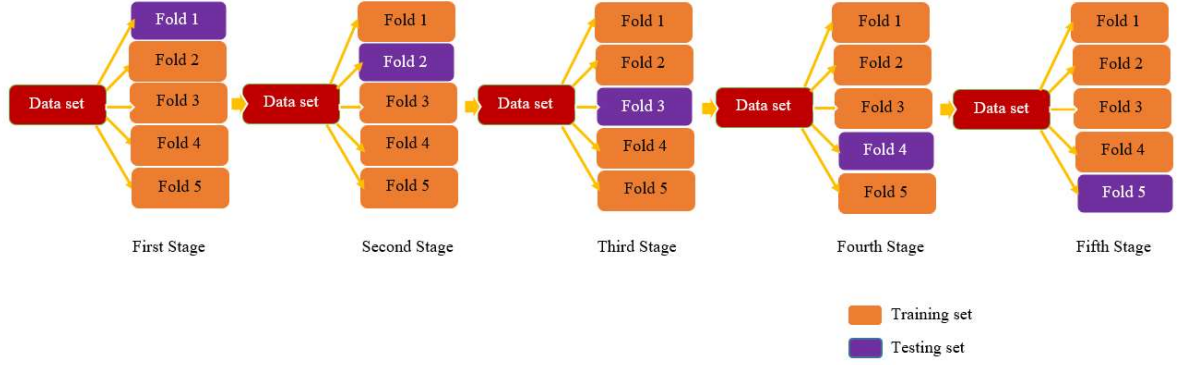


Şekil 3.28 GPU yolu

3.7. K-Kat Çapraz Doğrulama

Sınıflandırma problemlerinde veri setimiz önce eğitim ve test setlerine ayrılır ve daha sonra eğitim veri seti üzerinde bir model oluşturularak yapılan tahminler test seti üzerinde sınanır. Ancak, eğitim/test ayrımı ile ilgili bazı sorunlar olabilir. Örneğin, veri setini rastgele ayırmak mümkün olmayabilir veya model oluşturma için yalnızca bazı veri setleri seçilmiş olabilir, bu da aşırı uyum sorununa neden olur. Bu sorun çapraz doğrulama ile çözülebilir (Hussain ve ark., 2021). Çapraz doğrulama, modelin yüksek performansının rastgele olup olmadığını gösterecektir. K-Kat Çapraz Doğrulamada, veriler alt gruplara ayrılır ve k-1 alt kümesi verileri eğitmek için kullanılır ve son alt küme test verisi olarak bırakılır. K deney sonucunda elde edilen hatanın ortalama değeri modelimizin doğruluğunu gösterir (Yadav ve Shukla, 2016). Veri kümesi K sayıda bölüme/katlara ayrılır ve her kat nokta sayısı kadar test kümesi olarak kullanılır. Mevcut çalışmada beş kat doğrulama senaryosu kullanılmıştır (K=5). Burada veri seti beş kata bölünmüştür. İlk aşamada, ilk kat modeli test etmek için, geri kalanlar ise modeli eğitmek için kullanılır. İkinci aşamada, ikinci kat test kümesi olarak

kullanılırken diğerleri eğitim kümeleri olarak kullanılır. Bu işlem, beş katın her biri test için veri seti olarak kullanılabileceği kadar tekrarlanır (Şekil 3.29).



Şekil 3.29 K-kat çapraz doğrulama

4. BULGULAR VE TARTIŞMA

Bu çalışmada, sekiz yüz ifadesini tanımak için iki yaklaşım kullanıldı.

İlk yaklaşım, önceden eğitilmiş evrişimli sinir ağlarıyla transfer öğrenme ve k-kat çarpaz doğrulama yöntemlerini kullanıyor.

İkinci yaklaşım ise, ECOC-MSVM, KNN ve DT gibi makine öğrenimi algoritmalarını sınıflandırıcı olarak kullanan önceden eğitilmiş evrişimli sinir ağlarıyla transfer öğrenme üzerine kuruludur.

4.1. Transfer Öğrenme ve K-Kat Çarpaz Doğrulama Kullanılarak Elde Edilen Sonuçlar

Bu bölümde, 8 yüz ifadesini temsil etmek için 70 kişiden 6720 görüntü kullandık, veri setimiz ise 8 yüz ifadesini temsil etmek için 83 öğrenciden 7968 görüntü içermektedir. Bu bölümde, AlexNet, GoogleNet, ResNet18, ResNet50, MobileNetV2 ve VGG16 gibi CNN'ler (Evrişimli Sinir Ağları) kullanarak transfer öğrenme (TL) ve k-kat çarpaz doğrulama (KCV) yöntemlerini kullandık. Aşağıdaki Şekil 4.1'de görülen akış şeması, bu bölümde kullanılan algoritmayı göstermektedir. Bu algoritma aşağıdaki adımlarla açıklanabilir:

Birinci adım: 70 kişinin yüz ifadelerini temsil eden 6720 görüntüden oluşan veri setimizi yükleme.

İkinci adım: Başlangıç sayacı değerini $K=1$ olarak ayarlama.

Üçüncü adım: Veri setini beş parçaya bölmek için beş kat (5-fold) eğitim ve test verileri oluşturma (veri setinin %80' i eğitim için, %20'si test için).

Dördüncü adım: Önceden eğitilmiş CNN ağını yükleme.

Beşinci adım: Verilerin Derin Evrişimli Sinir Ağına (Deep Convolutional Neural Network) uygun hale gelmesi için ön işleme aşaması (görüntülerin yeniden boyutlandırılması).

Altıncı adım: Bu süreçte CNN ağındaki son katmanları, transfer öğrenme kullanarak yeni katmanlarla değiştirme.

Yedinci adım: CNN ağını eğitim veri seti üzerinde eğitme.

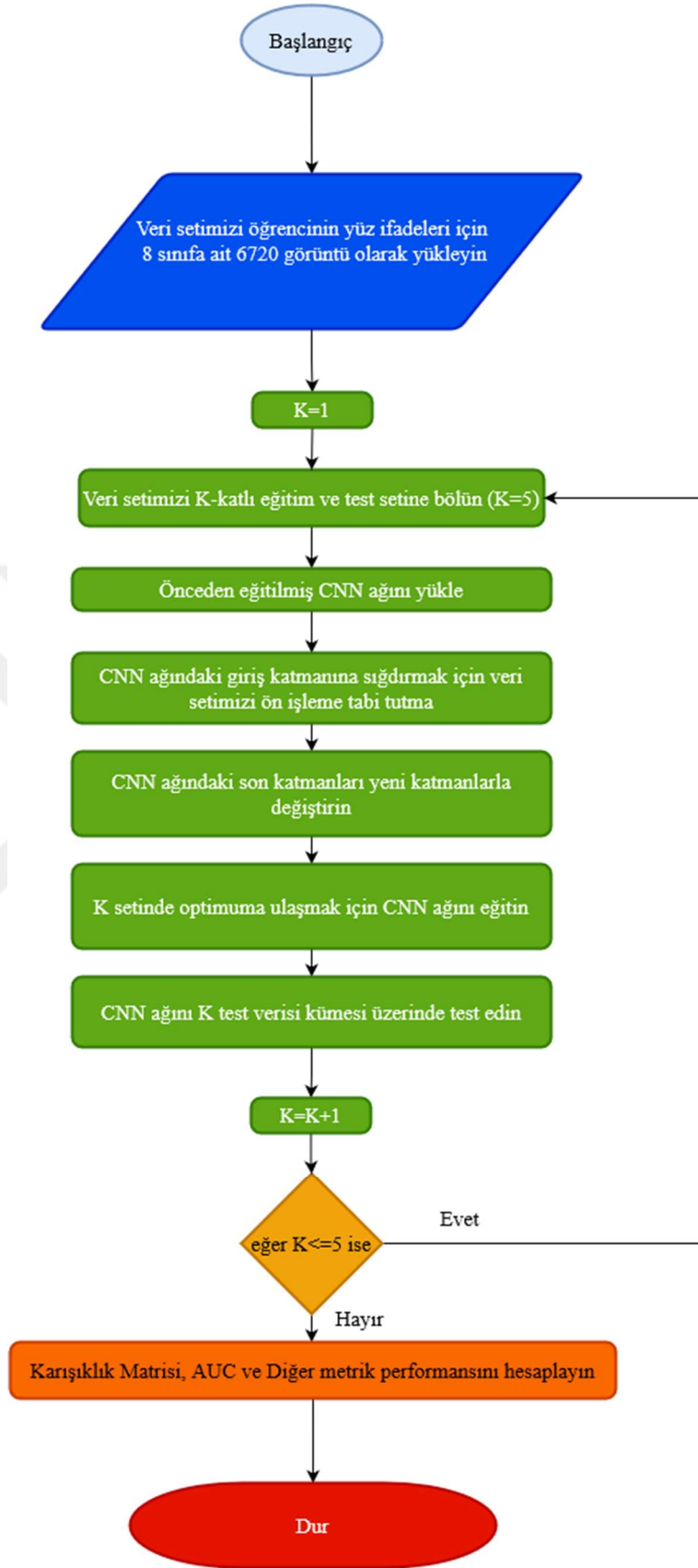
Sekizinci adım: CNN ağını test veri seti üzerinde test etme.

Dokuzuncu adım: Sayacın deęerini bir artırma.

Onuncu adım: Yeni sayacın deęerini 5 ile karşılařtırma, eęer yeni deęer 5'e eřit veya daha kúçükse 3 ile 8. adımları tekrarla.

Aksi takdirde (sayacın deęeri 5'ten büyükse), döngüden çık ve performans metriklerini hesapla (bu, beř kez tekrarlanan sonuçların ortalama oranıdır. Konfüzyon matrisleri ve ROC-AUC eęrileri de çıkarıldı).





Şekil 4.1 Bu bölümde kullanılan algoritmanın akış şeması

Bu bölümde, 6720 görüntü altı önceden eğitilmiş CNN ağı, transfer öğrenme ile K-kat çarpaz doğrulama ile uygulanmıştır. Bu ağlar AlexNet + TL + KFCV, GoogleNet + TL + KFCV, ResNet 18 + TL + KFCV, ResNet50 + TL + KFCV, MobileNetV2 + TL + KFCV ve VGG16 + TL + KFCV olarak adlandırılmaktadır. Transfer öğrenme (TL) + 5-kat çarpaz doğrulama süreci, önceki altı eğitilmiş derin öğrenme mimarisi (AlexNet, GoogleNet, ResNet 18, ResNet 50, MobileNetV2 ve VGG16) üzerinde uygulanmıştır. Eğitim görüntülerinden çıkarılan özellikler, bu mimarilerin daha iyi parametrelenmesi için modelleri eğitmek için kullanılmıştır. Transfer öğrenme algoritması kullanılarak, altı ağ veri setimiz üzerinde eğitilir ve böylece yeni parametrelere sahip olurken, eski parametreler kaldırılır. Bu bölüm için veri seti üniversite öğrencilerinin yüz görüntülerinden oluşmaktadır ve her bir kategori için 840 görüntü olmak üzere sekiz ifade kategorisine (öfke, iğrenme, korku, mutluluk, tarafsızlık, üzüntü, dikkat ve şaşkınlık) dağılmıştır. Pratik deneyler altı ağ üzerinde ayrı ayrı gerçekleştirilmiştir. Çizelge 4.1, her bir ağ için eğitim ve test sürelerini göstermektedir. Eğitim ve test süresinin her bir ağ tarafından diğerlerinden farklı maksimum parametre sayısı ile ilişkili olduğu gözlemlenebilir (Çizelge 4.2).

Çizelge 4.1 GPU kullanılarak çeşitli CNN ağları için eğitim ve test aşamalarında geçen süre

CNN Ağı+TL+KFCV (K=5)	Eğitim ve test için geçen süre (Dakika)
AlexNet+TL+KFCV	≈147
GoogleNet+TL+KFCV	≈110
ResNet18+TL+KFCV	≈105
ResNet50+TL+KFCV	≈300
MobileNetV2+TL+KFCV	≈250
VGG16+TL+KFCV	≈660

Çizelge 4.2 Çeşitli CNN ağları için maksimum parametre sayısı

CNN Ağı	Maksimum veri büyüklüğü	Parametreler (Milyon)	Giriş görüntüsü boyutu	Katman derinliği *
AlexNet (Krizhevsky ve ark., 2017)	227-by-227	61.0	227 MB	8
GoogleNet (Guo ve ark., 2017)	224-by-224	7.0	27 MB	22
ResNet18 (He ve ark., 2016)	224-by-224	11.7	44 MB	18
ResNet50 (He ve ark., 2016)	224-by-224	25.6	96 MB	50
MobileNetV2 (Sandler ve ark., 2018)	224-by-224	3.5	13 MB	53
VGG16 (Simonyan ve Zisserman, 2015)	224-by-224	138	515 MB	16

* Giriş katmanından çıkış katmanına kadar olan yol üzerinde, ağ derinliği ardışık evrişimli veya tam bağlı katmanların en büyük sayısı olarak tanımlanır.

4.1.1. İlk kısım için performans ölçümleri

Modelin performansını değerlendirmede birçok değerin yardımcı olduğu, çünkü bu değerlerin, bir sınıflandırma modelinin (yani sınıflandırıcı) performansının değerlendirilmesinin temelini oluşturan konfüzyon matrisinden çıkarıldığı anlaşılabilir. Bu kare matris, bir sınıflandırıcı tarafından yapılan Gerçek pozitif (TP), Gerçek negatif (TN), Yanlış pozitif (FP) ve Yanlış negatif (FN) tahminlerin sayılarını bildirerek öğrenme algoritmasının performansını gösterir. Bu çalışmada, modellerin (AlexNet + TL + KFCV, GoogleNet + TL + KFCV, ResNet 18 + TL + KFCV, ResNet50 + TL + KFCV, MobileNetV2 + TL + KFCV ve VGG16 + TL + KFCV) eğitildiği konfüzyon matrisleri gösterilmiştir ((Şekil 4.2 – Şekil 4.7).

Confusion Matrix: alexnet									
Output Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	
	822 12.2%	3 0.0%	5 0.1%	0 0.0%	1 0.0%	2 0.0%	2 0.0%	1 0.0%	98.3% 1.7%
	7 0.1%	797 11.9%	1 0.0%	8 0.1%	4 0.1%	25 0.4%	8 0.1%	10 0.1%	92.7% 7.3%
	6 0.1%	0 0.0%	824 12.3%	0 0.0%	3 0.0%	0 0.0%	9 0.1%	0 0.0%	97.9% 2.1%
	2 0.0%	3 0.0%	0 0.0%	824 12.3%	2 0.0%	2 0.0%	2 0.0%	1 0.0%	98.6% 1.4%
	0 0.0%	0 0.0%	1 0.0%	1 0.0%	812 12.1%	0 0.0%	0 0.0%	0 0.0%	99.8% 0.2%
	2 0.0%	26 0.4%	5 0.1%	3 0.0%	8 0.1%	800 11.9%	17 0.3%	4 0.1%	92.5% 7.5%
	0 0.0%	7 0.1%	4 0.1%	1 0.0%	6 0.1%	10 0.1%	802 11.9%	2 0.0%	96.4% 3.6%
	1 0.0%	4 0.1%	0 0.0%	3 0.0%	4 0.1%	1 0.0%	0 0.0%	822 12.2%	98.4% 1.6%
									97.9% 2.1%
									94.9% 5.1%
									98.1% 1.9%
									98.1% 1.9%
									96.7% 3.3%
									95.2% 4.8%
									95.5% 4.5%
									97.9% 2.1%
									96.8% 3.2%
Target Class									

Şekil 4.2 AlexNet+TL+KFCV için konfüzyon matrisi

Confusion Matrix: GoogleNet									
Output Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	
	741 11.0%	38 0.6%	30 0.4%	17 0.3%	11 0.2%	15 0.2%	32 0.5%	7 0.1%	83.2% 16.8%
	42 0.6%	624 9.3%	9 0.1%	28 0.4%	6 0.1%	113 1.7%	75 1.1%	32 0.5%	67.2% 32.8%
	25 0.4%	10 0.1%	752 11.2%	8 0.1%	41 0.6%	6 0.1%	39 0.6%	0 0.0%	85.4% 14.6%
	10 0.1%	18 0.3%	1 0.0%	674 10.0%	6 0.1%	11 0.2%	6 0.1%	44 0.7%	87.5% 12.5%
	3 0.0%	1 0.0%	10 0.1%	12 0.2%	746 11.1%	19 0.3%	13 0.2%	9 0.1%	91.8% 8.2%
	7 0.1%	68 1.0%	10 0.1%	18 0.3%	17 0.3%	601 8.9%	41 0.6%	22 0.3%	76.7% 23.3%
	9 0.1%	67 1.0%	28 0.4%	13 0.2%	8 0.1%	66 1.0%	634 9.4%	6 0.1%	76.3% 23.7%
	3 0.0%	14 0.2%	0 0.0%	70 1.0%	5 0.1%	9 0.1%	0 0.0%	720 10.7%	87.7% 12.3%
									88.2% 11.8%
									74.3% 25.7%
									89.5% 10.5%
									80.2% 19.8%
									88.8% 11.2%
									71.5% 28.5%
									75.5% 24.5%
									85.7% 14.3%
									81.7% 18.3%
Target Class									

Şekil4.3 GoogleNet+TL+KFCV için konfüzyon matrisi

Confusion Matrix: ResNet18									
Output Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	
	837 12.5%	5 0.1%	1 0.0%	1 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.2% 0.8%
	0 0.0%	811 12.1%	0 0.0%	9 0.1%	0 0.0%	4 0.1%	0 0.0%	0 0.0%	98.4% 1.6%
	3 0.0%	1 0.0%	836 12.4%	0 0.0%	0 0.0%	0 0.0%	7 0.1%	0 0.0%	98.7% 1.3%
	0 0.0%	2 0.0%	0 0.0%	830 12.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.8% 0.2%
	0 0.0%	0 0.0%	0 0.0%	0 0.0%	839 12.5%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	0 0.0%	20 0.3%	0 0.0%	0 0.0%	0 0.0%	832 12.4%	6 0.1%	0 0.0%	97.0% 3.0%
	0 0.0%	1 0.0%	3 0.0%	0 0.0%	0 0.0%	4 0.1%	827 12.3%	0 0.0%	99.0% 1.0%
	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.0%	0 0.0%	0 0.0%	840 12.5%	99.9% 0.1%
	99.6% 0.4%	96.5% 3.5%	99.5% 0.5%	98.8% 1.2%	99.9% 0.1%	99.0% 1.0%	98.5% 1.5%	100% 0.0%	99.0% 1.0%
Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise		
Target Class									

Confusion Matrix: Mobilenetv2									
Output Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	
	829 12.3%	12 0.2%	4 0.1%	5 0.1%	0 0.0%	1 0.0%	0 0.0%	1 0.0%	97.3% 2.7%
	2 0.0%	767 11.4%	0 0.0%	6 0.1%	1 0.0%	12 0.2%	3 0.0%	0 0.0%	97.0% 3.0%
	5 0.1%	1 0.0%	823 12.2%	0 0.0%	2 0.0%	0 0.0%	9 0.1%	0 0.0%	98.0% 2.0%
	1 0.0%	4 0.1%	0 0.0%	825 12.3%	1 0.0%	0 0.0%	1 0.0%	2 0.0%	98.9% 1.1%
	0 0.0%	1 0.0%	0 0.0%	0 0.0%	834 12.4%	1 0.0%	0 0.0%	2 0.0%	99.5% 0.5%
	2 0.0%	31 0.5%	0 0.0%	2 0.0%	1 0.0%	802 11.9%	8 0.1%	2 0.0%	94.6% 5.4%
	1 0.0%	20 0.3%	13 0.2%	0 0.0%	0 0.0%	23 0.3%	819 12.2%	0 0.0%	93.5% 6.5%
	0 0.0%	4 0.1%	0 0.0%	2 0.0%	1 0.0%	1 0.0%	0 0.0%	833 12.4%	99.0% 1.0%
									98.7% 1.3%
									91.3% 8.7%
									98.0% 2.0%
									98.2% 1.8%
									99.3% 0.7%
									95.5% 4.5%
									97.5% 2.5%
									99.2% 0.8%
									97.2% 2.8%
Target Class									

Şekil 4.6 MobileNet +TL+KFCV için konfüzyon matrisi

Confusion Matrix: vgg16									
Output Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	
	827 12.3%	18 0.3%	4 0.1%	1 0.0%	0 0.0%	3 0.0%	8 0.1%	0 0.0%	96.1% 3.9%
	3 0.0%	742 11.0%	2 0.0%	4 0.1%	0 0.0%	30 0.4%	14 0.2%	6 0.1%	92.6% 7.4%
	7 0.1%	1 0.0%	827 12.3%	0 0.0%	0 0.0%	0 0.0%	18 0.3%	0 0.0%	97.0% 3.0%
	2 0.0%	6 0.1%	0 0.0%	771 11.5%	0 0.0%	2 0.0%	1 0.0%	10 0.1%	97.3% 2.7%
	0 0.0%	0 0.0%	2 0.0%	4 0.1%	833 12.4%	1 0.0%	1 0.0%	0 0.0%	99.0% 1.0%
	1 0.0%	48 0.7%	2 0.0%	7 0.1%	4 0.1%	792 11.8%	15 0.2%	2 0.0%	90.9% 9.1%
	0 0.0%	13 0.2%	3 0.0%	0 0.0%	3 0.0%	10 0.1%	781 11.6%	1 0.0%	96.3% 3.7%
	0 0.0%	12 0.2%	0 0.0%	53 0.8%	0 0.0%	2 0.0%	2 0.0%	821 12.2%	92.2% 7.8%
									98.5% 1.5%
									88.3% 11.7%
									98.5% 1.5%
									91.8% 8.2%
									99.2% 0.8%
									94.3% 5.7%
									93.0% 7.0%
									97.7% 2.3%
									95.1% 4.9%
Target Class									

Şekil 4.7 VGG 16 +TL+KFCV için konfüzyon matrisi

Konfüzyon matrisinden aşağıdaki performans metriklerinin birkaç parametresi çıkarılabilir (Abdullah ve Alkan,2022):

1. Doğruluk (ACC), bir sınıflandırıcının ne sıklıkla doğru tahmin yaptığını ölçer ve aşağıdaki denklem kullanılarak hesaplanır:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.1)$$

2. Duyarlılık (REC), aynı zamanda duyarlılık (SNS) veya doğru pozitif oranı (TPR) olarak da adlandırılır. Bir hedef için gözlenen tüm pozitif değerlerden gerçek pozitiflerin oranını tahmin eder:

$$SNS = \frac{TP}{TP + FN} \quad (4.2)$$

3. Kesinlik (PRC), kesin olarak tanımlanmış gerçek pozitiflerin oranını tahmin eder ve aşağıdaki gibi hesaplanır:

$$PRC = \frac{TP}{TP + FP} \quad (4.3)$$

4. Özgüllük (SPC), bir hedef için gözlenen tüm negatif değerlerden gerçek negatiflerin oranını tahmin eder, aynı zamanda doğru negatif oranı (TNR) olarak da bilinir.

$$SPC = \frac{TN}{TN + FP} \quad (4.4)$$

5. F1 Skoru, Hassasiyet ve Hatırlama değerlerinin harmonik ortalamasını gösterir. F1 Skoru, dengesiz bir sınıflandırıcıyı tahmin etmek için de uygun bir ölçüdür.

$$F1 - Skoru = \frac{2 \times PRC \times SNS}{PRC + SNS} \quad (4.5)$$

Şimdi, konfüzyon matrislerine dayanarak bu araştırmada kullanılan altı ağ için (AlexNet + TL + KFCV, GoogleNet + TL + KFCV, ResNet 18 + TL + KFCV, ResNet50 + TL + KFCV, MobileNetV2 + TL + KFCV ve VGG16 + TL + KFCV) performans metriklerini (ACC, SNS, PRC, SPC, F1-Skor) çıkarmak mümkündür. Performans metriklerinin değerlerini gösteren Çizelgeler (Çizelgeler 4.3 - 4.8) aşağıdaki gibi düzenlenmiştir.

Çizelge 4.3 AlexNet için performans ölçümlerinin değerleri

AlexNet	ACC	SNS	PRC	SPC	F1_score
Öfke	0.995	0.983	0.978	0.996	0.980
Dikkat	0.984	0.926	0.948	0.992	0.937
İğrenme	0.994	0.978	0.980	0.997	0.979
Korku	0.995	0.985	0.980	0.997	0.983
Mutluluk	0.995	0.997	0.966	0.995	0.981
Tarafsızlık	0.984	0.924	0.952	0.993	0.938
Üzüntü	0.989	0.963	0.954	0.993	0.959
Sürpriz	0.995	0.984	0.978	0.996	0.981
Ortalama	0.991	0.968	0.967	0.995	0.967

Çizelge 4.4 GoogleNet için performans ölçümlerinin değerleri

GoogleNet	ACC	SNS	PRC	SPC	F1_score
Öfke	0.962	0.831	0.882	0.983	0.856
Dikkat	0.922	0.671	0.742	0.962	0.705
İğrenme	0.967	0.853	0.895	0.984	0.873
Korku	0.961	0.875	0.802	0.972	0.837
Mutluluk	0.976	0.917	0.888	0.984	0.902
Tarafsızlık	0.937	0.766	0.715	0.959	0.740
Üzüntü	0.940	0.762	0.754	0.965	0.758
Sürpriz	0.967	0.876	0.857	0.979	0.866
Ortalama	0.954	0.819	0.817	0.973	0.817

Çizelge 4.5 ResNet18 için performans ölçümlerinin değerleri

ResNet18	ACC	SNS	PRC	SPC	F1_score
Öfke	0.998	0.991	0.996	0.999	0.994
Dikkat	0.993	0.984	0.965	0.995	0.974
İğrenme	0.997	0.987	0.995	0.999	0.991
Korku	0.998	0.997596	0.988	0.998	0.992
Mutluluk	0.999	1	0.998	0.999	0.999
Tarafsızlık	0.994	0.969	0.990	0.998	0.979

Üzüntü	0.996	0.990	0.984	0.997	0.987
Sürpriz	0.999	0.998	1	1	0.999
Ortalama	0.997	0.989	0.989	0.998	0.989

Çizelge 4.6 ResNet50 için performans ölçümlerinin değerleri

ResNet50	ACC	SNS	PRC	SPC	F1_score
Öfke	0.998	0.985	0.998	0.999	0.992
Dikkat	0.992	0.973	0.969	0.995	0.971
İğrenme	0.997	0.990	0.989	0.998	0.989
Korku	0.998	0.990	0.996	0.999	0.993
Mutluluk	0.999	0.997	1	1	0.998
Tarafsızlık	0.992	0.978	0.960	0.994	0.969
Üzüntü	0.994	0.975	0.979	0.997	0.977
Sürpriz	0.999	0.998	0.996	0.999	0.997
Ortalama	0.996	0.986	0.986	0.998	0.986

Çizelge 4.7 MobileNetV2 için performans ölçümlerinin değerleri

MobileNetV2	ACC	SNS	PRC	SPC	F1_score
Öfke	0.994	0.973	0.986	0.998	0.979
Dikkat	0.985	0.969	0.913	0.987	0.940
İğrenme	0.994	0.979	0.979	0.997	0.979
Korku	0.996	0.989	0.982	0.997	0.985
Mutluluk	0.998	0.995	0.992	0.998	0.994
Tarafsızlık	0.987	0.945	0.954	0.993	0.950
Üzüntü	0.988	0.934	0.975	0.996	0.954
Sürpriz	0.997	0.990	0.991	0.998	0.991
Ortalama	0.993	0.972	0.972	0.996	0.971

Çizelge 4.8 VGG16 için performans ölçümlerinin değerleri

VGG16	ACC	SNS	PRC	SPC	F1_score
Öfke	0.993	0.960	0.984	0.997	0.972
Dikkat	0.976	0.926	0.883	0.983	0.904
İğrenme	0.994	0.969	0.984	0.997	0.976

Korku	0.986	0.973	0.917	0.988	0.944
Mutluluk	0.997	0.990	0.991	0.998	0.991
Tarafsızlık	0.981	0.909	0.942	0.991	0.925
Üzüntü	0.986	0.963	0.929	0.990	0.946
Sürpriz	0.986	0.922	0.977	0.996	0.949
Ortalama	0.987	0.951	0.951	0.993	0.951

Çalışmadaki altı önceden eğitilmiş ağ arasındaki sonuçları karşılaştırmak için, her ağ için değerlendirme parametrelerinin ortalama değerleri ayrıldı. Bu değerler, sonuçların karşılaştırılmasını kolaylaştırmak için bir grafikte (Şekil 4.8) temsil edilebilir.



Şekil 4.8 Bu çalışmada kullanılan altı derin öğrenme yöntemi için performans ölçümlerinin ortalamalarının karşılaştırılması

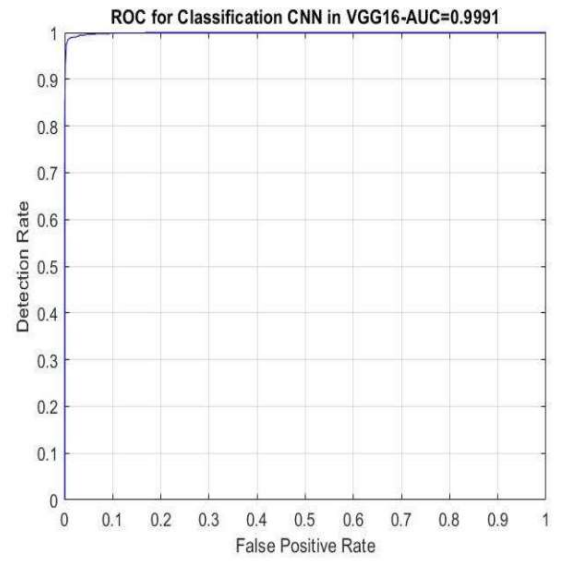
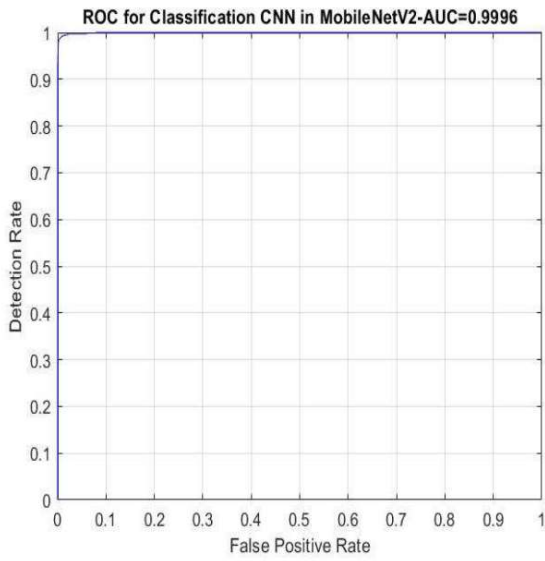
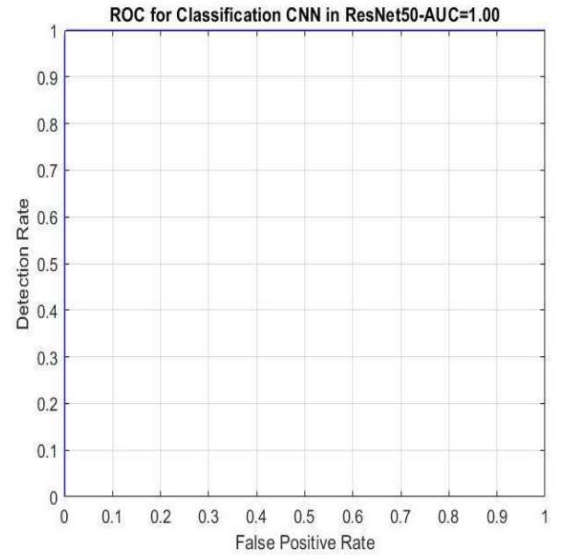
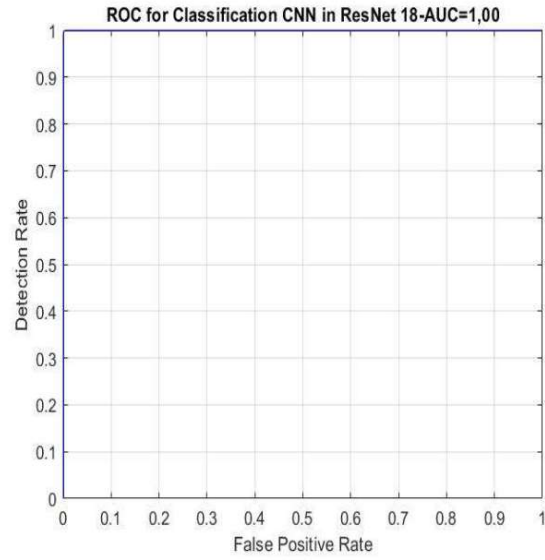
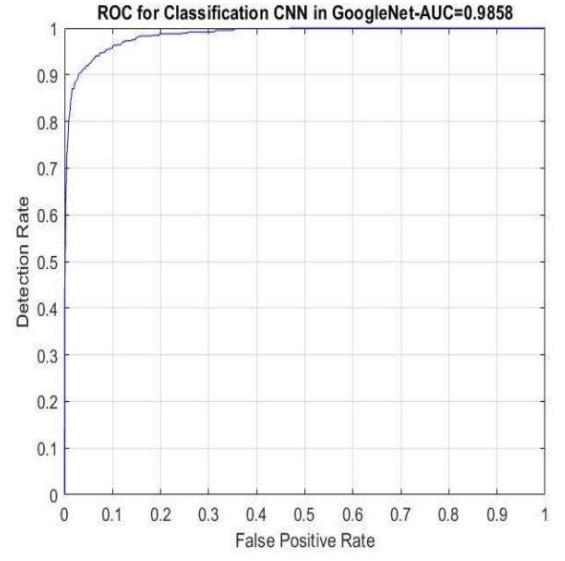
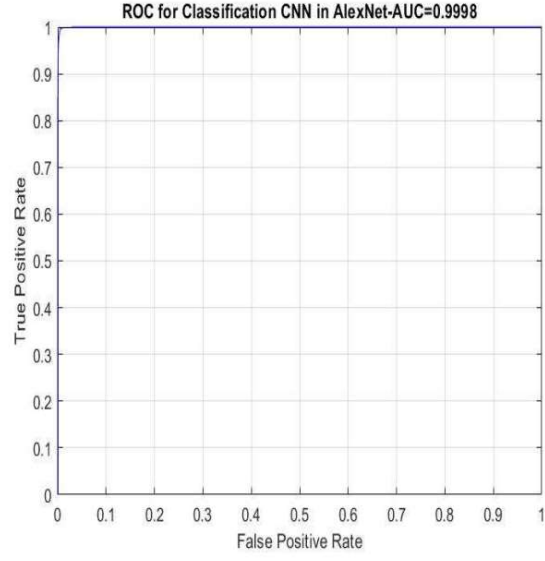
Önemleri nedeniyle, bu çalışmada AUC-ROC eğrileri de dahil edilmiştir (Şekil 4.9). AUC-ROC eğrisi, bu çalışmada kullanılan CNN ağlarının sınıflandırıcılarının performansını görselleştirmeye yardımcı olur ve ROC eğrisi, sınıflandırıcının tüm olası eşiklerdeki performansını özetler. ROC eğrisinin grafiği, tüm olası eşikler için yanlış pozitif oranı (FPR) x-ekseninde ve doğru pozitif oranını (TPR) y-ekseninde çizilir.

AUC: AUC, ROC eğrisinin altındaki alanı ifade eder. Sınıflandırıcı ne kadar iyiye, doğru pozitif oranı artacak ve eğrinin altındaki alan 1'e yakın olacaktır.

Konfüzyon matrislerini (Şekil 4.2'den 4.7'ye, Çizelgeler 4.3'ten 4.8'e) inceleyerek, en iyi sonuçların ResNet18 ve ona çok yakın olan ResNet50 tarafından kaydedildiği açıktır. ResNet18+TL+KFCV için performans metriklerinin (ACC, SNS, PRC, SPC, F1-Skor) ortalama değerleri sırasıyla 0.997, 0.989, 0.989, 0.998, 0.989 iken ResNet50+TL+KFCV için sırasıyla 0.996, 0.986, 0.986, 0.998, 0.986'dır. Bu sonuçları sırasıyla MobileNetV2+TL+KFCV, AlexNet+TL+KFCV ve VGG16+TL+KFCV izlerken, en düşük sonuçlar GoogleNet+TL+KFCV tarafından kaydedildi.

Ayrıca, ROC eğrilerini ve AUC değerlerini (Bkz. Şekil 4.9) inceleyerek, ResNet18+TL+KFCV ve ResNet50+TL+KFCV önceden eğitilmiş ağırları kullanıldığında en iyi sonuçların elde edildiği görülmektedir. Bu, öfke, iğrenme, korku, mutluluk, tarafsızlık, üzüntü, dikkat ve sürpriz yüz ifadelerinin otomatik tanınması için en iyi sonuçları elde ettikleri anlamına gelir.

Çizelgeleri (Çizelgeler 4.3, 4.7 ve 4.8) incelediğimizde AlexNet, MobileNetV2 ve VGG16 için öfke, iğrenme, korku, mutluluk ve sürpriz yüz ifadeleri için F1-Skor değerlerinin yüksek olduğu görülebilir. Bununla birlikte, tarafsızlık, dikkat ve üzüntü yüz ifadeleri için en yüksek değerdan yaklaşık %5 ila %7 oranında düşük oldukları görülebilir.



Şekil 4.9 Altı CNN modelin sınıflandırmadaki ROC-AUC grafikleri

GoogleNet için (Çizelge 4.4), öfke, iğrenme, korku, mutluluk ve sürpriz yüz ifadeleri için F1-Skor değerlerinin yaklaşık %85 olduğu görülebilir, ancak tarafsızlık, dikkat ve üzüntü yüz ifadeleri için bu değerlerin en yüksek değerden yaklaşık %20 düşük olduğu da görülebilir. ResNet18 ve ResNet50 için F1-Skor değerlerine (Çizelgeler 4.5 ve 4.6) baktığımızda, öfke, iğrenme, korku, mutluluk ve sürpriz yüz ifadeleri için skorların çok yüksek olduğu, ancak tarafsızlık, dikkat ve üzüntü yüz ifadeleri için yaklaşık %2 oranında düştüğü görülebilir.

Tarafsızlık, dikkat ve üzüntü yüz ifadelerinin bir kişiden diğerine farklılık gösterebileceği ve bir kişinin dikkat durumunu ve tarafsızlık durumunu tanımanın zor olabileceği mümkündür, bu nedenle bu yüz ifadeleri için düşük F1-Skor değerleri dikkate alınabilir.

Yukarıdaki tüm sonuçlardan, en iyi sonuçların ResNet18 ve ResNet50 tarafından kaydedildiği açıktır. ResNet18 ve ResNet50'yi eğitim ve test aşamalarının zamanı açısından karşılaştırarak, ResNet18'in daha az süre kaydettiği görülebilir (Bkz. Çizelge 4.1). Bu, ders boyunca öğrencilerin duygusal durumlarını otomatik olarak takip etmenin, öğreticiye anında geri bildirim sağlamak ve eğitim deneyimini iyileştirmeye yardımcı olmak için önerilen bir yaklaşım olabilir.

Bu sonuç bölümünde, 70 öğrencinin yüz ifadelerini temsil eden Mühendislik fakültesi öğrencilerinin yüz görüntülerinden oluşan kendi veritabanımız oluşturuldu. Öğrencilerin yüzlerine ait görüntü sayısı, yukarıda belirtilen sekiz yüz ifadesi için eşit olarak dağıtılmış olup, her kategori için 840 görüntüdür. Mevcut çalışmaların çoğu genel veritabanlarını kullandı ve araştırmamızda kullandığımız gibi özel verilere sahip değildi. Mevcut çalışmaların çoğu altı veya yedi kategoriye odaklanırken, çalışmamız sekiz kategoriye odaklandı. Ayrıca, araştırmamızda derin öğrenme, transfer öğrenme ve k-kat çarpaz doğrulama birlikte kullanıldı. Elde edilen sonuçlar, yüz ifadesi tanıma konusundaki benzer çalışmalarla karşılaştırıldı ve bu karşılaştırmaların özeti Çizelge 4.9'da gösterilmektedir. Çizelgeden, mevcut çalışmanın sonuçlarının yüksek doğruluk ve AUC değerleri verdiği açıktır. Ayrıca, performans metrikleri (Duyarlılık, Özgüllük, Kesinlik ve F1-Skor değeri) için elde edilen sonuçlar yüksektir. Bu karşılaştırmadan, bu çalışmanın eğitim amaçlı öğrencilerin yüz ifadelerinin tanınması alanına katkıda bulunabileceği görülebilir.

Çizelge 4.9 Mevcut çalışmanın yüz ifadesi tanıma için çeşitli derin öğrenme yöntemlerinin kullanıldığı diğer çalışmalarla karşılaştırılması

Çalışma	Yüz ifadeleri	Veriseti türü	Kullanılan yöntem	Genel Doğruluk	Ortalama Duyarlılık	Ortalama Kesinlik	Ortalama Özgüllük	Ortalama F1 skoru
(Fathalla ve ark., 2017)	Öfke İğrenme Mutluluk Tarafsızlık Üzüntü Sürpriz	CK+, MUG ve RAFD (Açık veriseti)	CNN VGG	93.33%	-	-	-	-
(Sang, ve ark., 2017)	Öfke İğrenme Korku Mutluluk Üzüntü Sürpriz Tarafsızlık	FERC-2013(Açık veriseti)	CNN VGG	71.9%	-	-	-	-
(Qayyum ve ark., 2017)	Öfke İğrenme Korku Mutluluk Üzüntü Sürpriz	JAFFE and CK+ (Açık veriseti)	ANN +Stationary Wavelet Transform Features	98.8% 96.6%	-	-	-	-
(Carcagn i ve ark., 2015)	Öfke İğrenme Korku Mutluluk Üzüntü Sürpriz Tarafsızlık	CK+, JAFFE, MMI	ORB+SV M	99%	-	-	-	-
(Chen ve ark., 2018)	Öfke İğrenme Korku Mutluluk Üzüntü Sürpriz Tarafsızlık	JAFFE and CK+ (Public dataset)	Softmax regression -based deep sparse autoencoder network (SRDSAN)	89.51%	-	-	-	-

(Prasada ve ark., 2019)	Öfke İğrenme Korku Mutluluk Üzüntü Sürpriz	Indian Face Database (Public dataset)	MSER (Maximall y Stable Extremal Regions) +SVM+A daboost	78%	-	-	-	-
(Chen ve ark., 2019)	Öfke İğrenme Korku Mutluluk Üzüntü Sürpriz	CK+ & BU- 4DFE (Public dataset)	VGG16	91.5 %	-	-	-	-
(Wu ve ark., 2021)	Öfke İğrenme Korku Mutluluk Üzüntü Sürpriz Tarafsızlık	CK+, JAFPE, SFEW2.0 (Public dataset)	Weight- Adapted Convoluti on Neural Network(WACNN)	94%	-	-	-	-
Bu çalışma	Mutluluk Üzüntü Öfke Korku İğrenme Dikkat Tarafsızlık Sürpriz	Kendi verisetimiz	AlexNet+ TL+KFC V	99%	97.1 %	97.7 %	99.5 %	96.8 %
Bu çalışma	Mutluluk Üzüntü Öfke Korku İğrenme Dikkat Tarafsızlık Sürpriz	Kendi verisetimiz	GoogleNe t+TL+KF CV	95.4 %	81.9 %	81.7 %	97.4 %	81.8 %
Bu çalışma	Mutluluk Üzüntü Öfke Korku İğrenme Dikkat Tarafsızlık Sürpriz	Kendi verisetimiz	ResNet18 +TL+KFC V	99.8 %	99%	99%	99.9 %	99%

Bu çalışma	Mutluluk Üzüntü Öfke Korku İğrenme Dikkat Tarafsızlık Sürpriz	Kendi verisetimiz	ResNet50 +TL+KFC V	99.7 %	98.6 %	98.6 %	99.8 %	98.6 %
Bu çalışma	Mutluluk Üzüntü Öfke Korku İğrenme Dikkat Tarafsızlık Sürpriz	Kendi verisetimiz	MobilNet V2+TL+K FCV	99.3 %	97.2 %	97.2 %	99.6 %	97.2 %
Bu çalışma	Mutluluk Üzüntü Öfke Korku İğrenme Dikkat Tarafsızlık Sürpriz	Kendi verisetimiz	VGG16+ TL+KFC V	98.8 %	95.2 %	95.2 %	99.3 %	95.1 %

4.2. Transfer Öğrenme ve ECOC-MSVM, KNN ve DT Gibi Makine Öğrenmesi Algoritmalarıyla Birlikte Önceden Eğitilmiş Evrişimli Sinir Ağları Kullanılarak Elde Edilen Sonuçlar

Araştırmanın ilk bölümünde elde edilen sonuçlar mükemmel ve çok etkiliydi, ancak Çizelge 4.1'de gösterildiği gibi eğitim ve doğrulama için gereken uzun süre, son katmanlarında karmaşık matematiksel işlemler gerçekleştiren evrişimli sinir ağlarının varlığından kaynaklanmaktadır. Ayrıca, zamanı artıran k-kat çarpaz doğrulama teknolojisinin kullanılması da bu süreyi artırır. Bu durum, ilk bölümde elde edilen etkili sonuçları korurken, eğitim ve test için gereken süreyi kısaltmak için bir çözüm aramamıza neden oldu. ECOC-MSVM, KNN ve DT gibi birçok makine öğrenmesi algoritmasının uygulanabileceğini keşfettik. Bu algoritmalar sınıflandırıcı olarak kullanılacaktır. Bu, evrişimli sinir ağlarının son katmanlarının silindiği ve bu tekniklerle değiştirildiği Çizelge 4.10' da görülebilir.

Bu sonuçlar bölümünde, transfer öğrenme AlexNet, GoogleNet, ResNet18, ResNet50 ve MobileNetV2 gibi önceden eğitilmiş ağlarda kullanıldı ve önceki ağların her biri için Çizelge 4.11'de gösterilen katmanlardan özellikler çıkarıldı. Bu özellikler daha sonra ECOC-MSVM,

KNN ve DT gibi makine öğrenimi yöntemlerine bağımsız olarak uygulandı ve bunlar sınıflandırıcı olarak kullanıldı.

Çizelge 4.10 GPU kullanılarak çeşitli CNN ağlarına ve birçok makine öğrenmesi algoritmasına ait eğitim ve test aşamalarının zamanları

CNN Ağı+TL+ Makine Öğrenimi Algoritması	Eğitim ve test için geçen süre (Dakika)
AlexNet+TL+ECOC-MSVM	≈20
AlexNet+TL+KNN	≈18
AlexNet+TL+DT	≈14
GoogleNet+TL+ ECOC-MSVM	≈ 16
GoogleNet+TL+ KNN	≈ 14
GoogleNet+TL+ DT	≈ 11
ResNet18+TL+ ECOC-MSVM	≈ 21
ResNet18+TL+ KNN	≈ 19
ResNet18+TL+ DT	≈ 15
ResNet50+TL+ ECOC-MSVM	≈ 23
ResNet50+TL+ KNN	≈ 22
ResNet50+TL+ DT	≈ 16
MobileNetV2+TL+ ECOC-MSVM	≈ 18
MobileNetV2+TL+ KNN	≈ 16
MobileNetV2+TL+ DT	≈ 13
AlexNet+GoogleNet+TL+ECOC-MSVM	≈ 24
AlexNet+ GoogleNet+TL+KNN	≈ 22
AlexNet+ GoogleNet+TL+DT	≈ 16
AlexNet+GoogleNet+ResNet18+TL+ECOC-MSVM	≈ 32
AlexNet+ GoogleNet+ ResNet18+TL+KNN	≈ 29
AlexNet+ GoogleNet+ ResNet18+TL+DT	≈ 23
AlexNet+GoogleNet+ResNet18+ResNet50+TL+ECOC-MSVM	≈ 40

AlexNet+ GoogleNet+ ResNet18+ ResNet50+TL+KNN	≈ 35
AlexNet+ GoogleNet+ ResNet18+ ResNet50+TL+DT	≈ 26
AlexNet+GoogleNet+ResNet18+ ResNet50+MobileNetV2+TL+ECOC- MSVM	≈ 44
AlexNet+ GoogleNet+ ResNet18+ ResNet50+ MobileNetV2+TL+KNN	≈ 40
AlexNet+ GoogleNet+ ResNet18+ ResNet50+ MobileNetV2+TL+DT	≈ 32

Çizelge 4.11 CNN ağlarının özellik çıkarılan katmanları

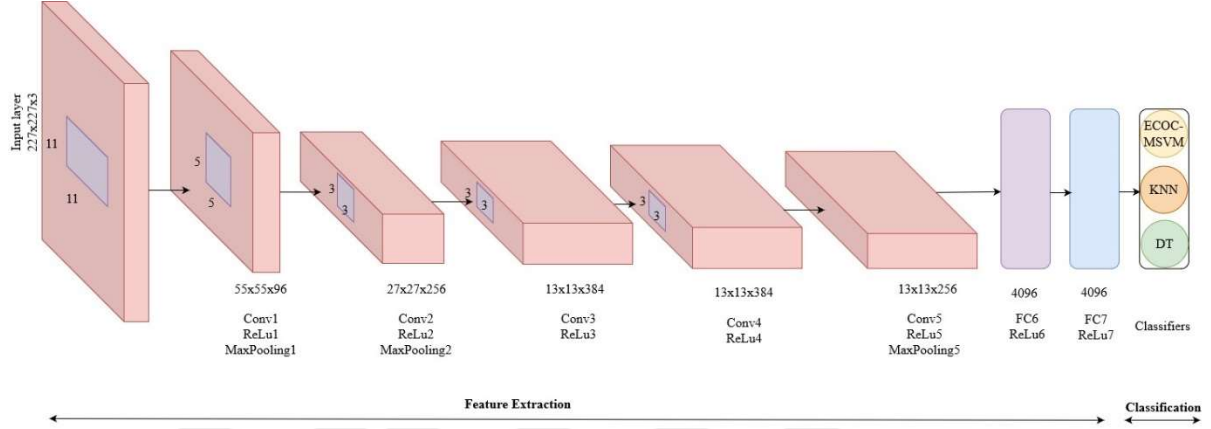
CNN Ağı	Özelliklerin çıkarıldığı katman
AlexNet	FC7
GoogleNet	Pool5_7x7_S1
ResNet18	Pool5
ResNet50	Average Pool
MobileNetV2	Global_average_Pooling2d

4.2.1. AlexNet tabanlı ECOC-MSVM, KNN ve DT

7968 yüz ifadesini (öfke, iğrenme, korku, mutluluk, tarafsızlık, üzüntü, dikkat ve sürpriz) içeren görüntülerden oluşan veri setimiz, önceden eğitilmiş ağ olan AlexNet'e girildi ve ardından bu ağa transfer öğrenme uygulandı. Transfer öğrenme sonrasında FC7 katmanından özellikler çıkarıldı ve önceki aşamadan elde edilen özellikler ECOC-MSVM, KNN veya DT gibi makine öğrenmesi algoritmalarından birine girildi. Bu, sınıflandırıcı olarak hizmet eder. Önceki makine öğrenimi yöntemleri bağımsız olarak kullanılır (Şekil 4.10). Ardından performans metrikleri hesaplandı ve konfüzyon matrisi (Şekil 4.11) çıkarıldı. Veri setimizin %70' i ağ eğitim sürecine ayrılmıştır ve bu, sonuçların bu

bölümündeki tüm örnekler için geçerlidir. Ayrıca, veri setimizin %30' u ağ test sürecine ayrılmıştır ve bu durum yine, sonuçların bu bölümündeki tüm örnekler için geçerlidir.

Test aşamasındaki veri kümesindeki görüntüler sekiz farklı yüz ifadesine (Öfke, dikkat, iğrenme, korku, mutluluk, tarafsızlık, üzüntü ve sürpriz) ayrılmıştır ve test aşamasında her bir ifade için 299 görüntü bulunmaktadır.



Şekil 4.10 AlexNet tabanlı ECOC-MSVM, KNN ve DT modelleri

Confusion Matrix: ALENet+MSVM									
Output Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	
	297 12.4%	0 0.0%	0 0.0%	0 0.0%	4 0.2%	4 0.2%	4 0.2%	0 0.0%	96.1% 3.9%
	0 0.0%	294 12.3%	0 0.0%	6 0.3%	6 0.3%	1 0.0%	0 0.0%	0 0.0%	95.8% 4.2%
	2 0.1%	0 0.0%	298 12.5%	0 0.0%	0 0.0%	4 0.2%	2 0.1%	0 0.0%	97.4% 2.6%
	0 0.0%	0 0.0%	0 0.0%	293 12.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	0 0.0%	0 0.0%	0 0.0%	0 0.0%	289 12.1%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	0 0.0%	5 0.2%	0 0.0%	0 0.0%	0 0.0%	286 12.0%	0 0.0%	2 0.1%	97.6% 2.4%
	0 0.0%	0 0.0%	1 0.0%	0 0.0%	0 0.0%	4 0.2%	293 12.2%	0 0.0%	98.3% 1.7%
	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	297 12.4%	100% 0.0%
	99.3% 0.7%	98.3% 1.7%	99.7% 0.3%	98.0% 2.0%	96.7% 3.3%	95.7% 4.3%	98.0% 2.0%	99.3% 0.7%	98.1% 1.9%
Target Class									

Confusion Matrix: ALENet+KNN									
Output Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	
	299 12.5%	0 0.0%	2 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.3% 0.7%
	0 0.0%	295 12.3%	0 0.0%	6 0.3%	0 0.0%	3 0.1%	0 0.0%	0 0.0%	97.0% 3.0%
	0 0.0%	0 0.0%	297 12.4%	0 0.0%	0 0.0%	0 0.0%	2 0.1%	0 0.0%	99.3% 0.7%
	0 0.0%	0 0.0%	0 0.0%	293 12.2%	0 0.0%	0 0.0%	0 0.0%	2 0.1%	99.3% 0.7%
	0 0.0%	0 0.0%	0 0.0%	0 0.0%	291 12.2%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	0 0.0%	4 0.2%	0 0.0%	0 0.0%	0 0.0%	292 12.2%	2 0.1%	0 0.0%	98.0% 2.0%
	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 0.1%	0 0.0%	295 12.3%	0 0.0%	99.3% 0.7%
	0 0.0%	0 0.0%	0 0.0%	0 0.0%	6 0.3%	4 0.2%	0 0.0%	297 12.4%	96.7% 3.3%
	100% 0.0%	98.7% 1.3%	99.3% 0.7%	98.0% 2.0%	97.3% 2.7%	97.7% 2.3%	98.7% 1.3%	99.3% 0.7%	98.6% 1.4%
Target Class									

Confusion Matrix: ALENet+Decision Trees									
Output Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	
	279 11.7%	11 0.5%	12 0.5%	8 0.3%	1 0.0%	14 0.6%	14 0.6%	5 0.2%	81.1% 18.9%
	11 0.5%	271 11.3%	3 0.1%	0 0.0%	3 0.1%	18 0.8%	6 0.3%	9 0.4%	84.4% 15.6%
	0 0.0%	2 0.1%	278 11.6%	9 0.4%	0 0.0%	4 0.2%	5 0.2%	6 0.3%	91.4% 8.6%
	5 0.2%	3 0.1%	3 0.1%	273 11.4%	3 0.1%	3 0.1%	3 0.1%	6 0.3%	91.3% 8.7%
	4 0.2%	0 0.0%	2 0.1%	0 0.0%	282 11.8%	4 0.2%	0 0.0%	6 0.3%	94.6% 5.4%
	0 0.0%	5 0.2%	0 0.0%	4 0.2%	6 0.3%	248 10.4%	0 0.0%	6 0.3%	92.2% 7.8%
	0 0.0%	7 0.3%	1 0.0%	4 0.2%	2 0.1%	6 0.3%	271 11.3%	5 0.2%	91.6% 8.4%
	0 0.0%	0 0.0%	0 0.0%	1 0.0%	2 0.1%	2 0.1%	0 0.0%	256 10.7%	98.1% 1.9%
	93.3% 6.7%	90.6% 9.4%	93.0% 7.0%	91.3% 8.7%	94.3% 5.7%	82.9% 17.1%	90.6% 9.4%	85.6% 14.4%	90.2% 9.8%
Target Class									

Şekil 4.11 AlexNet tabanlı ECOC-MSVM, KNN ve DT için konfüzyon matrisleri

Şekil 4.11'de sunulan konfüzyon matrisinden (4.1) ile (4.5) denklemlerine göre ek performans metriklerini hesaplayabilir ve bunları sistem değerlendirmesi için kullanabiliriz. Bu değerler, Çizelge 4.12 ve 4.14 arasında gösterilmiştir.

Çizelge 4.12 AlexNet+ECOC-MSVM için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.994147	0.961165	0.99904	0.993311	0.976974
Dikkat	0.992475	0.957655	0.997602	0.983278	0.970297
İğrenme	0.996237	0.973856	0.999521	0.996656	0.985124

Korku	0.997492	1	0.997141	0.979933	0.989865
Mutluluk	0.995819	1	0.995245	0.966555	0.982993
Tarafsızlık	0.991639	0.976109	0.993807	0.956522	0.966216
Üzüntü	0.995401	0.983221	0.997135	0.979933	0.981575
Sürpriz	0.999164	1	0.999045	0.993311	0.996644
Ortalama	0.995297	0.981501	0.997317	0.981187	0.981211

Çizelge 4.13 AlexNet+KNN için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.999164	0.993355	1	1	0.996667
Dikkat	0.994565	0.970395	0.998084	0.986622	0.978441
İğrenme	0.998328	0.993311	0.999044	0.993311	0.993311
Korku	0.996656	0.99322	0.997139	0.979933	0.986532
Mutluluk	0.996656	1	0.996192	0.973244	0.986441
Tarafsızlık	0.994565	0.979866	0.996657	0.976589	0.978224
Üzüntü	0.997492	0.993266	0.998091	0.986622	0.989933
Sürpriz	0.994983	0.967427	0.999041	0.993311	0.980198
Ortalama	0.996551	0.986355	0.998031	0.986204	0.986218

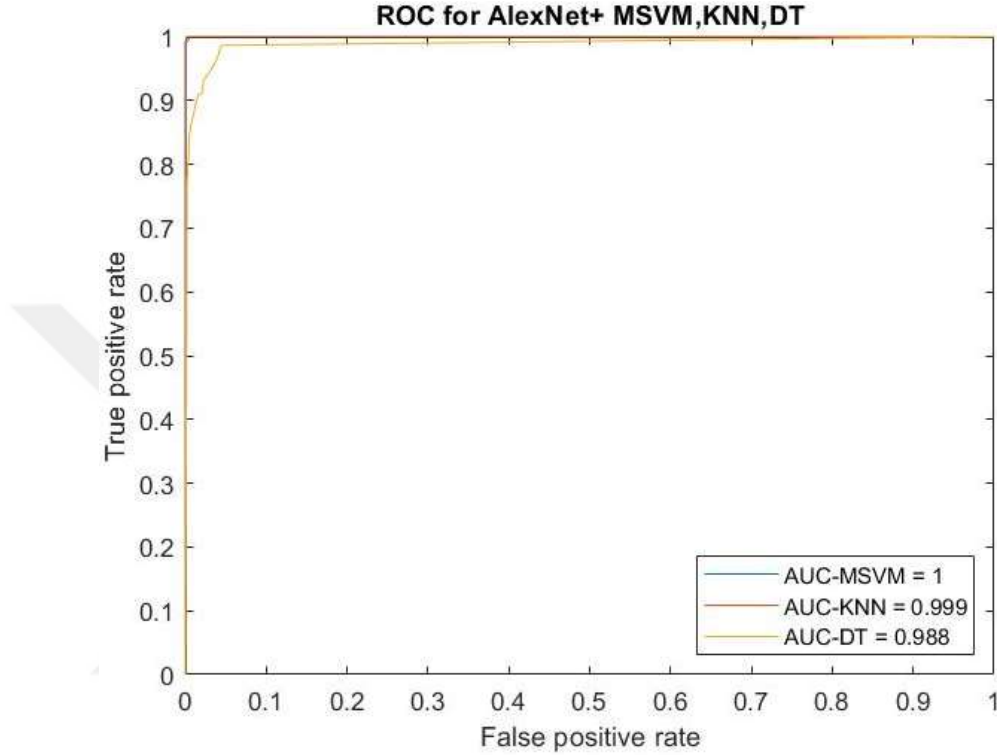
Çizelge 4.14 AlexNet+DT için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.964465	0.811047	0.990234	0.93311	0.867807
Dikkat	0.967391	0.844237	0.98648	0.906355	0.874194
İğrenme	0.980351	0.914474	0.989943	0.929766	0.922056
Korku	0.978261	0.913043	0.987578	0.913043	0.913043
Mutluluk	0.986204	0.946309	0.991882	0.943144	0.944724
Tarafsızlık	0.9699	0.921933	0.975977	0.829431	0.873239
Üzüntü	0.977843	0.915541	0.986641	0.906355	0.910924
Sürpriz	0.979933	0.980843	0.979822	0.856187	0.914286
Ortalama	0.975543	0.905928	0.98607	0.902174	0.902534

AUC-ROC eğrisi, çalışmanın sonuçlarının bu bölümünde kullanılan ağ sınıflandırıcılarının performansını göstermede yardımcı olurken, ROC eğrisi sınıflandırıcının tüm olası

eşiklerdeki performansını özetler. Tüm olası eşikler için ROC eğrisi grafiği, y ekseninde doğru pozitif oranı (TPR) ve x ekseninde yanlış pozitif oranı (FPR) ile belirtilir.

AUC, ROC eğrisinin altında kalan alanın kısaltmasıdır. Eğer sınıflandırıcı mükemmelse, doğru pozitif oranı yükselecek ve eğrinin altındaki alan 1'e yaklaşacaktır.



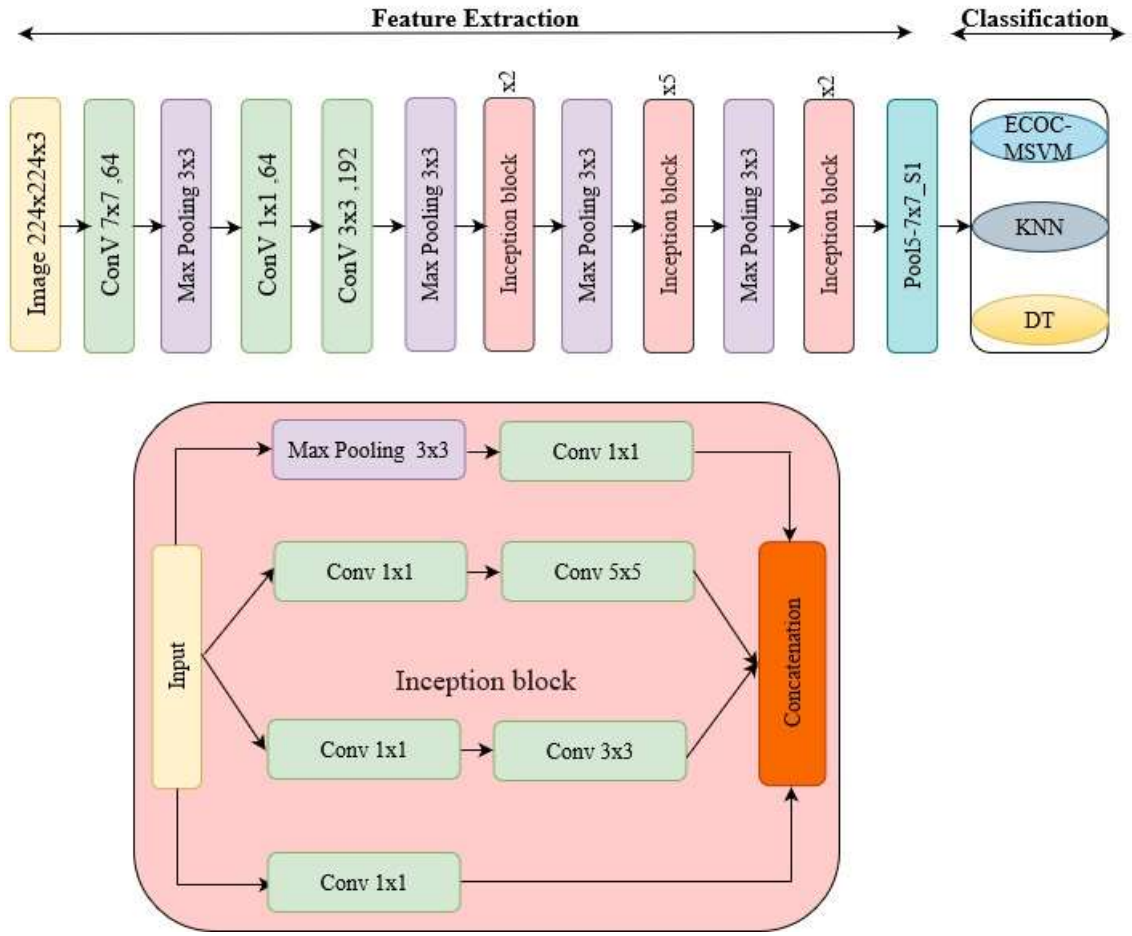
Şekil 4.12 AlexNet tabanlı makine öğrenimi yöntemleri için ROC eğrileri

Şekil 4.11 ve 4.12 ve Çizelge 4.11 ila 4.13 açıkça görülmektedir ki en iyi sonuçlar AlexNet+KNN ile kaydedilmiştir. KNN sınıflandırıcısı en iyi sınıflandırıcı olarak öne çıkmıştır ve onu AlexNet+ECOC-MSVM izlemektedir.

4.2.2. GoogleNet tabanlı ECOC-MSVM, KNN ve DT

7968 yüz ifadesini (öfke, iğrenme, korku, mutluluk, tarafsızlık, üzüntü, dikkat ve sürpriz) içeren görüntülerden oluşan veri setimiz, önceden eğitilmiş ağ olan GoogleNet'e girildi ve ardından bu ağa transfer öğrenme uygulandı. Transfer öğrenme sonrasında Pool5-7x7_S1 katmanından çıkarılan özellikler ECOC-MSVM, KNN veya DT gibi makine öğrenmesi algoritmalarından birine girildi. Bu, sınıflandırıcı olarak işlev görmektedir. Şekil 4.13, önceki makine öğrenimi yaklaşımlarının ayrı ayrı nasıl uygulandığını göstermektedir.

Ardından, performans ölçütleri belirlendi ve Şekil 4.14'te gösterilen konfüzyon matrisinde sunuldu.



Şekil 4.13 GoogleNet tabanlı ECOC-MSVM, KNN ve DT modelleri

Confusion Matrix: GoogleNet+MSVM									
Output Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	
	297 12.4%	0 0.0%	0 0.0%	4 0.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	98.7% 1.3%
	0 0.0%	287 12.0%	0 0.0%	7 0.3%	0 0.0%	2 0.1%	4 0.2%	0 0.0%	95.7% 4.3%
	2 0.1%	0 0.0%	296 12.4%	0 0.0%	4 0.2%	0 0.0%	5 0.2%	0 0.0%	96.4% 3.6%
	0 0.0%	0 0.0%	0 0.0%	288 12.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	0 0.0%	0 0.0%	1 0.0%	0 0.0%	295 12.3%	0 0.0%	0 0.0%	0 0.0%	99.7% 0.3%
	0 0.0%	12 0.5%	0 0.0%	0 0.0%	0 0.0%	294 12.3%	0 0.0%	0 0.0%	96.1% 3.9%
	0 0.0%	0 0.0%	2 0.1%	0 0.0%	0 0.0%	3 0.1%	290 12.1%	0 0.0%	98.3% 1.7%
	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	299 12.5%	100% 0.0%
									99.3% 0.7%
									96.0% 4.0%
									99.0% 1.0%
									96.3% 3.7%
									98.7% 1.3%
									98.3% 1.7%
									97.0% 3.0%
									100% 0.0%
									98.1% 1.9%
Target Class									

Confusion Matrix: GoogleNet+KNN									
Output Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	
	299 12.5%	0 0.0%	2 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.3% 0.7%
	0 0.0%	295 12.3%	0 0.0%	0 0.0%	0 0.0%	4 0.2%	0 0.0%	0 0.0%	98.7% 1.3%
	0 0.0%	0 0.0%	297 12.4%	0 0.0%	0 0.0%	0 0.0%	3 0.1%	0 0.0%	99.0% 1.0%
	0 0.0%	0 0.0%	0 0.0%	296 12.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	0 0.0%	0 0.0%	0 0.0%	0 0.0%	299 12.5%	0 0.0%	0 0.0%	4 0.2%	98.7% 1.3%
	0 0.0%	4 0.2%	0 0.0%	0 0.0%	0 0.0%	295 12.3%	1 0.0%	0 0.0%	98.3% 1.7%
	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	295 12.3%	0 0.0%	100% 0.0%
	0 0.0%	0 0.0%	0 0.0%	3 0.1%	0 0.0%	0 0.0%	0 0.0%	295 12.3%	99.0% 1.0%
									100% 0.0%
									98.7% 1.3%
									99.3% 0.7%
									99.0% 1.0%
									100% 0.0%
									98.7% 1.3%
									98.7% 1.3%
									99.1% 0.9%
Target Class									

Confusion Matrix: GoogleNet+Decision Tree									
Output Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	
	268 11.2%	5 0.2%	10 0.4%	4 0.2%	2 0.1%	12 0.5%	9 0.4%	0 0.0%	86.5% 13.5%
	7 0.3%	286 12.0%	2 0.1%	11 0.5%	8 0.3%	15 0.6%	3 0.1%	6 0.3%	84.6% 15.4%
	0 0.0%	5 0.2%	268 11.2%	1 0.0%	0 0.0%	5 0.2%	17 0.7%	1 0.0%	90.2% 9.8%
	3 0.1%	0 0.0%	10 0.4%	268 11.2%	3 0.1%	0 0.0%	5 0.2%	0 0.0%	92.7% 7.3%
	7 0.3%	0 0.0%	1 0.0%	7 0.3%	277 11.6%	0 0.0%	6 0.3%	5 0.2%	91.4% 8.6%
	8 0.3%	3 0.1%	4 0.2%	2 0.1%	4 0.2%	259 10.8%	3 0.1%	3 0.1%	90.6% 9.4%
	3 0.1%	0 0.0%	4 0.2%	5 0.2%	1 0.0%	0 0.0%	256 10.7%	0 0.0%	95.2% 4.8%
	3 0.1%	0 0.0%	0 0.0%	1 0.0%	4 0.2%	8 0.3%	0 0.0%	284 11.9%	94.7% 5.3%
									89.6% 10.4%
									95.7% 4.3%
									89.6% 10.4%
									89.6% 10.4%
									92.6% 7.4%
									86.6% 13.4%
									85.6% 14.4%
									95.0% 5.0%
									90.6% 9.4%
Target Class									

Şekil 4.14 GoogleNet tabanlı ECOC-MSVM, KNN ve DT için konfüzyon matrisleri

Şekil 4.14’deki konfüzyon matrislerine dayanarak (4.1) ile (4.5) denklemlerine göre diğer performans metriklerini hesaplayabilir ve ardından çeşitli makine öğrenme algoritmalarıyla çalışan ağı değerlendirebiliriz. Bu değerler Çizelge 4.15 ila 4.17 arasında verilmiştir.

Çizelge 4.15 GoogleNet+ECOC-MSVM için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.997492	0.986711	0.999044	0.993311	0.99
Dikkat	0.989548	0.956667	0.994264	0.959866	0.958264
İğrenme	0.994147	0.964169	0.998561	0.989967	0.976898
Korku	0.995401	1	0.994772	0.963211	0.981261
Mutluluk	0.99791	0.996622	0.998092	0.986622	0.991597
Tarafsızlık	0.992893	0.960784	0.997603	0.983278	0.971901
Üzüntü	0.994147	0.983051	0.995708	0.9699	0.976431
Sürpriz	1	1	1	1	1
Ortalama	0.995192	0.981	0.997255	0.980769	0.980794

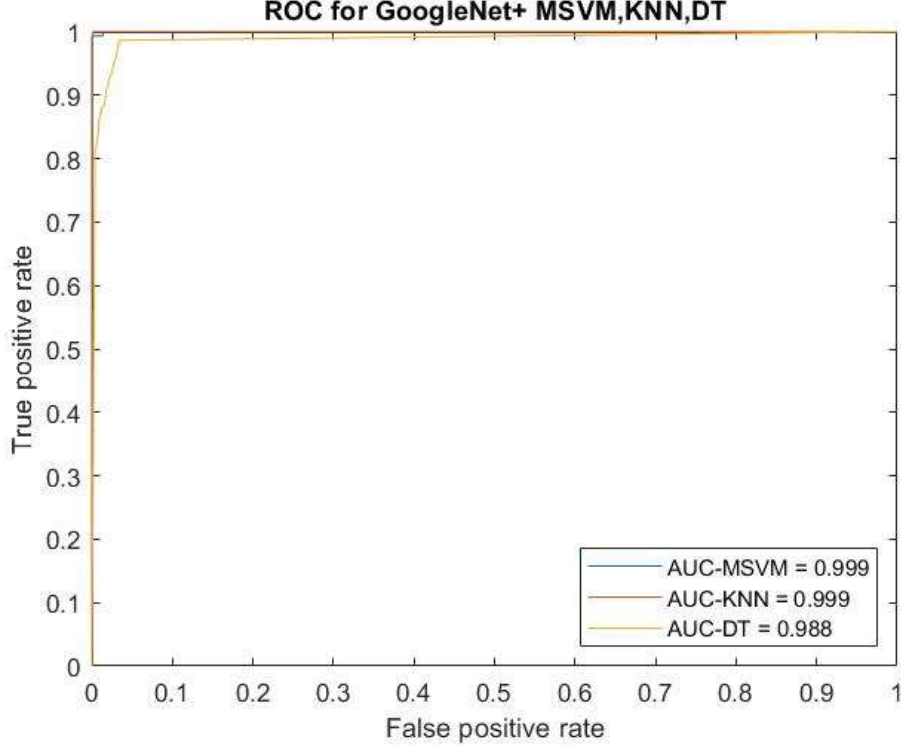
Çizelge 4.16 GoogleNet+KNN için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.999164	0.993355	1	1	0.996667
Dikkat	0.996656	0.986622	0.998089	0.986622	0.986622
İğrenme	0.99791	0.99	0.999044	0.993311	0.991653
Korku	0.998746	1	0.998569	0.989967	0.994958
Mutluluk	0.998328	0.986799	1	1	0.993355
Tarafsızlık	0.996237	0.983333	0.998088	0.986622	0.984975
Üzüntü	0.998328	1	0.998093	0.986622	0.993266
Sürpriz	0.997074	0.989933	0.99809	0.986622	0.988275
Ortalama	0.997805	0.991255	0.998746	0.991221	0.991221

Çizelge 4.17 GoogleNet+DT için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.969482	0.864516	0.98511	0.896321	0.880131
Dikkat	0.972826	0.846154	0.993671	0.956522	0.897959
İğrenme	0.974916	0.902357	0.985203	0.896321	0.899329
Korku	0.978261	0.927336	0.985259	0.896321	0.911565
Mutluluk	0.979933	0.914191	0.989469	0.926421	0.920266
Tarafsızlık	0.97199	0.905594	0.981007	0.866221	0.88547

Üzüntü	0.976589	0.951673	0.979746	0.856187	0.901408
Sürpriz	0.98704	0.946667	0.99283	0.949833	0.948247
Ortalama	0.97638	0.907311	0.986537	0.905518	0.905547



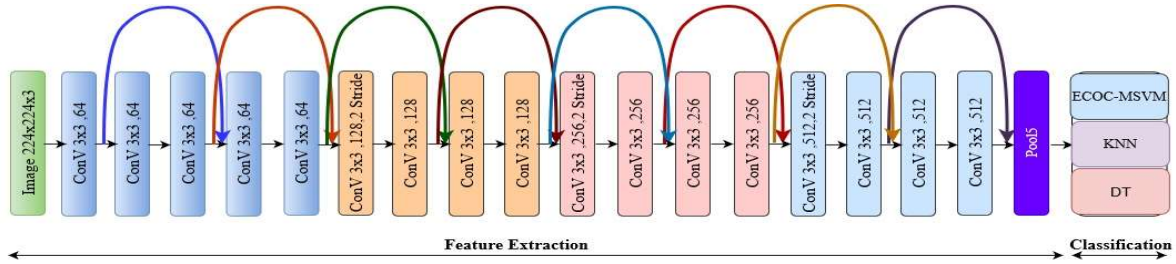
Şekil 4.15 GoogleNet tabanlı makine öğrenimi yöntemleri için ROC eğrileri

GoogleNet+KNN kullanılırken elde edilen sonuçlar, GoogleNet+ECOC-MSVM ile elde edilen sonuçlara yakındı. Bu durum Şekil 4.14 ve Çizelge 4.15 ve 4.17’ de görülebilir. GoogleNet+DT ile elde edilen sonuçlar Çizelge 4.17’ de görülmektedir. Şekil 4.15'teki ROC-AUC grafikleri, üç sınıflandırıcıda sınıflandırma sürecinin güvenilirliğini göstermektedir. Önceki üç sınıflandırıcıda sınıflandırmanın neredeyse ideal olduğu bulunmuştur.

4.2.3. ResNet18 tabanlı ECOC-MSVM, KNN ve DT

Öfke, iğrenme, korku, mutluluk, tarafsızlık, üzüntü, dikkat ve sürprizi ifade eden 7968 görüntüden oluşan veri setimiz, önceden eğitilmiş ağ olan ResNet18'e girildi ve ardından bu ağa transfer öğrenme uygulandı. Transfer öğrenme sonrasında Pool5 katmanından çıkarılan özellikler ECOC-MSVM, KNN veya DT gibi makine öğrenmesi algoritmalarından birine girildi. Bu, sınıflandırıcı olarak görev yapar. Şekil 4.16, önceki makine öğrenimi

yaklaşımlarının bağımsız olarak nasıl uygulandığını göstermektedir. Ardından, performans ölçütleri Şekil 4.17'de sunulan konfüzyon matrisinde belirlendi.



Şekil 4.16 ResNet18 tabanlı ECOC-MSVM, KNN ve DT modelleri

Confusion Matrix: ResNet18+MSVM									
Output Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	
Anger	298 12.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
Attention	0 0.0%	290 12.1%	0 0.0%	0 0.0%	0 0.0%	7 0.3%	0 0.0%	0 0.0%	97.6% 2.4%
Disgust	1 0.0%	0 0.0%	298 12.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.7% 0.3%
Fear	0 0.0%	2 0.1%	0 0.0%	299 12.5%	0 0.0%	0 0.0%	0 0.0%	4 0.2%	98.0% 2.0%
Happiness	0 0.0%	2 0.1%	0 0.0%	0 0.0%	297 12.4%	0 0.0%	0 0.0%	0 0.0%	99.3% 0.7%
Neutrality	0 0.0%	3 0.1%	0 0.0%	0 0.0%	0 0.0%	288 12.0%	0 0.0%	0 0.0%	99.0% 1.0%
Sadness	0 0.0%	2 0.1%	1 0.0%	0 0.0%	0 0.0%	4 0.2%	299 12.5%	0 0.0%	97.7% 2.3%
Surprise	0 0.0%	0 0.0%	0 0.0%	2 0.1%	0 0.0%	0 0.0%	0 0.0%	295 12.3%	99.3% 0.7%
	99.7% 0.3%	97.0% 3.0%	99.7% 0.3%	100% 0.0%	99.3% 0.7%	96.3% 3.7%	100% 0.0%	98.7% 1.3%	98.8% 1.2%
Target Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	

Confusion Matrix: ResNet18+KNN									
Output Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	
Anger	299 12.5%	0 0.0%	3 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.0% 1.0%
Attention	0 0.0%	295 12.3%	0 0.0%	3 0.1%	0 0.0%	7 0.3%	0 0.0%	0 0.0%	96.7% 3.3%
Disgust	0 0.0%	0 0.0%	296 12.4%	0 0.0%	0 0.0%	0 0.0%	4 0.2%	0 0.0%	98.7% 1.3%
Fear	0 0.0%	0 0.0%	0 0.0%	292 12.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
Happiness	0 0.0%	0 0.0%	0 0.0%	0 0.0%	299 12.5%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
Neutrality	0 0.0%	4 0.2%	0 0.0%	0 0.0%	0 0.0%	292 12.2%	4 0.2%	0 0.0%	97.3% 2.7%
Sadness	0 0.0%	0 0.0%	0 0.0%	4 0.2%	0 0.0%	0 0.0%	291 12.2%	0 0.0%	98.6% 1.4%
Surprise	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	299 12.5%	100% 0.0%
	100% 0.0%	98.7% 1.3%	99.0% 1.0%	97.7% 2.3%	100% 0.0%	97.7% 2.3%	97.3% 2.7%	100% 0.0%	98.8% 1.2%
Target Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	

Confusion Matrix: ResNet18+DT									
Output Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	
Anger	271 11.3%	14 0.6%	4 0.2%	2 0.1%	7 0.3%	1 0.0%	11 0.5%	8 0.3%	85.2% 14.8%
Attention	5 0.2%	249 10.4%	0 0.0%	10 0.4%	9 0.4%	9 0.4%	0 0.0%	6 0.3%	86.5% 13.5%
Disgust	3 0.1%	3 0.1%	279 11.7%	8 0.3%	5 0.2%	8 0.3%	4 0.2%	7 0.3%	88.0% 12.0%
Fear	0 0.0%	3 0.1%	4 0.2%	269 11.2%	4 0.2%	0 0.0%	6 0.3%	5 0.2%	92.4% 7.6%
Happiness	2 0.1%	5 0.2%	3 0.1%	0 0.0%	267 11.2%	8 0.3%	2 0.1%	6 0.3%	91.1% 8.9%
Neutrality	12 0.5%	15 0.6%	0 0.0%	3 0.1%	0 0.0%	267 11.2%	4 0.2%	2 0.1%	88.1% 11.9%
Sadness	0 0.0%	3 0.1%	4 0.2%	4 0.2%	2 0.1%	1 0.0%	267 11.2%	0 0.0%	95.0% 5.0%
Surprise	6 0.3%	7 0.3%	5 0.2%	3 0.1%	5 0.2%	5 0.2%	5 0.2%	265 11.1%	88.0% 12.0%
	90.6% 9.4%	83.3% 16.7%	93.3% 6.7%	90.0% 10.0%	89.3% 10.7%	89.3% 10.7%	89.3% 10.7%	88.6% 11.4%	89.2% 10.8%
Target Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	

Şekil 4.17 ResNet18 tabanlı ECOC-MSVM, KNN ve DT için konfüzyon matrisleri

Çizelge 4.18 ResNet18+ECOC-MSVM için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.999582	1	0.999522	0.996656	0.998325
Dikkat	0.993311	0.976431	0.995704	0.9699	0.973154
İğrenme	0.999164	0.996656	0.999522	0.996656	0.996656
Korku	0.997492	0.980328	1	1	0.990066
Mutluluk	0.998328	0.993311	0.999044	0.993311	0.993311
Tarafsızlık	0.994147	0.989691	0.994764	0.963211	0.976271
Üzüntü	0.997074	0.977124	1	1	0.98843
Sürpriz	0.997492	0.993266	0.998091	0.986622	0.989933
Ortalama	0.997074	0.988351	0.998331	0.988294	0.988268

Çizelge 4.19 ResNet18+KNN için performans ölçümlerinin değerleri

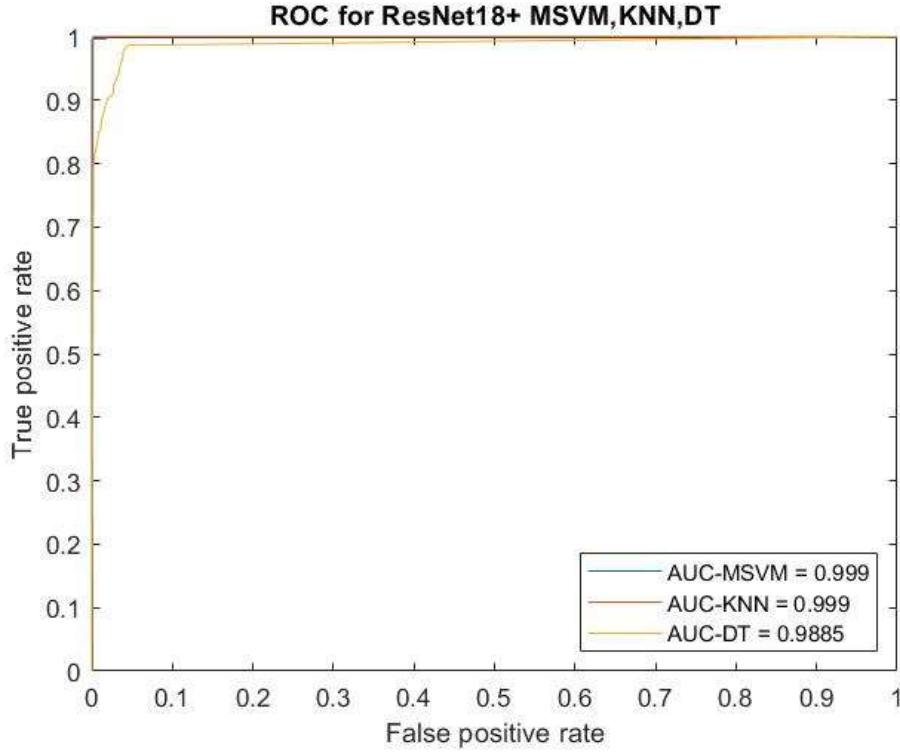
	ACC	SNS	SPC	PRC	F1_score
Öfke	0.998746	0.990066	1	1	0.995008
Dikkat	0.994147	0.967213	0.998083	0.986622	0.976821
İğrenme	0.997074	0.986667	0.998566	0.989967	0.988314
Korku	0.997074	1	0.996667	0.976589	0.988156
Mutluluk	1	1	1	1	1
Tarafsızlık	0.993729	0.973333	0.996654	0.976589	0.974958
Üzüntü	0.994983	0.986441	0.996185	0.973244	0.979798
Sürpriz	1	1	1	1	1
Ortalama	0.996969	0.987965	0.998269	0.987876	0.987882

Çizelge 4.20 ResNet18+DT için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.968645	0.852201	0.9865	0.906355	0.878444
Dikkat	0.962793	0.864583	0.976236	0.832776	0.848382
İğrenme	0.975753	0.880126	0.990361	0.93311	0.905844
Korku	0.978261	0.924399	0.985721	0.899666	0.911864
Mutluluk	0.975753	0.911263	0.984755	0.892977	0.902027
Tarafsızlık	0.971572	0.881188	0.984682	0.892977	0.887043

Üzüntü	0.980769	0.950178	0.984841	0.892977	0.92069
Sürpriz	0.970736	0.880399	0.98374	0.886288	0.883333
Ortalama	0.973035	0.893042	0.984604	0.89214	0.892203

Şekil 4.17, Çizelge 4.18 ve 4.20, ECOC-SVM ve KNN tarafından kaydedilen sonuçların benzer olduğunu göstermektedir. Ayrıca, bunların genel olarak en iyi sonuçlar olmasıyla birlikte DT tarafından kaydedilen sonuçların da iyi olduğu görülebilir. Ancak çizelgeler en iyi sonuçlardan %8,8 daha düşük olduğunu göstermektedir.



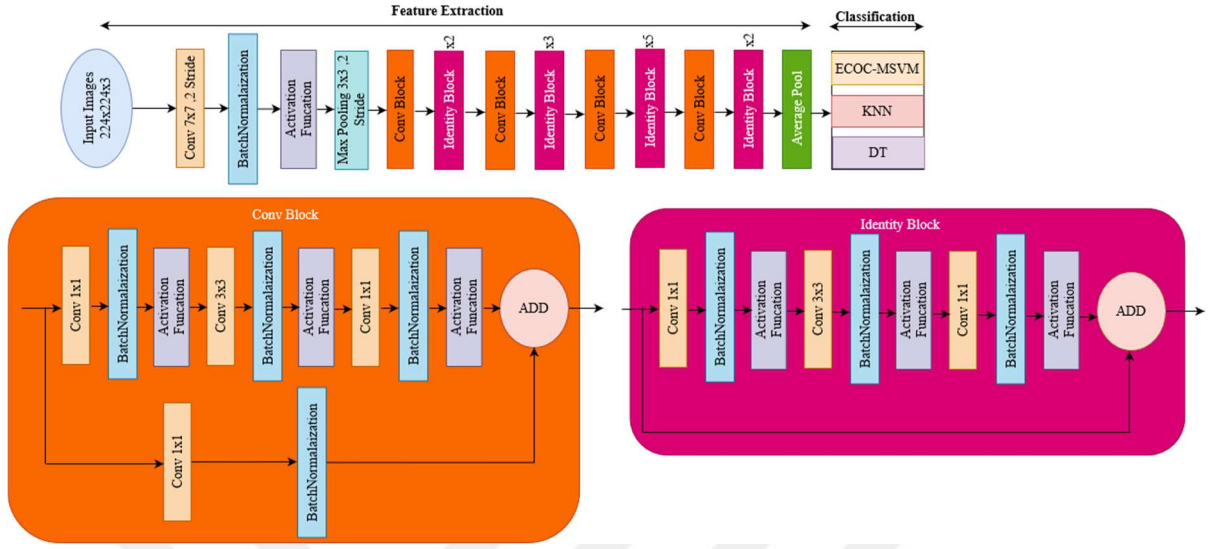
Şekil 4.18 ResNet18 tabanlı makine öğrenimi yöntemleri için ROC Eğrileri

Şekil 4.18'deki ROC eğrisinden görüldüğü gibi üç sınıflandırıcı için de AUC değerleri 0,99 civarındadır. Böylece, ResNet 18 tabanlı ECOC-SVM, KNN ve DT ile test edilen üç sınıflandırıcıda başarılı bir sınıflandırma doğruluğu sunmaktadır.

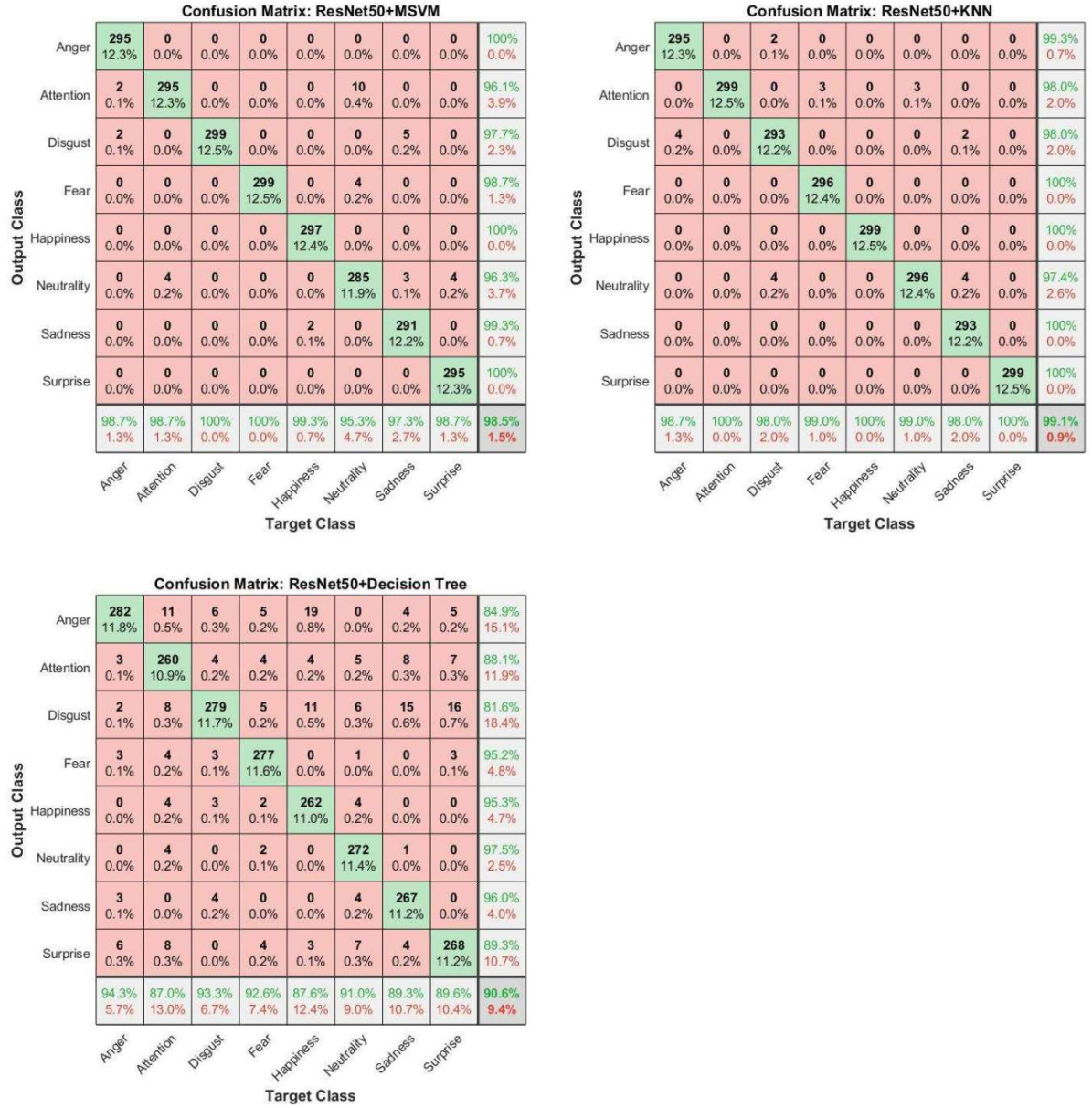
4.2.4. ResNet50 tabanlı ECOC-MSVM, KNN ve DT

Sekiz yüz duygusunu (öfke, tiksinti, korku, mutluluk, tarafsızlık, üzüntü, dikkat ve sürpriz) içeren 7968 görüntüden oluşan veri setimiz, önceden eğitilmiş ResNet50 ağına uygulanmıştır. Daha sonra bu ağa transfer öğrenme uygulandı ve Ortalama Havuz katmanından çıkarılan özellikler ECOC-MSVM, KNN veya DT gibi makine öğrenme algoritmalarından her birine ayrı ayrı uygulandı. Burada, Şekil 4.19, önceki makine öğrenimi

algoritmalarının bağımsız dağıtımını göstermektedir. Bunu takiben, Şekil 4.20'de elde edilen konfüzyon matrisleri sunulmaktadır.



Şekil 4.19 ResNet50 tabanlı ECOC-MSVM, KNN ve DT modelleri



Şekil 4.20 ResNet50 tabanlı ECOC-MSVM, KNN ve DT için konfüzyon matrisleri

Çizelge 4.21 ResNet50+ECOC-MSVM için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.998328	1	0.998093	0.986622	0.993266
Dikkat	0.993311	0.960912	0.998082	0.986622	0.973597
İğrenme	0.997074	0.977124	1	1	0.98843
Korku	0.998328	0.986799	1	1	0.993355
Mutluluk	0.999164	1	0.999045	0.993311	0.996644
Tarafsızlık	0.989548	0.962838	0.993321	0.953177	0.957983

Üzüntü	0.995819	0.993174	0.996189	0.973244	0.983108
Sürpriz	0.998328	1	0.998093	0.986622	0.993266
Ortalama	0.996237	0.985106	0.997853	0.98495	0.984956

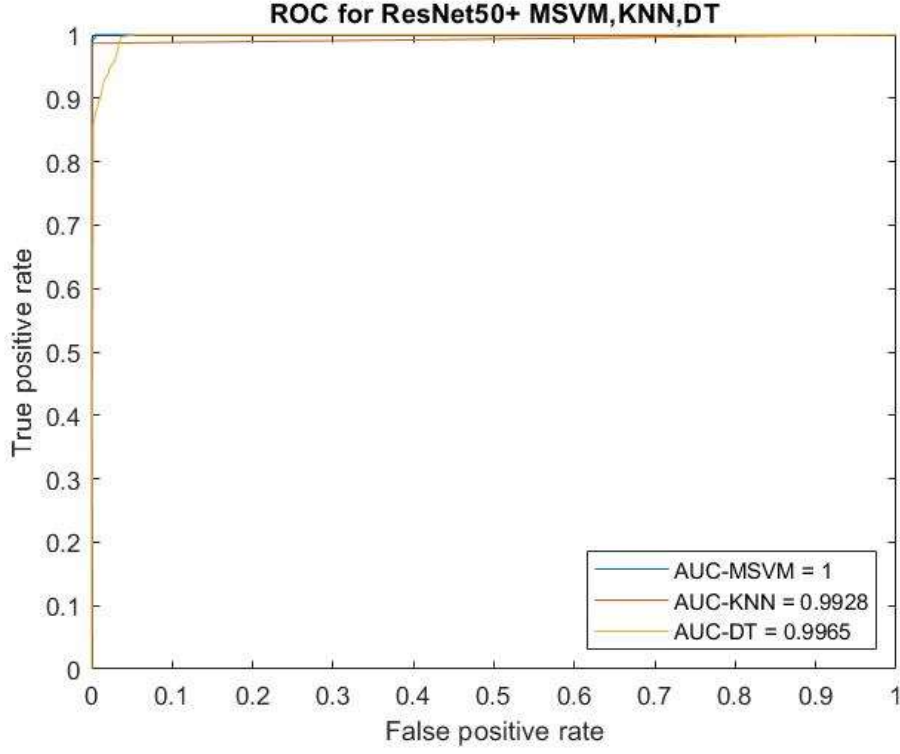
Çizelge 4.22 ResNet50+KNN için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.997492	0.993266	0.998091	0.986622	0.989933
Dikkat	0.997492	0.980328	1	1	0.990066
İğrenme	0.994983	0.979933	0.997133	0.979933	0.979933
Korku	0.998746	1	0.998569	0.989967	0.994958
Mutluluk	1	1	1	1	1
Tarafsızlık	0.995401	0.973684	0.998563	0.989967	0.981758
Üzüntü	0.997492	1	0.997141	0.979933	0.989865
Sürpriz	1	1	1	1	1
Ortalama	0.997701	0.990901	0.998687	0.990803	0.990814

Çizelge 4.23 ResNet50+DT için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.97199	0.849398	0.991748	0.943144	0.893819
Dikkat	0.969064	0.881356	0.981402	0.869565	0.875421
İğrenme	0.965301	0.815789	0.990244	0.93311	0.870515
Korku	0.98495	0.95189	0.989529	0.926421	0.938983
Mutluluk	0.979097	0.952727	0.982522	0.876254	0.912892
Tarafsızlık	0.985786	0.97491	0.987222	0.909699	0.941176
Üzüntü	0.982023	0.960432	0.984863	0.892977	0.925477
Sürpriz	0.973662	0.893333	0.985182	0.896321	0.894825
Ortalama	0.976484	0.909979	0.986589	0.905936	0.906638

ResNet50-KNN modeli, genel F1 Puanı için ResNet-50-ECOC-MSVM mimarisini %0,6 ve ResNet-50-DT mimarisini %8,5 oranında geride bırakmıştır. Bu, Şekil 4.20'de belirgin olarak gözükmemektedir ve ayrıca 4.21'den 4.23'e kadar olan Çizelgelerde bu durumu açıkça göstermektedir.

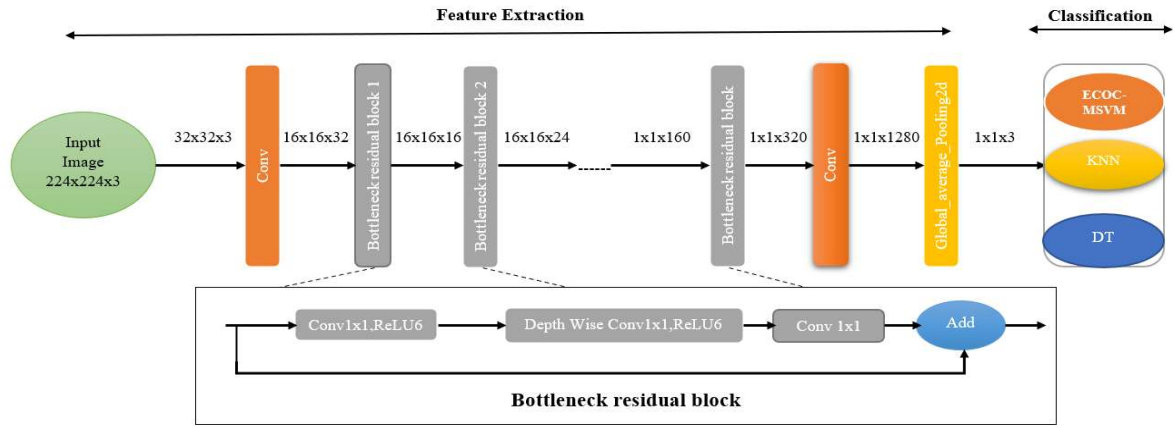


Şekil 4.21 ResNet50 tabanlı makine öğrenimi yöntemleri için ROC Eğrileri

ResNet50 tabanlı ECOC-SVM sınıflandırıcısı, AUC için 1 değerini kaydederek ideallığe ulaşmış, ancak KNN ve DT sınıflandırıcıları, AUC için 0,99 değerine erişerek ideale yaklaşmıştır. Bu değerler Şekil 4.21'de gösterilmiştir.

4.2.5. MobileNetV2 tabanlı ECOC-MSVM, KNN ve DT

Sekiz yüz duygusunu (öfke, tiksinti, korku, mutluluk, tarafsızlık, üzüntü, dikkat ve sürpriz) ifade eden 7968 görüntüden oluşan veri setimiz, önceden eğitilmiş MobileNetV2 ağına burada da uygulanmıştır. **Global_average_pooling2d** katmanından çıkarılan özellikler, ECOC-MSVM, KNN veya DT gibi makine öğrenimi algoritmaları kullanılarak sınıflandırılmıştır. Şekil 4.22, önceki makine öğrenimi algoritmalarının bağımsız dağıtımını göstermektedir. Performans ölçümleri daha sonra Şekil 4.23'te gösterilen konfüzyon matrislerinde sunulmaktadır.



Şekil 4.22 MobileNetV2 tabanlı ECOC-MSVM, KNN ve DT modelleri

Confusion Matrix: MobileNetV2+MSVM

Output Class \ Target Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	
Anger	297 12.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
Attention	0 0.0%	297 12.4%	0 0.0%	0 0.0%	6 0.3%	4 0.2%	0 0.0%	0 0.0%	96.7% 3.3%
Disgust	2 0.1%	0 0.0%	299 12.5%	0 0.0%	0 0.0%	0 0.0%	3 0.1%	0 0.0%	98.4% 1.6%
Fear	0 0.0%	1 0.0%	0 0.0%	293 12.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.7% 0.3%
Happiness	0 0.0%	0 0.0%	0 0.0%	0 0.0%	299 12.5%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
Neutrality	0 0.0%	1 0.0%	0 0.0%	6 0.3%	0 0.0%	293 12.2%	1 0.0%	0 0.0%	97.3% 2.7%
Sadness	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	291 12.2%	0 0.0%	100% 0.0%
Surprise	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	299 12.5%	100% 0.0%
	99.3% 0.7%	99.3% 0.7%	100% 0.0%	98.0% 2.0%	100% 0.0%	98.0% 2.0%	97.3% 2.7%	100% 0.0%	99.0% 1.0%

Confusion Matrix: MobileNetV2+KNN

Output Class \ Target Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	
Anger	299 12.5%	0 0.0%	2 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.3% 0.7%
Attention	0 0.0%	295 12.3%	0 0.0%	7 0.3%	0 0.0%	8 0.3%	0 0.0%	0 0.0%	95.2% 4.8%
Disgust	0 0.0%	0 0.0%	289 12.1%	0 0.0%	0 0.0%	0 0.0%	1 0.0%	0 0.0%	99.7% 0.3%
Fear	0 0.0%	0 0.0%	0 0.0%	292 12.2%	4 0.2%	0 0.0%	0 0.0%	0 0.0%	98.6% 1.4%
Happiness	0 0.0%	0 0.0%	0 0.0%	0 0.0%	295 12.3%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
Neutrality	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	291 12.2%	3 0.1%	0 0.0%	99.0% 1.0%
Sadness	0 0.0%	0 0.0%	8 0.3%	0 0.0%	0 0.0%	0 0.0%	295 12.3%	0 0.0%	97.4% 2.6%
Surprise	0 0.0%	4 0.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	299 12.5%	98.7% 1.3%
	100% 0.0%	98.7% 1.3%	96.7% 3.3%	97.7% 2.3%	98.7% 1.3%	97.3% 2.7%	98.7% 1.3%	100% 0.0%	98.5% 1.5%

Confusion Matrix: MobileNetV2+Decision Tree

Output Class \ Target Class	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	
Anger	262 11.0%	9 0.4%	5 0.2%	1 0.0%	2 0.1%	8 0.3%	0 0.0%	11 0.5%	87.9% 12.1%
Attention	5 0.2%	262 11.0%	9 0.4%	2 0.1%	6 0.3%	3 0.1%	6 0.3%	5 0.2%	87.9% 12.1%
Disgust	3 0.1%	4 0.2%	259 10.8%	0 0.0%	6 0.3%	6 0.3%	10 0.4%	5 0.2%	88.4% 11.6%
Fear	2 0.1%	3 0.1%	0 0.0%	281 11.7%	0 0.0%	3 0.1%	6 0.3%	0 0.0%	95.3% 4.7%
Happiness	16 0.7%	2 0.1%	0 0.0%	11 0.5%	282 11.8%	5 0.2%	7 0.3%	2 0.1%	86.8% 13.2%
Neutrality	9 0.4%	5 0.2%	2 0.1%	2 0.1%	1 0.0%	268 11.2%	2 0.1%	0 0.0%	92.7% 7.3%
Sadness	2 0.1%	8 0.3%	17 0.7%	2 0.1%	1 0.0%	0 0.0%	259 10.8%	1 0.0%	89.3% 10.7%
Surprise	0 0.0%	6 0.3%	7 0.3%	0 0.0%	1 0.0%	6 0.3%	9 0.4%	275 11.5%	90.5% 9.5%
	87.6% 12.4%	87.6% 12.4%	86.6% 13.4%	94.0% 6.0%	94.3% 5.7%	89.6% 10.4%	86.6% 13.4%	92.0% 8.0%	89.8% 10.2%

Şekil 4.23 MobileNetV2 tabanlı ECOC-MSVM, KNN ve DT için konfüzyon matrisleri

Çizelge 4.24 MobileNetV2+ECOC-MSVM için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.999164	1	0.999045	0.993311	0.996644
Dikkat	0.994983	0.967427	0.999041	0.993311	0.980198
İğrenme	0.99791	0.983553	1	1	0.991708
Korku	0.997074	0.996599	0.99714	0.979933	0.988196
Mutluluk	1	1	1	1	1
Tarafsızlık	0.994147	0.973422	0.997131	0.979933	0.976667
Üzüntü	0.996656	1	0.996192	0.973244	0.986441
Sürpriz	1	1	1	1	1
Ortalama	0.997492	0.990125	0.998569	0.989967	0.989982

Çizelge 4.25 MobileNetV2+KNN için performans ölçümlerinin değerleri

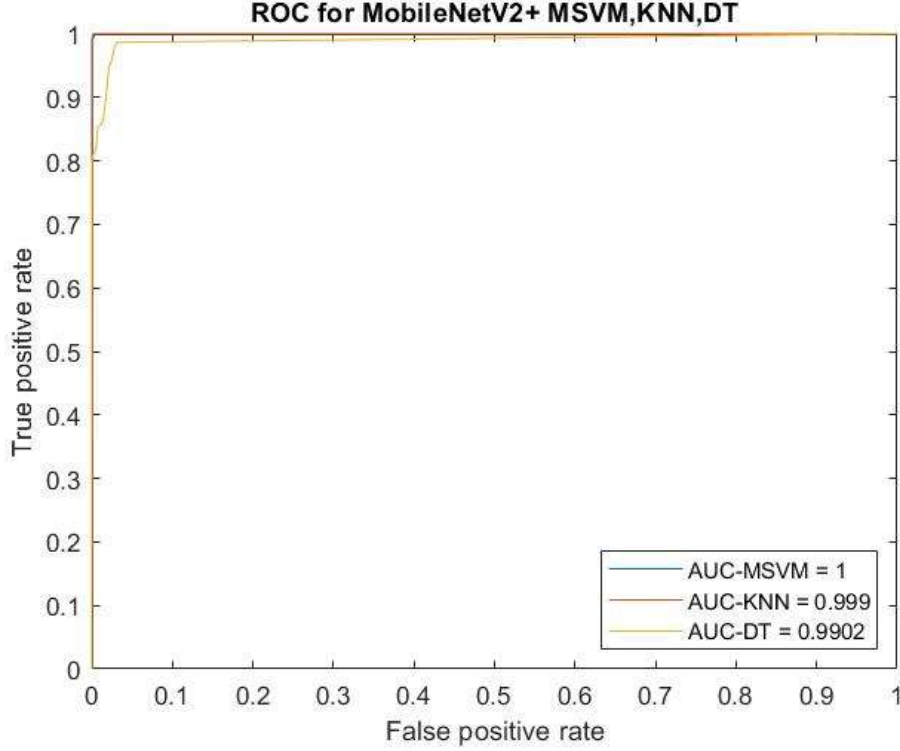
	ACC	SNS	SPC	PRC	F1_score
Öfke	0.999164	0.993355	1	1	0.996667
Dikkat	0.992057	0.951613	0.998079	0.986622	0.968801
İğrenme	0.995401	0.996552	0.995243	0.966555	0.981324
Korku	0.995401	0.986486	0.99666	0.976589	0.981513
Mutluluk	0.998328	1	0.998093	0.986622	0.993266
Tarafsızlık	0.995401	0.989796	0.996187	0.973244	0.98145
Üzüntü	0.994983	0.973597	0.998085	0.986622	0.980066
Sürpriz	0.998328	0.986799	1	1	0.993355
Ortalama	0.996133	0.984775	0.997793	0.984532	0.984555

Çizelge 4.26 MobileNetV2+DT için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.969482	0.879195	0.98233	0.876254	0.877722
Dikkat	0.969482	0.879195	0.98233	0.876254	0.877722
İğrenme	0.969064	0.883959	0.980943	0.866221	0.875
Korku	0.986622	0.952542	0.991416	0.939799	0.946128
Mutluluk	0.974916	0.867692	0.991776	0.943144	0.903846
Tarafsızlık	0.978261	0.927336	0.985259	0.896321	0.911565

Üzüntü	0.970318	0.893103	0.980971	0.866221	0.879457
Sürpriz	0.977843	0.904605	0.988506	0.919732	0.912106
Ortalama	0.974498	0.898453	0.985441	0.897993	0.897943

Şekil 4.23, Çizelge 4.24-26'da, sonuçlar MobileNetV2-ECOC-MSVM modelinin MobileNetV2-KNN'den %0,5 ve MobileNetV2-DT'den %9,2 daha iyi performans gösterdiğini ortaya koymaktadır.



Şekil 4.24 MobileNetV2 tabanlı makine öğrenimi yöntemleri için ROC Eğrileri

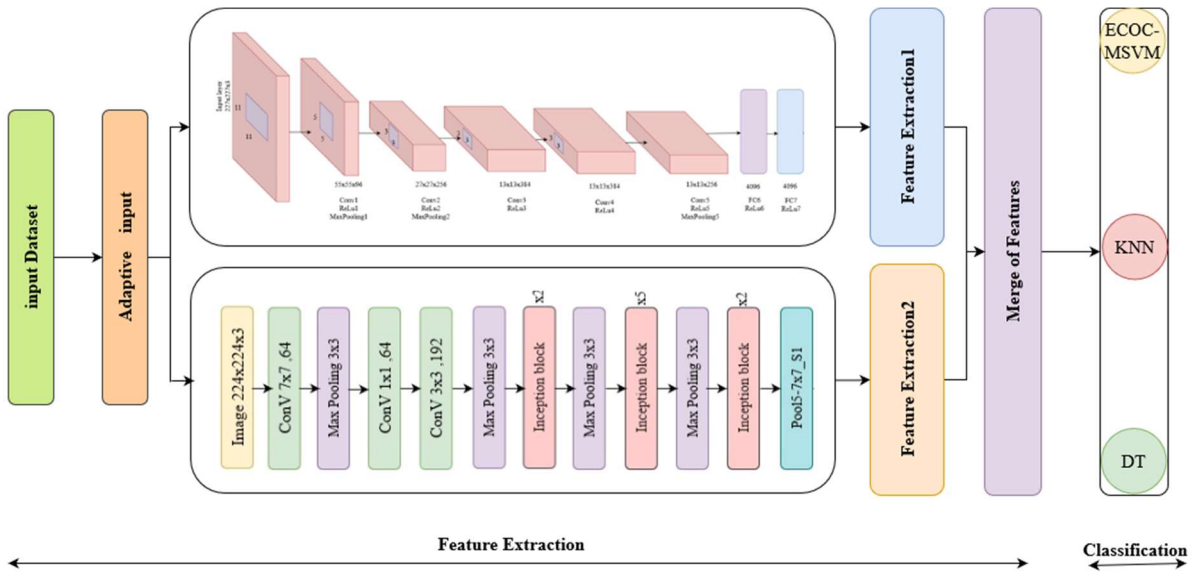
Şekil 4.24 ayrıca ECOC-MSVM sınıflandırıcısının AUC için 1 ve KNN ve DT için ayrı ayrı yaklaşık 0,99 değerini kaydederek ideallığe ulaştığını göstermektedir.

4.2.6. Çoklu CNN ağlarının kombinasyonu ve birleştirme özellikleri

Burada, Çizelge 4.11'e göre seçilen katmanlardan öznitelikler çıkarılmış ve ardından seçilen ağlar için öznitelikler birleştirilmiştir. Daha sonra sınıflandırıcı olarak kullanılan ECOC-MSVM, KNN ve DT makine öğrenme algoritmalarından birine uygulanmaktadır. Çizelge 4.11'de belirtilen katmanlardan özellikler çıkarıldığında, her ağda aşağıdaki katmanlar çıkarılmış demektir: son katmanlar, tam bağlantılı katmanlar ve sınıflandırma katmanı SoftMax. Tam bağlantılı katmanlar, eğitim ve test aşamalarında uzun zaman gerektiren çok sayıda karmaşık hesaplama işlemi gerektiren bir sinir ağıdır.

4.2.6.1. AlexNet ve GoogleNet

Yüzle ilgili sekiz duyguyu (öfke, tiksinti, korku, mutluluk, tarafsızlık, üzüntü, dikkat ve sürpriz) içeren 7968 görüntüden oluşan veri setimiz, görüntülerin boyutunu önceki giriş katmanına uyacak şekilde değiştiren uyarlanabilir bir giriş katmanına sağlanmaktadır (AlexNet ve GoogleNet için sırasıyla 227x227x3 ve 224x224x3). AlexNet ve GoogleNet mimarilerinin birleştirilmesi amacıyla, FC7 katmanındaki özellikler AlexNet'te, Pool5_7x7_S1 katmanındaki özellikler de GoogleNet'te çıkarılır. Ardından bu özellikler birleştirilerek makine öğrenimi algoritmaları olan ECOC-MSVM, KNN ve DT'ye ayrı ayrı uygulanır (Şekil 4.25).



Şekil 4.25 ECOC-MSVM, KNN ve DT ile AlexNet+GoogleNet mimarisi

Confusion Matrix: AlexNet+GoogleNet+MSVM

Anger	291 12.2%	4 0.2%	5 0.2%	0 0.0%	2 0.1%	0 0.0%	0 0.0%	0 0.0%	96.4% 3.6%
Attention	4 0.2%	285 11.9%	4 0.2%	0 0.0%	0 0.0%	2 0.1%	0 0.0%	0 0.0%	96.6% 3.4%
Disgust	4 0.2%	0 0.0%	289 12.1%	0 0.0%	0 0.0%	0 0.0%	1 0.0%	0 0.0%	98.3% 1.7%
Fear	0 0.0%	3 0.1%	0 0.0%	299 12.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.0% 1.0%
Happiness	0 0.0%	0 0.0%	0 0.0%	0 0.0%	297 12.4%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
Neutrality	0 0.0%	7 0.3%	1 0.0%	0 0.0%	0 0.0%	291 12.2%	11 0.5%	4 0.2%	92.7% 7.3%
Sadness	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	6 0.3%	287 12.0%	0 0.0%	98.0% 2.0%
Surprise	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	295 12.3%	100% 0.0%
	97.3% 2.7%	95.3% 4.7%	96.7% 3.3%	100% 0.0%	99.3% 0.7%	97.3% 2.7%	96.0% 4.0%	98.7% 1.3%	97.6% 2.4%
	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	

Target Class

Confusion Matrix: AlexNet+GoogleNet+KNN

Anger	299 12.5%	0 0.0%	2 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.3% 0.7%
Attention	0 0.0%	292 12.2%	0 0.0%	4 0.2%	0 0.0%	5 0.2%	0 0.0%	5 0.2%	95.4% 4.6%
Disgust	0 0.0%	0 0.0%	297 12.4%	0 0.0%	0 0.0%	0 0.0%	6 0.3%	0 0.0%	98.0% 2.0%
Fear	0 0.0%	0 0.0%	0 0.0%	295 12.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
Happiness	0 0.0%	0 0.0%	0 0.0%	0 0.0%	297 12.4%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
Neutrality	0 0.0%	1 0.0%	0 0.0%	0 0.0%	0 0.0%	290 12.1%	2 0.1%	0 0.0%	99.0% 1.0%
Sadness	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	291 12.2%	0 0.0%	100% 0.0%
Surprise	0 0.0%	6 0.3%	0 0.0%	0 0.0%	2 0.1%	4 0.2%	0 0.0%	294 12.3%	96.1% 3.9%
	100% 0.0%	97.7% 2.3%	99.3% 0.7%	98.7% 1.3%	99.3% 0.7%	97.0% 3.0%	97.3% 2.7%	98.3% 1.7%	98.5% 1.5%
	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	

Target Class

Confusion Matrix: AlexNet+GoogleNet+Decision Tree

Anger	269 11.2%	2 0.1%	2 0.1%	0 0.0%	2 0.1%	3 0.1%	2 0.1%	3 0.1%	95.1% 4.9%
Attention	0 0.0%	275 11.5%	4 0.2%	0 0.0%	9 0.4%	17 0.7%	3 0.1%	3 0.1%	88.4% 11.6%
Disgust	6 0.3%	7 0.3%	277 11.6%	3 0.1%	4 0.2%	3 0.1%	13 0.5%	8 0.3%	86.3% 13.7%
Fear	5 0.2%	5 0.2%	2 0.1%	290 12.1%	11 0.5%	1 0.0%	6 0.3%	4 0.2%	89.5% 10.5%
Happiness	10 0.4%	1 0.0%	6 0.3%	0 0.0%	269 11.2%	2 0.1%	2 0.1%	6 0.3%	90.9% 9.1%
Neutrality	0 0.0%	6 0.3%	0 0.0%	0 0.0%	0 0.0%	264 11.0%	4 0.2%	5 0.2%	94.6% 5.4%
Sadness	8 0.3%	3 0.1%	7 0.3%	6 0.3%	2 0.1%	6 0.3%	267 11.2%	2 0.1%	88.7% 11.3%
Surprise	1 0.0%	0 0.0%	1 0.0%	0 0.0%	2 0.1%	3 0.1%	2 0.1%	268 11.2%	96.8% 3.2%
	90.0% 10.0%	92.0% 8.0%	92.6% 7.4%	97.0% 3.0%	90.0% 10.0%	88.3% 11.7%	89.3% 10.7%	89.6% 10.4%	91.1% 8.9%
	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	

Target Class

Şekil 4.26 AlexNet+GoogleNet tabanlı ECOC-MSVM, KNN ve DT için konfüzyon matrisleri

Çizelge 4.27 AlexNet+GoogleNet+ECOC-MSVM için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.992057	0.963576	0.996172	0.973244	0.968386
Dikkat	0.989967	0.966102	0.993324	0.953177	0.959596
İğrenme	0.993729	0.982993	0.995234	0.966555	0.974705
Korku	0.998746	0.990066	1	1	0.995008
Mutluluk	0.999164	1	0.999045	0.993311	0.996644
Tarafsızlık	0.98704	0.926752	0.99615	0.973244	0.949429

Üzüntü	0.992475	0.979522	0.994283	0.959866	0.969595
Sürpriz	0.998328	1	0.998093	0.986622	0.993266
Ortalama	0.993938	0.976126	0.996538	0.975753	0.975829

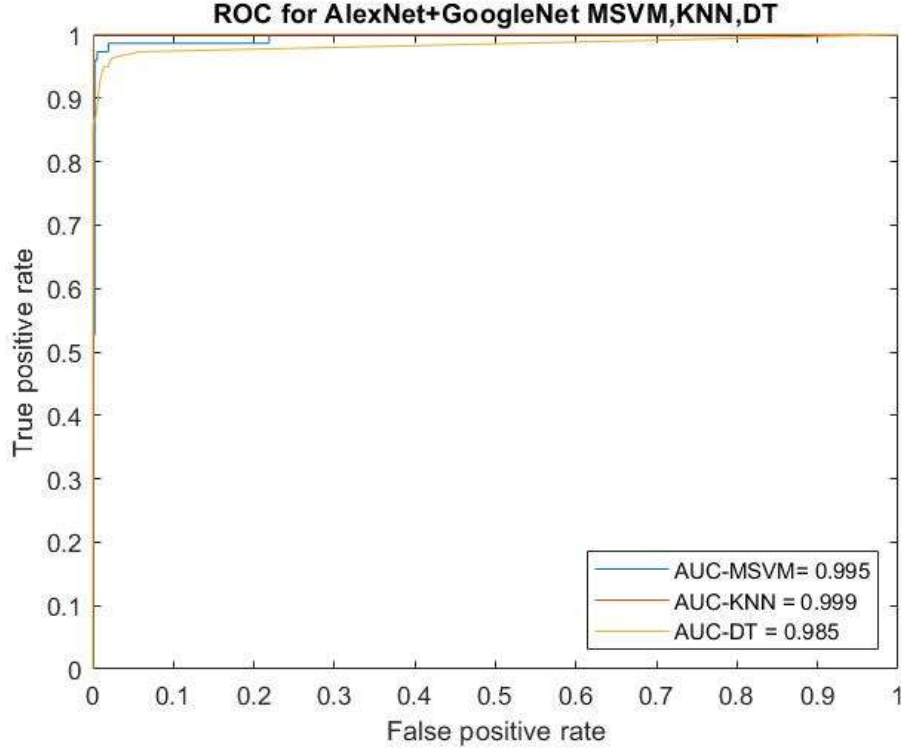
Çizelge 4.28 AlexNet+GoogleNet+KNN için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.999164	0.993355	1	1	0.996667
Dikkat	0.991221	0.954248	0.996644	0.976589	0.965289
İğrenme	0.996656	0.980198	0.999043	0.993311	0.986711
Korku	0.998328	1	0.998093	0.986622	0.993266
Mutluluk	0.999164	1	0.999045	0.993311	0.996644
Tarafsızlık	0.994983	0.989761	0.995712	0.9699	0.97973
Üzüntü	0.996656	1	0.996192	0.973244	0.986441
Sürpriz	0.992893	0.960784	0.997603	0.983278	0.971901
Ortalama	0.996133	0.984793	0.997792	0.984532	0.984581

Çizelge 4.29 AlexNet+GoogleNet+DT için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.981605	0.95053	0.985775	0.899666	0.924399
Dikkat	0.974916	0.884244	0.988467	0.919732	0.901639
İğrenme	0.972408	0.862928	0.989377	0.926421	0.893548
Korku	0.982023	0.895062	0.995648	0.9699	0.930979
Mutluluk	0.976171	0.908784	0.985687	0.899666	0.904202
Tarafsızlık	0.979097	0.946237	0.983436	0.882943	0.913495
Üzüntü	0.972408	0.887043	0.984696	0.892977	0.89
Sürpriz	0.983278	0.967509	0.985343	0.896321	0.930556
Ortalama	0.977738	0.912792	0.987304	0.910953	0.911102

Birleştirilmiş AlexNet+GoogleNet ağında, KNN'nin ECOC-MSVM'den %0,9 ve DT'den de %7,4 daha üstün olduğu açıktır. Bu, Şekil 4.26 ve Çizelge 4.27-29'da görülmektedir.

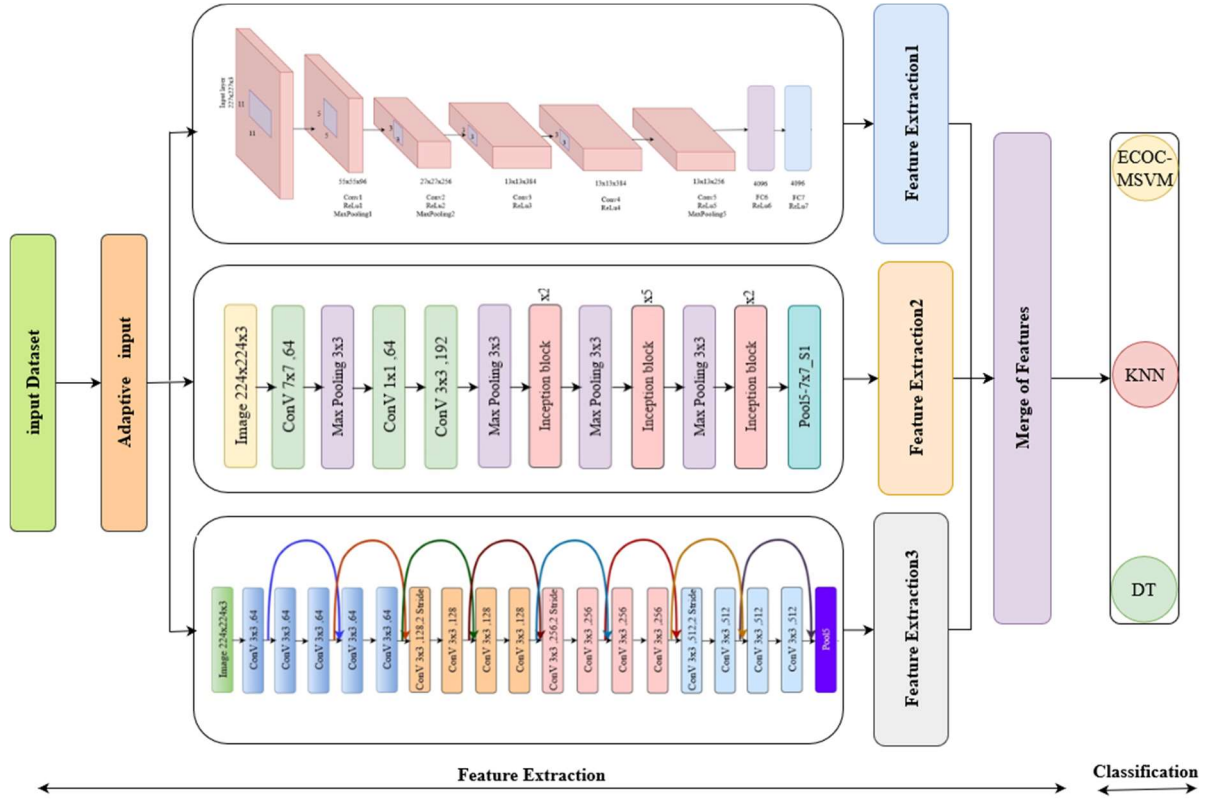


Şekil 4.27 AlexNet+GoogleNet tabanlı makine öğrenimi yöntemleri için ROC Eğrileri

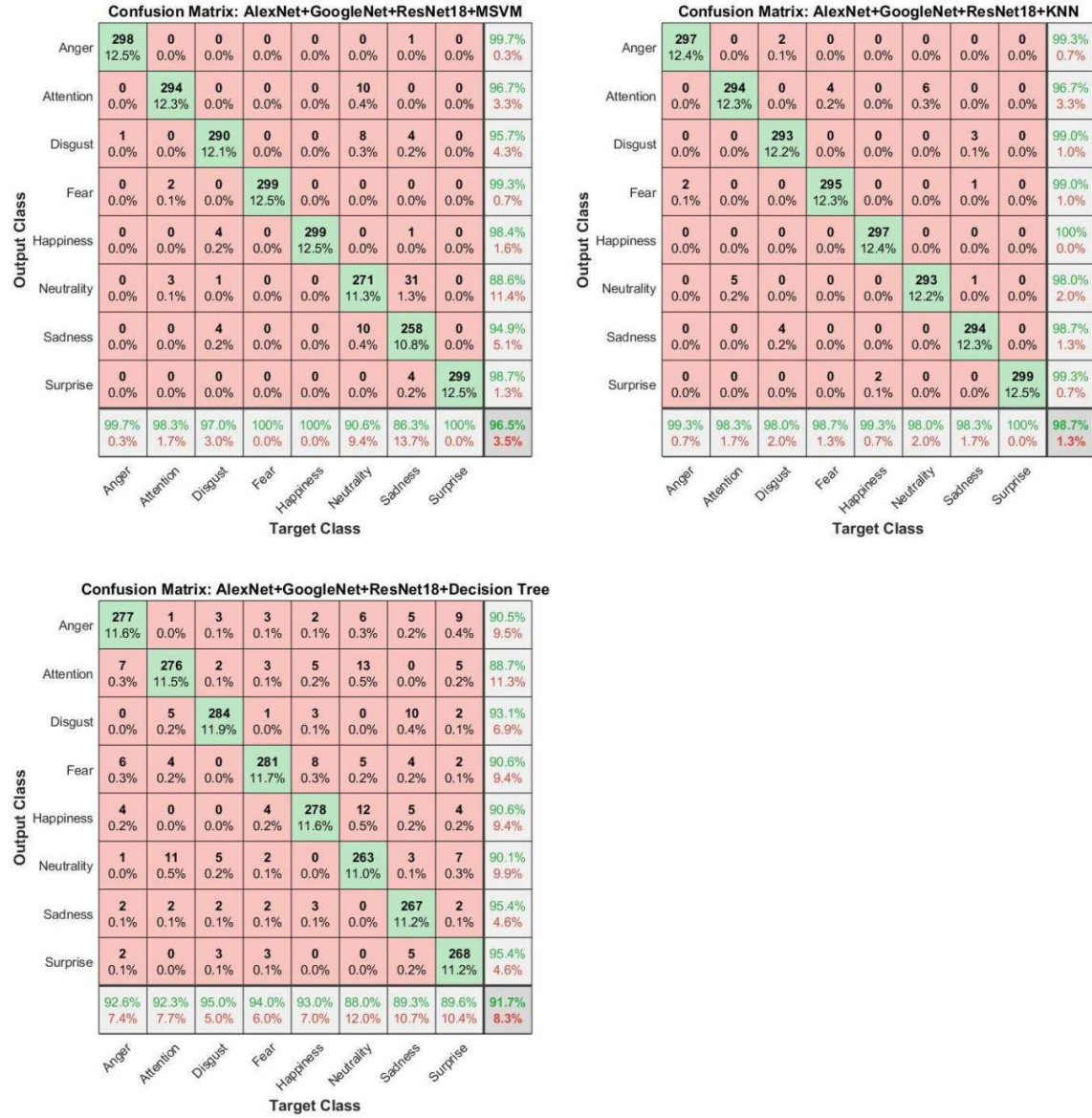
Şekil 4.27, ECOC-MSVM ve KNN sınıflandırıcılarının AUC değerleri için ayrı ayrı 0,99'a ulaştığını ve bu nedenle ideale yakın olduklarını göstermektedir. Ayrıca, DT sınıflandırıcısının da 0,98 AUC değeriyle ideale ulaştığı görülmektedir.

4.2.6.2. AlexNet, GoogleNet ve ResNet18

Yüzle ilgili sekiz duyguyu (öfke, tiksinti, korku, mutluluk, tarafsızlık, üzüntü, dikkat ve sürpriz) içeren ve 7968 görüntüden oluşan veri setimiz, görüntülerin boyutunu önceki giriş katmanına uyacak şekilde değiştiren uyarlanabilir bir giriş katmanına benzer olarak sağlanmaktadır (AlexNet, GoogleNet ve ResNet18 için sırasıyla 227x227x3, 224x224x3 ve 224x224x3). AlexNet, GoogleNet ve ResNet18 mimarilerini birleştirmek amacıyla, AlexNet'in FC7 katmanındaki özellikler, GoogleNet'in Pool5_7x7_S1 katmanındaki özellikler ve ResNet18'in Pool5 katmanındaki özellikler çıkarılır. Ardından bu özellikler birleştirilerek çalışmada kullanılan makine öğrenimi algoritmalarına (ECOC-MSVM, KNN ve DT) uygulanır (Şekil 4.28).



Şekil 4.28 ECOC-MSVM, KNN ve DT ile AlexNet+GoogleNet+ResNet18 mimarisi



Şekil 4.29 AlexNet+GoogleNet+ResNet18 tabanlı ECOC-MSVM, KNN ve DT için konfüzyon matrisleri

Çizelge 4.30 AlexNet+GoogleNet+ResNet18+ECOC-MSVM için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.999164	0.996656	0.999522	0.996656	0.996656
Dikkat	0.993729	0.967105	0.997605	0.983278	0.975124
İğrenme	0.990803	0.957096	0.995692	0.9699	0.963455
Korku	0.999164	0.993355	1	1	0.996667
Mutluluk	0.99791	0.983553	1	1	0.991708
Tarafsızlık	0.973662	0.885621	0.986577	0.906355	0.895868

Üzüntü	0.977007	0.948529	0.98066	0.862876	0.903678
Sürpriz	0.998328	0.986799	1	1	0.993355
Ortalama	0.991221	0.964839	0.995007	0.964883	0.964564

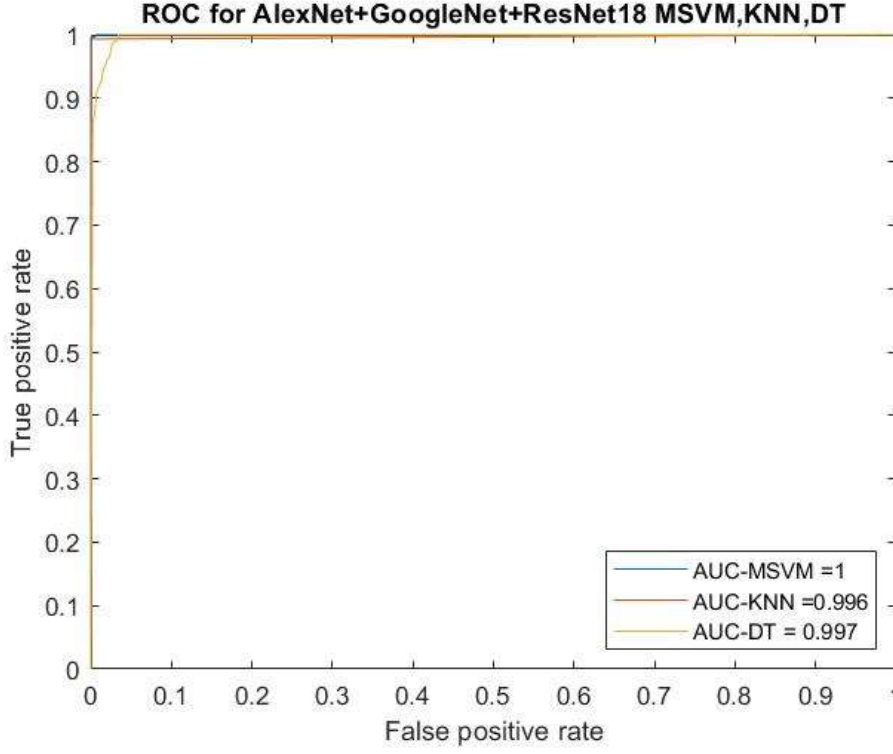
Çizelge 4.31 AlexNet+GoogleNet+ResNet18+KNN için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.998328	0.993311	0.999044	0.993311	0.993311
Dikkat	0.993729	0.967105	0.997605	0.983278	0.975124
İğrenme	0.996237	0.989865	0.997137	0.979933	0.984874
Korku	0.997074	0.989933	0.99809	0.986622	0.988275
Mutluluk	0.999164	1	0.999045	0.993311	0.996644
Tarafsızlık	0.994983	0.979933	0.997133	0.979933	0.979933
Üzüntü	0.996237	0.986577	0.997612	0.983278	0.984925
Sürpriz	0.999164	0.993355	1	1	0.996667
Ortalama	0.996865	0.98751	0.998208	0.987458	0.987469

Çizelge 4.32 AlexNet+GoogleNet+ResNet18+DT için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.978679	0.905229	0.989454	0.926421	0.915702
Dikkat	0.975753	0.88746	0.988948	0.923077	0.904918
İğrenme	0.98495	0.931148	0.992813	0.949833	0.940397
Korku	0.980351	0.906452	0.991354	0.939799	0.922824
Mutluluk	0.979097	0.905537	0.989928	0.929766	0.917492
Tarafsızlık	0.972826	0.900685	0.982857	0.879599	0.890017
Üzüntü	0.981187	0.953571	0.984848	0.892977	0.92228
Sürpriz	0.981605	0.953737	0.985315	0.896321	0.924138
Ortalama	0.979306	0.917977	0.98819	0.917224	0.917221

Şekil 4.29 ve Çizelge 4.30-32'den KNN sınıflandırıcısının ECOC-MSVM'den %2,2 ve DT'den de %7 üstün olduğu görülmektedir.

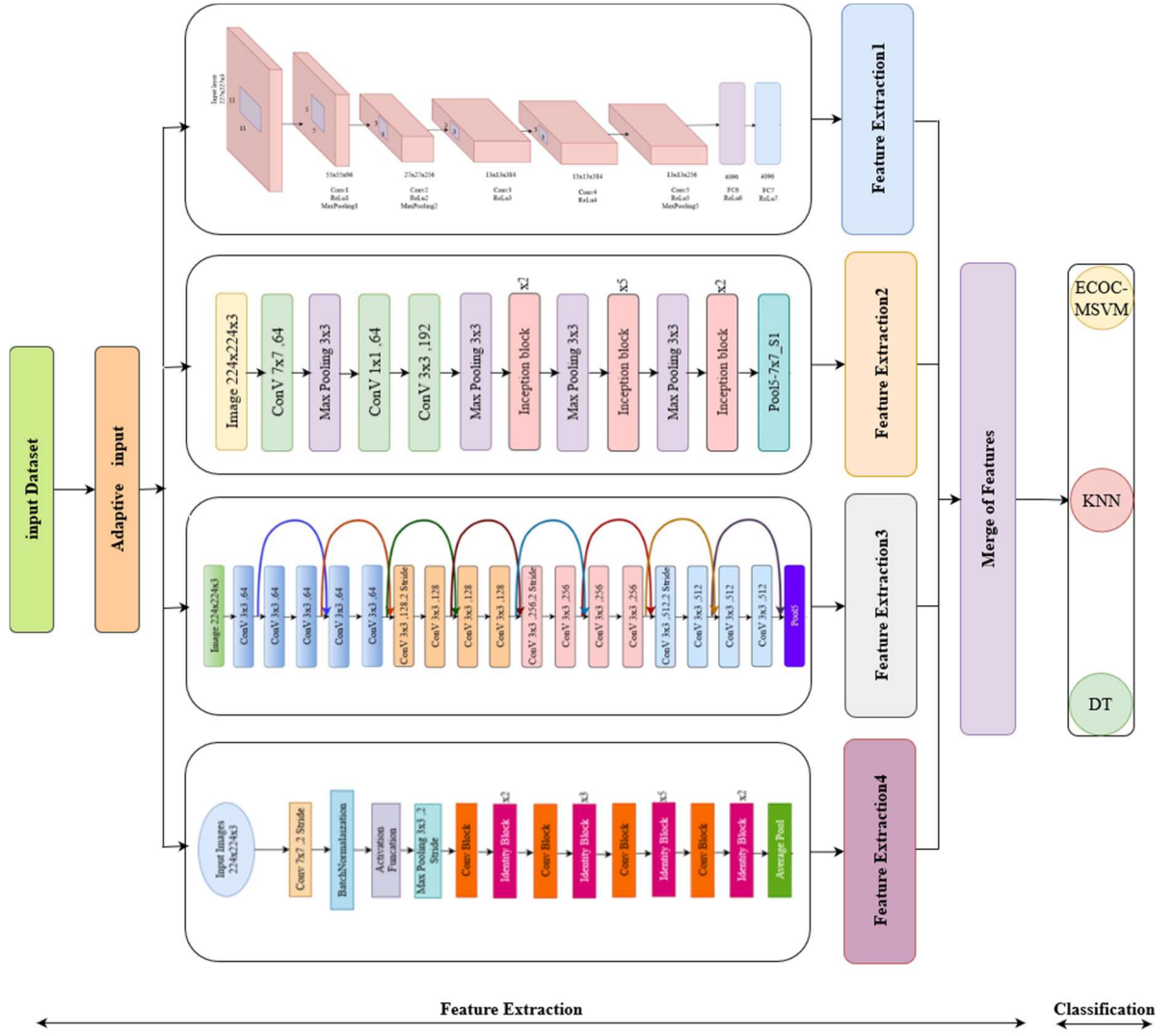


Şekil 4.30 AlexNet+GoogleNet+ResNet18 tabanlı makine öğrenimi yöntemleri için ROC eğrileri

Şekil 4.30'da, ECOC-MSVM sınıflandırıcısı 1 AUC değerine sahipken, KNN ve DT için yaklaşık olarak 0,99 AUC değerleriyle ideallığe yakın oldukları görülmektedir.

4.2.6.3. AlexNet, GoogleNet, ResNet18 ve ResNet50

Burada ayrıca, sekiz yüz duygusuna (öfke, tiksinti, korku, mutluluk, tarafsızlık, üzüntü, dikkat ve sürpriz) sahip ve 7968 görüntüden oluşan veri setimiz, görüntülerin boyutunu giriş katmanına uyacak şekilde değiştiren uyarlanabilir bir giriş katmanına sağlanmaktadır (AlexNet, GoogleNet, ResNet18 ve ResNet50 için sırasıyla 227x227x3, 224x224x3, 224x224x3 ve 224x224x3). Bu amaçla, FC7 katmanındaki özellikler AlexNet'te, Pool5_7x7_S1 katmanındaki özellikler GoogleNet'te, Pool5 katmanındaki özellikler ResNet18'de ve Average Pool katmanındaki özellikler de ResNet50 mimarisinde çıkarılır. Ardından bu özellikler birleştirilerek sınıflandırıcı olarak kullanılan makine öğrenimi algoritmalarından ECOC-MSVM, KNN ve DT'ye ayrı ayrı uygulanmaktadır (Şekil 4.31'e bakınız).



Şekil 4.31 ECOC-MSVM, KNN ve DT ile AlexNet+GoogleNet+ResNet18+ResNet50 mimarisi

Confusion Matrix: AlexNet+GoogleNet+ResNet18+ResNet50+MSVM

Anger	298 12.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
Attention	0 0.0%	291 12.2%	0 0.0%	0 0.0%	0 0.0%	2 0.1%	0 0.0%	0 0.0%	99.3% 0.7%
Disgust	1 0.0%	0 0.0%	299 12.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.7% 0.3%
Fear	0 0.0%	2 0.1%	0 0.0%	291 12.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.3% 0.7%
Happiness	0 0.0%	0 0.0%	0 0.0%	0 0.0%	299 12.5%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
Neutrality	0 0.0%	6 0.3%	0 0.0%	8 0.3%	0 0.0%	297 12.4%	3 0.1%	0 0.0%	94.6% 5.4%
Sadness	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	296 12.4%	0 0.0%	100% 0.0%
Surprise	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	299 12.5%	100% 0.0%
	99.7% 0.3%	97.3% 2.7%	100% 0.0%	97.3% 2.7%	100% 0.0%	99.3% 0.7%	99.0% 1.0%	100% 0.0%	99.1% 0.9%
	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	

Target Class

Confusion Matrix: AlexNet+GoogleNet+ResNet18+ResNet50+KN

Anger	299 12.5%	0 0.0%	2 0.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.3% 0.7%
Attention	0 0.0%	299 12.5%	0 0.0%	2 0.1%	0 0.0%	7 0.3%	0 0.0%	0 0.0%	97.1% 2.9%
Disgust	0 0.0%	0 0.0%	292 12.2%	0 0.0%	0 0.0%	0 0.0%	7 0.3%	0 0.0%	97.7% 2.3%
Fear	0 0.0%	0 0.0%	0 0.0%	297 12.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
Happiness	0 0.0%	0 0.0%	0 0.0%	0 0.0%	299 12.5%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
Neutrality	0 0.0%	0 0.0%	1 0.0%	0 0.0%	0 0.0%	292 12.2%	2 0.1%	0 0.0%	99.0% 1.0%
Sadness	0 0.0%	0 0.0%	4 0.2%	0 0.0%	0 0.0%	0 0.0%	290 12.1%	0 0.0%	98.6% 1.4%
Surprise	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	299 12.5%	100% 0.0%
	100% 0.0%	100% 0.0%	97.7% 2.3%	99.3% 0.7%	100% 0.0%	97.7% 2.3%	97.0% 3.0%	100% 0.0%	99.0% 1.0%
	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	

Target Class

Confusion Matrix: AlexNet+GoogleNet+ResNet18+ResNet50+Decision¹

Anger	272 11.4%	2 0.1%	3 0.1%	5 0.2%	3 0.1%	7 0.3%	9 0.4%	3 0.1%	89.5% 10.5%
Attention	2 0.1%	268 11.2%	5 0.2%	6 0.3%	1 0.0%	8 0.3%	1 0.0%	3 0.1%	91.2% 8.8%
Disgust	7 0.3%	5 0.2%	276 11.5%	0 0.0%	6 0.3%	2 0.1%	4 0.2%	9 0.4%	89.3% 10.7%
Fear	7 0.3%	6 0.3%	8 0.3%	270 11.3%	6 0.3%	0 0.0%	3 0.1%	5 0.2%	88.5% 11.5%
Happiness	0 0.0%	0 0.0%	1 0.0%	4 0.2%	277 11.6%	5 0.2%	0 0.0%	4 0.2%	95.2% 4.8%
Neutrality	3 0.1%	6 0.3%	1 0.0%	2 0.1%	2 0.1%	270 11.3%	6 0.3%	0 0.0%	93.1% 6.9%
Sadness	6 0.3%	10 0.4%	1 0.0%	3 0.1%	4 0.2%	2 0.1%	273 11.4%	3 0.1%	90.4% 9.6%
Surprise	2 0.1%	2 0.1%	4 0.2%	9 0.4%	0 0.0%	5 0.2%	3 0.1%	272 11.4%	91.6% 8.4%
	91.0% 9.0%	89.6% 10.4%	92.3% 7.7%	90.3% 9.7%	92.6% 7.4%	90.3% 9.7%	91.3% 8.7%	91.0% 9.0%	91.1% 8.9%
	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	

Target Class

Şekil 4.32 AlexNet+GoogleNet+ResNet18+ResNet50 tabanlı ECOC-MSVM, KNN ve DT için konfüzyon matrisleri

Çizelge 4.33 AlexNet+GoogleNet+ResNet18+ResNet50+ECOC-MSVM için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.999582	1	0.999522	0.996656	0.998325
Dikkat	0.995819	0.993174	0.996189	0.973244	0.983108
İğrenme	0.999582	0.996667	1	1	0.998331
Korku	0.995819	0.993174	0.996189	0.973244	0.983108
Mutluluk	1	1	1	1	1

Tarafsızlık	0.992057	0.94586	0.999038	0.993311	0.969005
Üzüntü	0.998746	1	0.998569	0.989967	0.994958
Sürpriz	1	1	1	1	1
Ortalama	0.997701	0.991109	0.998688	0.990803	0.990854

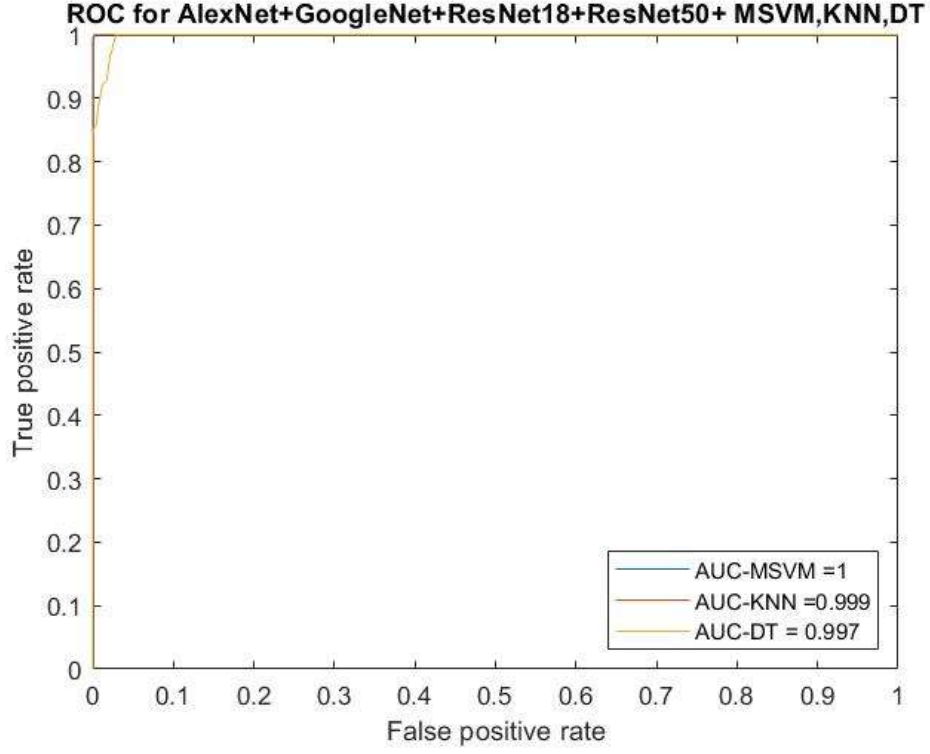
Çizelge 4.34 AlexNet+GoogleNet+ResNet18+ResNet50+KNN için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.999164	0.993355	1	1	0.996667
Dikkat	0.996237	0.970779	1	1	0.985173
İğrenme	0.994147	0.976589	0.996656	0.976589	0.976589
Korku	0.999164	1	0.999045	0.993311	0.996644
Mutluluk	1	1	1	1	1
Tarafsızlık	0.995819	0.989831	0.996662	0.976589	0.983165
Üzüntü	0.994565	0.986395	0.99571	0.9699	0.978078
Sürpriz	1	1	1	1	1
Ortalama	0.997387	0.989619	0.998509	0.989548	0.989539

Çizelge 4.35 AlexNet+GoogleNet+ResNet18+ResNet50+DT için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.975334	0.894737	0.987069	0.909699	0.902156
Dikkat	0.976171	0.911565	0.985224	0.896321	0.903879
İğrenme	0.976589	0.893204	0.988958	0.923077	0.907895
Korku	0.973244	0.885246	0.986104	0.90301	0.89404
Mutluluk	0.98495	0.95189	0.989529	0.926421	0.938983
Tarafsızlık	0.979515	0.931034	0.986204	0.90301	0.916808
Üzüntü	0.977007	0.903974	0.98756	0.913043	0.908486
Sürpriz	0.978261	0.915825	0.987112	0.909699	0.912752
Ortalama	0.977634	0.910934	0.98722	0.910535	0.910625

Şekil 4.32 ve Çizelge 4.33-35, ECOC-MSVM sınıflandırıcısının KNN sınıflandırıcısından %0,1 ve DT sınıflandırıcısından da %8 daha iyi performans gösterdiğini sunmaktadır.

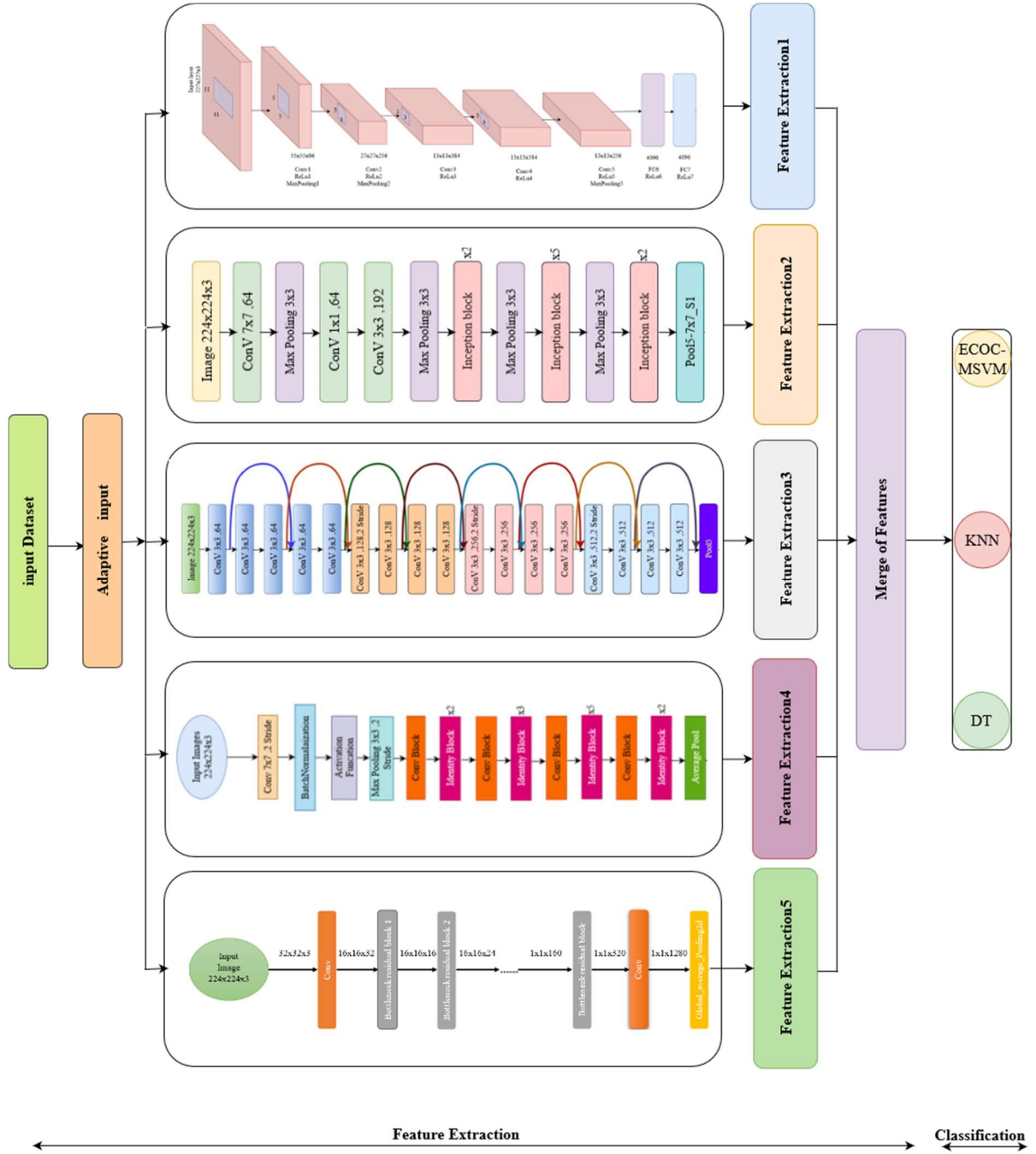


Şekil 4.33 AlexNet+GoogleNet+ResNet18+ResNet50 tabanlı makine öğrenimi yöntemleri için ROC eğrileri

Şekil 4.33 ayrıca ECOC-MSVM sınıflandırıcısının AUC değeri için 1 ve KNN ve DT için de ayrı ayrı yaklaşık 0,99 değerine erişerek ideallığe ulaştığını göstermektedir.

4.2.6.4 AlexNet, GoogleNet, ResNet18, ResNet50 ve MobileNetV2

Burada da, sekiz yüz duygusunu (öfke, tiksinti, korku, mutluluk, tarafsızlık, üzüntü, dikkat ve sürpriz) içeren 7968 görüntüden oluşan veri setimiz, görüntülerin boyutunu giriş katmanına uyacak şekilde değiştiren uyarlanabilir bir giriş katmanına sağlanmaktadır (AlexNet, GoogleNet, ResNet18, ResNet50 ve MobileNetV2 için sırasıyla 227x227x3, 224x224x3, 224x224x3, 224x224x3 ve 224x224x3). Burada AlexNet, GoogleNet, ResNet18, ResNet50 ve MobileNetV2 mimarileri için sırasıyla FC7, Pool5_7x7_S1, Pool5, Average Pool ve Global_average_pooling2d katmanlarından özellikler çıkarılmaktadır. Ardından bu özellikler birleştirilerek ECOC-MSVM, KNN ve DT'ye ayrı ayrı uygulanmaktadır (Şekil 4.34).



Şekil 4.33 ECOC-MSVM, KNN ve DT ile AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2 mimarisi

AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2+MSVM

Anger	295 12.3%	0 0.0%	0 0.0%	0 0.0%	2 0.1%	0 0.0%	0 0.0%	0 0.0%	99.3% 0.7%
Attention	0 0.0%	291 12.2%	0 0.0%	0 0.0%	0 0.0%	2 0.1%	0 0.0%	0 0.0%	99.3% 0.7%
Disgust	4 0.2%	0 0.0%	297 12.4%	0 0.0%	0 0.0%	0 0.0%	2 0.1%	0 0.0%	98.0% 2.0%
Fear	0 0.0%	4 0.2%	0 0.0%	295 12.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	98.7% 1.3%
Happiness	0 0.0%	0 0.0%	0 0.0%	0 0.0%	297 12.4%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
Neutrality	0 0.0%	4 0.2%	1 0.0%	0 0.0%	0 0.0%	297 12.4%	1 0.0%	0 0.0%	98.0% 2.0%
Sadness	0 0.0%	0 0.0%	1 0.0%	0 0.0%	0 0.0%	0 0.0%	296 12.4%	0 0.0%	99.7% 0.3%
Surprise	0 0.0%	0 0.0%	0 0.0%	4 0.2%	0 0.0%	0 0.0%	0 0.0%	299 12.5%	98.7% 1.3%
	98.7% 1.3%	97.3% 2.7%	99.3% 0.7%	98.7% 1.3%	99.3% 0.7%	99.3% 0.7%	99.0% 1.0%	100% 0.0%	99.0% 1.0%
	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	

Target Class

AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2+KNN

Anger	299 12.5%	0 0.0%	1 0.0%	0 0.0%	0 0.0%	0 0.0%	4 0.2%	0 0.0%	98.4% 1.6%
Attention	0 0.0%	299 12.5%	0 0.0%	3 0.1%	0 0.0%	8 0.3%	0 0.0%	0 0.0%	96.5% 3.5%
Disgust	0 0.0%	0 0.0%	297 12.4%	0 0.0%	0 0.0%	0 0.0%	2 0.1%	0 0.0%	99.3% 0.7%
Fear	0 0.0%	0 0.0%	0 0.0%	296 12.4%	2 0.1%	0 0.0%	1 0.0%	0 0.0%	99.0% 1.0%
Happiness	0 0.0%	0 0.0%	0 0.0%	0 0.0%	297 12.4%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
Neutrality	0 0.0%	0 0.0%	1 0.0%	0 0.0%	0 0.0%	291 12.2%	1 0.0%	0 0.0%	99.3% 0.7%
Sadness	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	291 12.2%	0 0.0%	100% 0.0%
Surprise	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	299 12.5%	100% 0.0%
	100% 0.0%	100% 0.0%	99.3% 0.7%	99.0% 1.0%	99.3% 0.7%	97.3% 2.7%	97.3% 2.7%	100% 0.0%	99.0% 1.0%
	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	

Target Class

AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2+DT

Anger	287 12.0%	2 0.1%	2 0.1%	4 0.2%	6 0.3%	0 0.0%	6 0.3%	2 0.1%	92.9% 7.1%
Attention	2 0.1%	286 12.0%	7 0.3%	4 0.2%	6 0.3%	21 0.9%	14 0.6%	0 0.0%	84.1% 15.9%
Disgust	7 0.3%	2 0.1%	270 11.3%	1 0.0%	3 0.1%	8 0.3%	4 0.2%	2 0.1%	90.9% 9.1%
Fear	1 0.0%	4 0.2%	0 0.0%	281 11.7%	7 0.3%	0 0.0%	1 0.0%	10 0.4%	92.4% 7.6%
Happiness	0 0.0%	0 0.0%	5 0.2%	0 0.0%	268 11.2%	0 0.0%	8 0.3%	2 0.1%	94.7% 5.3%
Neutrality	0 0.0%	4 0.2%	2 0.1%	0 0.0%	3 0.1%	253 10.6%	0 0.0%	9 0.4%	93.4% 6.6%
Sadness	1 0.0%	0 0.0%	7 0.3%	3 0.1%	4 0.2%	11 0.5%	260 10.9%	3 0.1%	90.0% 10.0%
Surprise	1 0.0%	1 0.0%	6 0.3%	6 0.3%	2 0.1%	6 0.3%	6 0.3%	271 11.3%	90.6% 9.4%
	96.0% 4.0%	95.7% 4.3%	90.3% 9.7%	94.0% 6.0%	89.6% 10.4%	84.6% 15.4%	87.0% 13.0%	90.6% 9.4%	91.0% 9.0%
	Anger	Attention	Disgust	Fear	Happiness	Neutrality	Sadness	Surprise	

Target Class

Şekil 4.34 AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2 tabanlı ECOC-MSVM, KNN ve DT için konfüzyon matrisleri

Çizelge 4.36 AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2+ECOC-MSVM için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.997492	0.993266	0.998091	0.986622	0.989933
Dikkat	0.995819	0.993174	0.996189	0.973244	0.983108
İğrenme	0.996656	0.980198	0.999043	0.993311	0.986711
Korku	0.996656	0.986622	0.998089	0.986622	0.986622
Mutluluk	0.999164	1	0.999045	0.993311	0.996644
Tarafsızlık	0.996656	0.980198	0.999043	0.993311	0.986711

Üzüntü	0.998328	0.996633	0.998568	0.989967	0.993289
Sürpriz	0.998328	0.986799	1	1	0.993355
Ortalama	0.997387	0.989611	0.998508	0.989548	0.989547

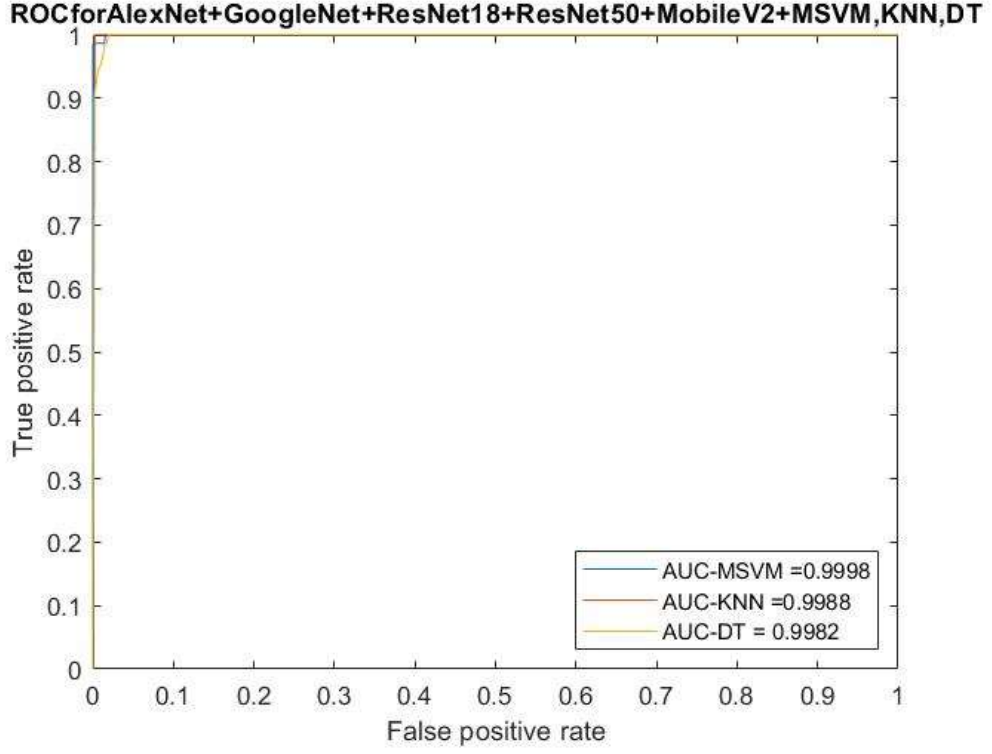
Çizelge 4.37 AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2+KNN için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.99791	0.983553	1	1	0.991708
Dikkat	0.995401	0.964516	1	1	0.981938
İğrenme	0.998328	0.993311	0.999044	0.993311	0.993311
Korku	0.997492	0.989967	0.998567	0.989967	0.989967
Mutluluk	0.999164	1	0.999045	0.993311	0.996644
Tarafsızlık	0.995819	0.993174	0.996189	0.973244	0.983108
Üzüntü	0.996656	1	0.996192	0.973244	0.986441
Sürpriz	1	1	1	1	1
Ortalama	0.997596	0.990565	0.99863	0.990385	0.99039

Çizelge 4.38 AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2+DT için performans ölçümlerinin değerleri

	ACC	SNS	SPC	PRC	F1_score
Öfke	0.985786	0.928803	0.994239	0.959866	0.944079
Dikkat	0.97199	0.841176	0.993665	0.956522	0.895149
İğrenme	0.976589	0.909091	0.986158	0.90301	0.90604
Korku	0.98286	0.924342	0.991379	0.939799	0.932007
Mutluluk	0.980769	0.946996	0.985301	0.896321	0.920962
Tarafsızlık	0.973244	0.933579	0.978312	0.846154	0.887719
Üzüntü	0.971572	0.899654	0.981455	0.869565	0.884354
Sürpriz	0.976589	0.906355	0.986622	0.906355	0.906355
Ortalama	0.977425	0.91125	0.987141	0.909699	0.909583

Şekil 4.34 ve Çizelge 4.36-4.38, ECOC-MSVM sınıflandırıcısının genel F1-skor değeri açısından KNN sınıflandırıcısına eşdeğer olduğunu ve DT sınıflandırıcısını da %8 oranla geride bıraktığını göstermektedir.



Şekil 4.35 AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2 tabanlı makine öğrenimi yöntemleri için ROC Eğrileri

Şekil 4.35, ECOC-MSVM, KNN ve DT sınıflandırıcılarının AUC değerlerinin ayrı ayrı 0.99'a ulaştığını ve dolayısıyla ideallığe yakın olduklarını göstermektedir.

Burada, ayrı ayrı kullanılan ECOC-MSVM, KNN ve DT makine öğrenimi algoritmaları ile önceden eğitilmiş ağlar (AlexNet, GoogleNet, ResNet18, ResNet50 ve MobileNetV2) kullanıldığında birleştirme işleminden önce elde edilen sonuçlar karşılaştırılmaktadır. Bu bağlamda, en iyi sonuçların MobileNetV2+ECOC-MSVM kullanıldığında elde edildiği ve bunların ResNet50+KNN tarafından erişilen sonuçlara neredeyse eşit olduğu görülmektedir. Ayrıca, GoogleNet+KNN tarafından kaydedilen sonuçlara benzer olduğu ortaya görülmektedir. Ek olarak, bu modeller ResNet18+DT tarafından kaydedilen en düşük sonuçtan %9,3 daha iyi performans göstermektedirler.

Önceden eğitilmiş iki, üç, dört veya beş ağın (AlexNet, GoogleNet, ResNet18, ResNet50 ve MobileNetV2) birleştirilmesinden sonra elde edilen sonuçlar ise bu paragrafta karşılaştırılmaktadır. Bu bağlamda, en iyi sonuçların AlexNet+GoogleNet+ResNet18+ResNet50+MobileNetV2 mimarisinde elde edildiği görülmektedir. Burada ECOC-MSVM sınıflandırıcısı kullanıldığında genel F1-Skor değeri için %99 ve KNN sınıflandırıcısı için de %99'dur. Bu bulgular,

AlexNet+GoogleNet+ResNet18+ResNet50 tabanlı ECOC-MSVM ve KNN sınıflandırıcıları ile karşılaştırılabilir. Ayrıca, Çizelge 4.10'da, dört veya beş mimarinin oluşturduğu ağı eğitim ve test aşamalarının süreleri karşılaştırılabilir. Sonuçlar dört ağı içeren mimari tarafından harcanan sürenin beş ağına göre 4-6 dakika daha düşük olduğunu göstermektedir. Dört ağı içeren mimarinin, eğitim ve test için daha az zaman gerektirmesi açısından beş ağı kapsayan modele göre biraz daha avantajlı olduğu ifade edilebilir.

Çizelge 4.10'da verilen ECOC-MSVM ve KNN ile birleştirilmiş ağların performans ölçütleri ve eğitim-test aşamaları için gereken zaman açısından kontrol edildiğinde, daha yüksek sonuçlar elde edildiği ve K-çapraz doğrulama kullanıldığında gereken zamandan daha az olduğu görülmektedir.

Son olarak, makine öğrenimi algoritmalarından (ECOC-MSVM, KNN) birinin kullanımıyla, ağları birleştirme işlemi yüksek sonuçlar ve eğitim-test aşamalarında harcanan süreyi azaltmak için kullanılabilir. Eğitim ve test aşamaları için fazla zaman gerektirmemesinin birinci nedeni, verilerin tüm birleştirilmiş ağlarda paralel olarak akması ve aynı anda çıkarma katmanlarına ulaşmasıdır. Daha sonra çıkarılan özellikler, sınıflandırıcı olarak kullanılan makine öğrenimi algoritmalarından birine (ECOC-MSVM, KNN ve DT) uygulanmak üzere tek bir katmanda birleştirilir. İkinci neden, karmaşık matematiksel hesaplamalar gerektiren ve bu nedenle tamamlanması daha uzun süren CNN ağlarından tam bağlı katmanların çıkarılmasıdır.

5. SONUÇ VE ÖNERİLER

Derin öğrenmeyi kullanan yüz ifadesi tanıma sistemleri, yapay zekânın öne çıkan alanlarından birisidir. Öğrencilerin yüz ifadelerini (öfke, tiksinti, korku, mutluluk, tarafsızlık, üzüntü, dikkat ve sürpriz) otomatik olarak tanıyabilen derin öğrenme tabanlı bu sistem deneysel olarak iki farklı aşamada gerçekleştirilmiştir.

İlkinde, k-katlı çapraz doğrulama ve altı farklı derin öğrenme ağı kullanılmaktadır. Burada, %99,7 doğruluk ve %99 F1 skor değeri ile en yüksek sonuçlara sahip olan ResNet18+TL+KFCV ön plana çıkmaktadır.

İkinci olarak, birkaç derin öğrenme ağından elde edilen özelliklerin birleştirilmesi ve daha sonra bu birleştirilmiş özelliklerin sınıflandırıcı olarak kullanılan makine öğrenimi algoritmalarından (ECOC-MSVM, KNN ve DT) birine uygulanmasıyla en yüksek sonuçlar, AlexNet+GoogleNet+ResNet18+ResNet50+TL+KNN tarafından elde edilmiştir. Bu durumda, doğruluk oranı %99,7 ve F1-Skor değeri ise %99 olmaktadır.

Bu yöntemlerdeki kaydedilen yüksek sonuçlar karşılaştırıldığında, doğruluk ve F1 skor değerleri açısından benzer olduklarını görmekteyiz. Ancak ikinci yöntem, eğitim ve test için gereken süre açısından birinci yöntemi geride bırakmaktadır. Birinci yöntemde, eğitim ve test için 105 dakika (Bkz. Tablo 4.1) gerekmektedirken, ikinci yöntemde sadece 35 dakikaya ihtiyaç duyulmaktadır (Bkz. Tablo 4.10). Ayrıca ikinci yöntemde, eğitim ve test aşamaları için veri kümesinde kullanılan görüntü sayısı 7968 iken, birinci yöntemde bu sayı 6720'dir.

Ancak her iki yöntem de ilgili literatür çalışmalarına göre oldukça etkilidir.

Bu araştırmanın bulguları her iki eğitim türü için de önemli çıktılarına sahiptir. Günümüzde, dünyanın farklı ülkelerindeki çoğu eğitim kurumu, öğrenciler için yüz yüze eğitime alternatif bir seçenek olarak çevrimiçi eğitimi tercih etmeye başlamıştır. Bu çalışma, hem normal eğitim hem de çevrimiçi eğitim için kullanılabilir. Bu sistem, öğrencilerin yüz ifadelerini otomatik ve hızlı bir şekilde tanıyarak, önerilen sistemi alandaki öğretim görevlileri ve araştırmacılar için kullanışlı hale getirir. Öğrencilerin eğitim alanında ki yüz ifadeleri, onların ders deneyimleri hakkında geri bildirim verebilir. Otomatik çevrimiçi yüz tanımaya izin veren kurumlar, çevrimiçi eğitimde önemli pratik uygulamalara sahip olabilir. Bu tür teknikler, öğretim görevlileri tarafından sınıf düzeyinde katılım ve anlayış hakkında geri bildirim toplamak için kullanılabilir. Bu geri bildirimin birçok hususun değerlendirilmesinde faydalı olabileceğine inanmaktayız. Daha geniş ölçekte, öğrencilerin yüz ifadelerine ilişkin

bilgiler, yükseköğretim bakanlıkları, üniversite yöneticileri ve yükseköğretimdeki karar vericiler tarafından yapılan planlamayı daha stratejik bir hale getirebilir. Bir diğer taraftan, bu çalışma mevcut araştırma ve donanımlar nedeniyle bazı sınırlamalara sahiptir. Verilerin işlenmesinde belirli kısıtlamalar vardır ancak, diğer çalışmalarla karşılaştırıldığında, performans metrikleri açısından elde edilen bulgular başarılıdır. Gelecekteki çalışmalarda, kurumdaki tüm üniversite öğrencilerinin yüz ifadelerini verisetine dâhil etmeyi ve öğretim elemanlarına ders sırasında öğrencilerin durumları hakkında geri bildirim sağlamaya yardımcı olmayı amaçlamaktayız. Ayrıca, yüz ifadesi tanımayı, önerilen modeli daha da geliştirmek amacıyla çevrimiçi öğrenme eğitim platformlarına entegre etmeyi hedeflemekteyiz.



KAYNAKLAR

- Sathik, M., Jonathan, S. G. (2013). Effect of facial expressions on student's comprehension recognition in virtual educational environments. *SpringerPlus*, 2(1). <https://doi.org/10.1186/2193-1801-2-455>
- Utami, P., Hartanto, R., Soesanti, I. (2019). A Study on Facial Expression Recognition in Assessing Teaching Skills: Datasets and Methods. *Procedia Computer Science*, 161, 544–552. <https://doi.org/10.1016/j.procs.2019.11.154>
- Pise, A. A., Vadapalli, H., Sanders, I. (2022). Estimation of Learning Affects Experienced by Learners: An Approach Using Relational Reasoning and Adaptive Mapping. *Wireless Communications and Mobile Computing*, 2022, 1–14. <https://doi.org/10.1155/2022/8808283>
- Bouhlal, M., Aarika, K., Abdelouahid, R. A., Elfilali, S., Benlahmar, E. (2020). Emotions recognition as innovative tool for improving students' performance and learning approaches. *Procedia Computer Science*, 175, 597–602. <https://doi.org/10.1016/j.procs.2020.07.086>
- Li, S., Deng, W. (2022). Deep Facial Expression Recognition: A Survey. *IEEE Transactions on Affective Computing*, 13(3), 1195–1215. <https://doi.org/10.1109/taffc.2020.2981446>
- Sarkar, P., Etemad, A. (2022). Self-Supervised ECG Representation Learning for Emotion Recognition. *IEEE Transactions on Affective Computing*, 13(3), 1541–1554. <https://doi.org/10.1109/taffc.2020.3014842>
- Li, Z., Wang, C., Liu, X., Wang, Y. (2021). Facial expression description and recognition based on fuzzy semantic concepts. *Future Generation Computer Systems*, 114, 619–628. <https://doi.org/10.1016/j.future.2020.08.034>
- Maheshwari, D., Ghosh, S., Tripathy, R., Sharma, M., Acharya, U. R. (2021). Automated accurate emotion recognition system using rhythm-specific deep convolutional neural network technique with multi-channel EEG signals. *Computers in Biology and Medicine*, 134, 104428. <https://doi.org/10.1016/j.compbiomed.2021.104428>
- Huang, Q., Huang, C., Wang, X., Jiang, F. (2021). Facial expression recognition with grid-wise attention and visual transformer. *Information Sciences*, 580, 35–54. <https://doi.org/10.1016/j.ins.2021.08.043>

- Wang, Z., Zeng, F., Liu, S., Zeng, B. (2021). OAENet: Oriented attention ensemble for accurate facial expression recognition. *Pattern Recognition*, 112, 107694. <https://doi.org/10.1016/j.patcog.2020.107694>
- Chen, B., Guan, W., Li, P., Ikeda, N., Hirasawa, K., Lu, H. (2021). Residual multi-task learning for facial landmark localization and expression recognition. *Pattern Recognition*, 115, 107893. <https://doi.org/10.1016/j.patcog.2021.107893>
- Ghazouani, H. (2021). A genetic programming-based feature selection and fusion for facial expression recognition. *Applied Soft Computing*, 103, 107173. <https://doi.org/10.1016/j.asoc.2021.107173>
- Barman, A., Dutta, P. (2021). Facial expression recognition using distance and shape signature features. *Pattern Recognition Letters*, 145, 254–261. <https://doi.org/10.1016/j.patrec.2017.06.018>
- Sun, X., Xia, P., Ren, F. (2021). Multi-attention based Deep Neural Network with hybrid features for Dynamic Sequential Facial Expression Recognition. *Neurocomputing*, 444, 378–389. <https://doi.org/10.1016/j.neucom.2019.11.127>
- Tuncer, T., Dogan, S., Subasi, A. (2021). A new fractal pattern feature generation function based emotion recognition method using EEG. *Chaos, Solitons & Fractals*, 144, 110671. <https://doi.org/10.1016/j.chaos.2021.110671>
- Zhou, J., Zhang, X., Lin, Y., Liu, Y. (2021). Facial expression recognition using frequency multiplication network with uniform rectangular features. *Journal of Visual Communication and Image Representation*, 75, 103018. <https://doi.org/10.1016/j.jvcir.2020.103018>
- Liu, T., Wang, J., Yang, B., Wang, X. (2021). Facial expression recognition method with multi-label distribution learning for non-verbal behavior understanding in the classroom. *Infrared Physics & Technology*, 112, 103594. <https://doi.org/10.1016/j.infrared.2020.103594>

- Lee, J. R., Wang, L., Wong, A. (2021). EmotionNet Nano: An Efficient Deep Convolutional Neural Network Design for Real-Time Facial Expression Recognition. *Frontiers in Artificial Intelligence*, 3. <https://doi.org/10.3389/frai.2020.609673>
- Minaee, S., Minaei, M., Abdolrashidi, A. (2021). Deep-Emotion: Facial Expression Recognition Using Attentional Convolutional Network. *Sensors*, 21(9), 3046. <https://doi.org/10.3390/s21093046>
- Uppal, H., Sepas-Moghaddam, A., Greenspan, M., Etemad, A. (2021). Depth as Attention for Face Representation Learning. *IEEE Transactions on Information Forensics and Security*, 16, 2461–2476. <https://doi.org/10.1109/tifs.2021.3053458>
- Altameem, T., Altameem, A. (2020). Facial expression recognition using human machine interaction and multi-modal visualization analysis for healthcare applications. *Image and Vision Computing*, 103, 104044. <https://doi.org/10.1016/j.imavis.2020.104044>
- Sun, X., Xia, P., Zhang, L., Shao, L. (2020). A ROI-guided deep architecture for robust facial expressions recognition. *Information Sciences*, 522, 35–48. <https://doi.org/10.1016/j.ins.2020.02.047>
- Zeng, N., Zhang, H., Song, B., Liu, W., Li, Y., Dobaie, A. M. (2018). Facial expression recognition via learning deep sparse autoencoders. *Neurocomputing*, 273, 643–649. <https://doi.org/10.1016/j.neucom.2017.08.043>
- Sun, W., Zhao, H., Jin, Z. (2018). A visual attention based ROI detection method for facial expression recognition. *Neurocomputing*, 296, 12–22. <https://doi.org/10.1016/j.neucom.2018.03.034>
- Kuo, C. M., Lai, S. H., Sarkis, M. (2018). A Compact Deep Learning Model for Robust Facial Expression Recognition. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. <https://doi.org/10.1109/cvprw.2018.00286>
- Hasani, B., Mahoor, M. H. (2017). Spatio-Temporal Facial Expression Recognition Using Convolutional Neural Networks and Conditional Random Fields. *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*. <https://doi.org/10.1109/fg.2017.99>

- Wang, F., Lv, J., Ying, G., Chen, S., Zhang, C. (2019). Facial expression recognition from image based on hybrid features understanding. *Journal of Visual Communication and Image Representation*, 59, 84–88. <https://doi.org/10.1016/j.jvcir.2018.11.010>
- Chen, J., Lv, Y., Xu, R., Xu, C. (2019). Automatic social signal analysis: Facial expression recognition using difference convolution neural network. *Journal of Parallel and Distributed Computing*, 131, 97–102. <https://doi.org/10.1016/j.jpdc.2019.04.017>
- Rajendra Kurup, A., Ajith, M., Martínez Ramón, M. (2019). Semi-supervised facial expression recognition using reduced spatial features and Deep Belief Networks. *Neurocomputing*, 367, 188–197. <https://doi.org/10.1016/j.neucom.2019.08.029>
- Makhmudkhujaev, F., Abdullah-Al-Wadud, M., Iqbal, M. T. B., Ryu, B., Chae, O. (2019). Facial expression recognition with local prominent directional pattern. *Signal Processing: Image Communication*, 74, 1–12. <https://doi.org/10.1016/j.image.2019.01.002>
- Renda, A., Barsacchi, M., Bechini, A., Marcelloni, F. (2019). Comparing ensemble strategies for deep learning: An application to facial expression recognition. *Expert Systems With Applications*, 136, 1–11. <https://doi.org/10.1016/j.eswa.2019.06.025>
- Ye, Y., Zhang, X., Lin, Y., Wang, H. (2019). Facial expression recognition via region-based convolutional fusion network. *Journal of Visual Communication and Image Representation*, 62, 1–11. <https://doi.org/10.1016/j.jvcir.2019.04.009>
- Sadeghi, H., Raie, A. A. (2019). Histogram distance metric learning for facial expression recognition. *Journal of Visual Communication and Image Representation*, 62, 152–165. <https://doi.org/10.1016/j.jvcir.2019.05.004>
- Xie, W., Jia, X., Shen, L., Yang, M. (2019). Sparse deep feature learning for facial expression recognition. *Pattern Recognition*, 96, 106966. <https://doi.org/10.1016/j.patcog.2019.106966>
- Ahmad Radzi, S., Itqan K, S., N.M, S. (2016). A MATLAB-based Convolutional Neural Network Approach for Face Recognition System. *Journal of Bioinformatics and Proteomics Review*, 2(1), 1–5. <https://doi.org/10.15436/2381-0793.16.009>

- Mayya, V., Pai, R. M., Manohara Pai, M. (2016). Automatic Facial Expression Recognition Using DCNN. *Procedia Computer Science*, 93, 453–461. <https://doi.org/10.1016/j.procs.2016.07.233>
- Fathallah, A., Abdi, L., Douik, A. (2017). Facial Expression Recognition via Deep Learning. *IEEE Conference Publication | IEEE Xplore*, (IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)), 745–750. <https://doi.org/10.1109/AICCSA.2017.124>
- Sang, D. V., Van Dat, N., Thuan, D. P. (2017). Facial expression recognition using deep convolutional neural networks. 2017 9th International Conference on Knowledge and Systems Engineering (KSE). <https://doi.org/10.1109/kse.2017.8119447>
- Arya, S., Agrawal, A., Indore, I. (2018). Face recognition with partial face recognition and convolutional neural network. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, 7(1), 91-94.
- Sharma, M., Anuradha, J., Manne, H. K., Kashyap, G. S. C. (2017). Facial detection using deep learning. IOP Conference Series: *Materials Science and Engineering*, 263, 042092. <https://doi.org/10.1088/1757-899x/263/4/042092>
- Li, J., Lam, E. Y. (2015). Facial expression recognition using deep neural networks. 2015 *IEEE International Conference on Imaging Systems and Techniques (IST)*. <https://doi.org/10.1109/ist.2015.7294547>
- Hussein, H., Angelini, F., Naqvi, M., Chambers, J. A. (2018). Deep-Learning Based Facial Expression Recognition System Evaluated on Three Spontaneous Databases. 2018 9th *International Symposium on Signal, Image, Video and Communications (ISIVC)*. <https://doi.org/10.1109/isivc.2018.8709224>
- Prasada Rao, K., Chandra Sekhara Rao, M., Hemanth Chowdary, N. (2019). An integrated approach to emotion recognition and gender classification. *Journal of Visual Communication and Image Representation*, 60, 339–345. <https://doi.org/10.1016/j.jvcir.2019.03.002>
- Tarnowski, P., Kołodziej, M., Majkowski, A., Rak, R. J. (2017). Emotion recognition using facial expressions. *Procedia Computer Science*, 108, 1175–1184. <https://doi.org/10.1016/j.procs.2017.05.025>

- Qayyum, H., Majid, M., Anwar, S. M., Khan, B. (2017). Facial expression recognition using stationary wavelet transform features. *Mathematical Problems in Engineering*. 2017: 1–9. <https://doi.org/10.1155/2017/9854050>
- Carcagnì, P., Del Coco, M., Leo, M., Distanto, C. (2015). Facial expression recognition and histograms of oriented gradients: a comprehensive study. *SpringerPlus*, 4(1). <https://doi.org/10.1186/s40064-015-1427-3>
- Waqas, U., Choe, T. Y. (2018). Good Portrait Selection Based on Deep Learning Using Facial Expressions. *Journal of Computers* .Vol. 29 No. 6, 2018, pp. 132-139 <https://doi.org/10.3966/199115992018122906012>
- Leo, M., Del Coco, M., Carcagnì, P., Distanto, C., Bernava, M., Pioggia, G., Palestra, G. (2015). Automatic emotion recognition in robot-children interaction for ASD treatment. In *Proceedings of the IEEE International Conference on Computer Vision Workshops* pp. 537-545, <https://doi.org/10.1109/ICCVW.2015.76>.
- Lopes, A. T., de Aguiar, E., De Souza, A. F., Oliveira-Santos, T. (2017). Facial expression recognition with Convolutional Neural Networks: Coping with few data and the training sample order. *Pattern Recognition*, 61, 610–628. <https://doi.org/10.1016/j.patcog.2016.07.026>
- Sajjad, M., Nasir, M., Ullah, F. U. M., Muhammad, K., Sangaiah, A. K., Baik, S. W. (2019). Raspberry Pi assisted facial expression recognition framework for smart security in law-enforcement services. *Information Sciences*, 479, 416–431. <https://doi.org/10.1016/j.ins.2018.07.027>
- Chen, L., Zhou, M., Su, W., Wu, M., She, J., Hirota, K. (2018). Softmax regression based deep sparse autoencoder network for facial emotion recognition in human-robot interaction. *Information Sciences*, 428, 49–61. <https://doi.org/10.1016/j.ins.2017.10.044>
- Surace, L., Patacchiola, M., Battini Sönmez, E., Spataro, W., Cangelosi, A. (2017). Emotion recognition in the wild using deep neural networks and Bayesian classifiers. *Proceedings of the 19th ACM International Conference on Multimodal Interaction*. <https://doi.org/10.1145/3136755.3143015>
- Wu, M., Su, W., Chen, L., Liu, Z., Cao, W., Hirota, K. (2019). Weight-Adapted Convolution Neural Network for Facial Expression Recognition in Human-Robot Interaction. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 1–12.

<https://doi.org/10.1109/tsmc.2019.2897330>

- Sadeghi, H., Raie, A. A. (2022). HistNet: Histogram-based convolutional neural network with Chi-squared deep metric learning for facial expression recognition. *Information Sciences*, 608, 472–488. <https://doi.org/10.1016/j.ins.2022.06.092>
- Zhang, J., Yu, H. (2022). Improving the Facial Expression Recognition and Its Interpretability via Generating Expression Pattern-map. *Pattern Recognition*, 129, 108737. <https://doi.org/10.1016/j.patcog.2022.108737>
- Ruan, L., Han, Y., Sun, J., Chen, Q., Li, J. (2022). Facial expression recognition in facial occlusion scenarios: A path selection multi-network. *Displays*, 74, 102245. <https://doi.org/10.1016/j.displa.2022.102245>
- Mahmoudi, M. A., Chetouani, A., Boufera, F., Tabia, H. (2022). Kernel-based convolution expansion for facial expression recognition. *Pattern Recognition Letters*, 160, 128–134. <https://doi.org/10.1016/j.patrec.2022.06.013>
- Ge, H., Zhu, Z., Dai, Y., Wang, B., Wu, X. (2022). Facial expression recognition based on deep learning. *Computer Methods and Programs in Biomedicine*, 215, 106621. <https://doi.org/10.1016/j.cmpb.2022.106621>
- Maithri, M., Raghavendra, U., Gudigar, A., Samanth, J., Datta Barua, P., Murugappan, M., . . . Acharya, U. R. (2022). Automated emotion recognition: Current trends and future perspectives. *Computer Methods and Programs in Biomedicine*, 215, 106646. <https://doi.org/10.1016/j.cmpb.2022.106646>
- Ma, X., Ma, Y. (2022). Relation and context augmentation network for facial expression recognition. *Image and Vision Computing*, 127, 104556. <https://doi.org/10.1016/j.imavis.2022.104556>
- Albraikan, A. A., Alzahrani, J. S., Alshahrani, R., Yafoz, A., Alsini, R., Hilal, A. M., . . . Gupta, D. (2022). Intelligent facial expression recognition and classification using optimal deep transfer learning model. *Image and Vision Computing*, 128, 104583. <https://doi.org/10.1016/j.imavis.2022.104583>
- Guo, Y., Huang, J., Xiong, M., Wang, Z., Hu, X., Wang, J., Hijji, M. (2022). Facial expressions recognition with multi-region divided attention networks for smart education cloud applications. *Neurocomputing*, 493, 119–128. <https://doi.org/10.1016/j.neucom.2022.04.052>
- Allaert, B., Ward, I., Bilasco, I., Djeraba, C., Bennamoun, M. (2022). A comparative study on optical flow for facial expression analysis. *Neurocomputing*, 500, 434–448.

- <https://doi.org/10.1016/j.neucom.2022.05.077>
- Nan, Y., Ju, J., Hua, Q., Zhang, H., Wang, B. (2022). A-MobileNet: An approach of facial expression recognition. *Alexandria Engineering Journal*, 61(6), 4435–4444. <https://doi.org/10.1016/j.aej.2021.09.066>
- Saxena, S., Tripathi, S., Sudarshan, T. (2022). An intelligent facial expression recognition system with emotion intensity classification. *Cognitive Systems Research*, 74, 39–52. <https://doi.org/10.1016/j.cogsys.2022.04.001>
- Barra, S., Hossain, S., Pero, C., Umer, S. (2022). A Facial Expression Recognition Approach for Social IoT Frameworks. *Big Data Research*, 30, 100353. <https://doi.org/10.1016/j.bdr.2022.100353>
- Zhu, D., Fu, Y., Zhao, X., Wang, X., Yi, H. (2022). Facial Emotion Recognition Using a Novel Fusion of Convolutional Neural Network and Local Binary Pattern in Crime Investigation. *Computational Intelligence and Neuroscience*, 2022, 1–14. <https://doi.org/10.1155/2022/2249417>
- Abdullah, M. U., Alkan, A. (2022). A Comparative Approach for Facial Expression Recognition in Higher Education Using Hybrid-Deep Learning from Students' Facial Images. *Traitement Du Signal*, 39(6), 1929–1941. <https://doi.org/10.18280/ts.390605>
- Li, S., Deng, Y. Q., Zhu, Z. L., Hua, H. L., & Tao, Z. Z. (2021). A Comprehensive Review on Radiomics and Deep Learning for Nasopharyngeal Carcinoma Imaging. *Diagnostics*, 11(9), 1523. <https://doi.org/10.3390/diagnostics11091523>
- Guan, D. (2020). Classical Architectures in CNN. Retrieved from <https://guandi1995.github.io/Classical-CNN-architecture/>. (Erişim tarihi:27.07.2020)
- Sayracı, B. (2021). DERİN ÖĞRENME VE EVRİŞİMSEL SİNİR AĞLARI. Retrieved from <https://medium.com/deep-learning-from-deepest/deri%CC%87n-%C3%B6%C4%9Frenme-ve-evri%CC%87%C5%9Fi%CC%87msel-si%CC%87ni%CC%87r-a%C4%9Flari-446dc8a8d2f>/. (Erişim tarihi:26.08.2021)
- Pandey, A. K. (2021). Convolution, Padding, Stride, and Pooling in CNN. Retrieved from <https://medium.com/analytics-vidhya/convolution-padding-stride-and-pooling-in-cnn-13dc1f3ada26> (Erişim tarihi:24.01.2021)

- Kalita, D. (2022, March 3). Basics of CNN in Deep Learning. Retrieved from <https://www.analyticsvidhya.com/blog/2022/03/basics-of-cnn-in-deep-learning/> (Erişim tarihi:03.03.2022)
- Yang, J., Yang, G. (2018). Modified Convolutional Neural Network Based on Dropout and the Stochastic Gradient Descent Optimizer. *Algorithms*, 11(3), 28. <https://doi.org/10.3390/a11030028>
- Gupta, D. (2020). Fundamentals of Deep Learning - Activation Functions and When to Use Them? Retrieved from <https://www.analyticsvidhya.com/blog/2020/01/fundamentals-deep-learning-activation-functions-when-to-use-them/> (Erişim tarihi:29.01.2020)
- Sharma, S. (2022). Activation Functions in Neural Networks. Retrieved from <https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6> (Erişim tarihi:20.11.2022)
- Gülüm, S. (2023). Softmax : Bir Aktivasyon Fonksiyonu. Retrieved from <https://medium.com/deeper-deep-learning-tr/softmax-bir-aktivasyon-fonksiyonu-da8382d8a281> (Erişim tarihi:03.07.2023)
- Krizhevsky, A., Sutskever, I., Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. <https://doi.org/10.1145/3065386>
- Guo, Z., Chen, Q., Wu, G., Xu, Y., Shibasaki, R., Shao, X. (2017). Village Building Identification Based on Ensemble Convolutional Neural Networks. *Sensors*, 17(11), 2487. <https://doi.org/10.3390/s17112487>
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L. C. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.1109/cvpr.2018.00474>
- He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. <https://doi.org/10.1109/cvpr.2016.90>

- Simonyan, K., Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations (ICLR 2015)*, 1–14. <https://doi.org/10.48550/arXiv.1409.1556>.
- Bengio, Y. (2012). Deep Learning of Representations for Unsupervised and Transfer Learning. Retrieved from <https://proceedings.mlr.press/v27/bengio12a.html>
- Hussain, L., Huang, P., Nguyen, T., Lone, K. J., Ali, A., Khan, M. S., Duong, T. Q. (2021). Machine learning classification of texture features of MRI breast tumor and peri-tumor of combined pre- and early treatment predicts pathologic complete response. *BioMedical Engineering OnLine*, 20(1). <https://doi.org/10.1186/s12938-021-00899-z>
- Yadav, S., Shukla, S. (2016). Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification. *2016 IEEE 6th International Conference on Advanced Computing (IACC)*. <https://doi.org/10.1109/iacc.2016.25>
- Li, A., An, L., Che, Z. (2020). A Facial Expression Recognition Model Based on Texture and Shape Features. *Traitement Du Signal*, 37(4), 627–632. <https://doi.org/10.18280/ts.370411>
- Lai, C. Q., Ibrahim, H., Abdullah, J. M., Azman, A., Abdullah, M. Z. (2021). Convolutional Neural Network Utilizing Error-Correcting Output Codes Support Vector Machine for Classification of Non-Severe Traumatic Brain Injury From Electroencephalogram Signal. *IEEE Access*, 9, 24946–24964. <https://doi.org/10.1109/access.2021.3056724>
- Ab Wahab, M. N., Nazir, A., Zhen Ren, A. T., Mohd Noor, M. H., Akbar, M. F., Mohamed, A. S. A. (2021). Efficientnet-Lite and Hybrid CNN-KNN Implementation for Facial Expression Recognition on Raspberry Pi. *IEEE Access*, 9, 134065–134080. <https://doi.org/10.1109/access.2021.3113337>
- Vikas Reddy, K., Rama Parvathy, L. (2022). Accurate Detection and Classification of Melanoma Skin Cancer Using Decision Tree Algorithm over CNN. *Advances in Parallel Computing Algorithms, Tools and Paradigms*. <https://doi.org/10.3233/apc220044>

- Kim, K. J., Tagkopoulos, I. (2019). Application of machine learning in rheumatic disease research. The Korean Journal of Internal Medicine, 34(4), 708–722.
<https://doi.org/10.3904/kjim.2018.349>
- Fumo, J. (2017). Types of Machine Learning Algorithms You Should Know. Retrieved from <https://towardsdatascience.com/types-of-machine-learning-algorithms-you-should-know-953a08248861> (Eriřim tarihi:17.08.2017)
- Candan, H. (2022). Adım Adım Makine Öğrenmesi Bölüm 3 : Denetimsiz Öğrenme Nedir? Retrieved from <https://medium.com/machine-learning-t%C3%BCrkiye/ad%C4%B1m-ad%C4%B1m-makine-%C3%B6%C4%9Frenmesi-b%C3%B6l%C3%BCm-3-denetimsiz-%C3%B6%C4%9Frenme-nedir-f890ada49a40> (Eriřim tarihi:02.01.2022)
- Bhatt, S. (2019). Reinforcement Learning 101. Retrieved from <https://towardsdatascience.com/reinforcement-learning-101-e24b50e1d292> (Eriřim tarihi:19.04.2019)