



**YAPAY ZEKA TEKNİKLERİ KULLANILARAK
MİKRO İFADELERİN TESPİTİ VE
SINIFLANDIRILMASI**

**2023
DOKTORA TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ**

Mehmet Zahit UZUN

**Tez Danışmanı
Doç. Dr. Yüksel ÇELİK**

**YAPAY ZEKA TEKNİKLERİ KULLANILARAK MİKRO İFADELERİN
TESPİTİ VE SINIFLANDIRILMASI**

Mehmet Zahit UZUN

**Tez Danışmanı
Doç. Dr. Yüksel ÇELİK**

**T.C.
Karabük Üniversitesi
Lisansüstü Eğitim Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalında
Doktora Tezi
Olarak Hazırlanmıştır**

KARABÜK

Haziran 2023

Mehmet Zahit UZUN tarafından hazırlanan “YAPAY ZEKA TEKNİKLERİ KULLANILARAK MİKRO İFADELERİN TESPİTİ VE SINIFLANDIRILMASI” başlıklı bu tezin Doktora Tezi olarak uygun olduğunu onaylarım.

Doç. Dr. Yüksel ÇELİK
Tez Danışmanı, Bilgisayar Mühendisliği Anabilim Dalı

Doç. Dr. Erdal BAŞARAN
İkinci Danışmanı, Ağrı İbrahim Çeçen Üniversitesi

Bu çalışma, jürimiz tarafından Oy Birliği ile Anabilim Dalınız Anabilim Dalında Doktora tezi olarak kabul edilmiştir. 22/06/2023

<u>Ünvanı, Adı SOYADI (Kurumu)</u>	<u>İmzası</u>
Başkan : Doç. Dr. Yüksel ÇELİK (KBÜ)	
Üye : Doç. Dr. Erdal BAŞARAN (AİÇÜ)	
Üye : Doç. Dr. İlker TÜRKER (KBÜ)	
Üye : Doç. Dr. Muhammed Kamil TURAN (KBÜ)	
Üye : Dr. Öğr. Üyesi Ali Burak ÖNCÜL (KÜ)	
Üye : Dr. Öğr. Üyesi Emrullah SONUÇ (KBÜ)	
Üye : Dr. Öğr. Üyesi Ferhat ATASOY (KBÜ)	

KBÜ Lisansüstü Eğitim Enstitüsü Yönetim Kurulu, bu tez ile, Doktora derecesini onamıştır.

Prof. Dr. Müslüm KUZU
Lisansüstü Eğitim Enstitüsü Müdürü



“Bu tezdeki tüm bilgilerin akademik kurallara ve etik ilkelere uygun olarak elde edildiğini ve sunulduğunu; ayrıca bu kuralların ve ilkelerin gerektirdiği şekilde, bu çalışmadan kaynaklanmayan bütün atıfları yaptığımı beyan ederim.”

Mehmet Zahit UZUN

ÖZET

Doktora Tezi

YAPAY ZEKA TEKNİKLERİ KULLANILARAK MİKRO İFADELERİN TESPİTİ VE SINIFLANDIRILMASI

Mehmet Zahit UZUN

Karabük Üniversitesi

Lisansüstü Eğitim Enstitüsü

Anabilim Dalınız Anabilim Dalı

Tez Danışmanı:

Doç. Dr. Yüksel ÇELİK

İkinci Tez Danışmanı:

Doç. Dr. Erdal BAŞARAN

Haziran 2023, 118 sayfa

Mikro ifade, duyguların maskelenmek istenip bastırılmaya çalışıldığı sırada kontrol edilemeyip istemsizce sızıntı şeklinde oluşan, tamamen gerçek duyguların yansıması, kısa süreli yüz kas hareketidir. Yüzün sadece bir kısmında görülen düşük yoğunluklu bu ifadeleri tespit etmek ve tanımak klinik, adli, ulusal güvenlik, gümrük, iletişim, eğitim, ticari, siyasi gibi birçok alanda hayati önem arz etmektedir. Bu alanda uzmanlaşmamış insanların bu ifadeleri fark etme oranları %32 gibi çok düşükken, uzman kişiler için bile mikro ifade belirleme oranı %47'dir ve bu oran çok yüksek değildir. Mikro ifade tanıma yapılamadığı takdirde klinik hastalarında intihar, gümrüklerde kaçakçılık, hukukta adaletsiz yargılama gibi hayati olaylar yanlış

yorumlanıp olumsuz neticelere sebebiyet verebilir. Mikro ifade tespit ve tanınması uzman tarafından video klip karelerinin saatler süren uzun ve detaylı bir şekilde analiz edilmesi gerekmektedir. Kamuda ve özel sektörde mikro ifade alanında yetişmiş uzman kişilerin varlığı çok olmadığı gibi saatler süren eğitim sonunda bile istenilen verim elde edilememektedir. Bu tez çalışmasında yapay zeka metotlarından faydalanılarak mikro ifade video klip görüntülerinde duyguların belirlenmesine çalışılmıştır. Önerilen model çalışmalarının ikisinde kamuya açık veri setlerinden en popüler olan CASME II, SAMM, SMIC de bulunan örnekler birleştirilip bileşik yeni bir veri seti kullanılırken, önerilen bir diğer model çalışmasında ise CASME II veri seti kullanılmıştır. İlk olarak SAMM veri setindeki örneklerde yüz algılama, hizalama ve kırpma işlemleri ile yüz bölgesini çıkarma, SMIC veri setindeki video kare dizilerinde ifadenin en yoğun olduğu tepe (apex) karesinin indeks konumları tespit etme işlemleri gerçekleştirilmiştir. Ardından TV-L1 ve Farneback optik akış teknikleri ile her video klip kare dizisi için tepe ve başlangıç(onset) kare farkından yararlanılarak mikro ifade hareket bilgisi elde edilmiştir. Görüntülere ait özellik haritalarının çıkarılmasında, veriye dayalı derin öğrenme yöntemlerinden VGG-16, AlexNet, SqueezeNet, MobilNetV2, EfficentNetB0, DenseNet121, DenseNet169, DenseNet201, Xception Evrişimsel Sinir Ağı (ESA) modelleri ve geleneksel yöntemlerden Gabor filtresinden deneysel çalışmalarda faydalanılmıştır. ESA modellerinde transfer öğrenme ile ImageNet'e ait öğrenilmiş ağırlıklarından yararlanılmıştır. Özellik seçimi için çalışmaya ait deneylerde Parçacık Sürü Optimizasyonu (PSO), Özyinelemeli Özellik Eleme (ÖÖE) ve Çapraz Doğrulama ile Özyinelemeli Özellik Eleme (ÇDÖÖE) algoritmaları, sınıflandırma sürecinde ise Destek Vektör Makineleri (DVM)'nin doğrusal, quadratic, finegaussian, cubic çekirdekleri kullanılmıştır. Sonuç olarak; önerilen üç model içerisinde en başarılı sonuçlar CASME II veri seti ile birlikte sırasıyla Xception ESA modeli, ÇDÖÖE özellik seçim algoritması ve doğrusal DVM sınıflandırıcısından oluşan mikro ifade duygu tanıma modelinden %92.48 doğruluk performansı elde edilmiştir.

Anahtar Sözcükler: Mikro ifade, yüz ifadeleri, yüz algılama, optik akış, özellik çıkarma, makine öğrenmesi, derin öğrenme, özellik seçme

Bilim Kodu : 92432



ABSTRACT

Ph. D. Thesis

DETECTION AND CLASSIFICATION OF MICROEXPRESSIONS USING ARTIFICIAL INTELLIGENCE TECHNIQUES

Mehmet Zahit UZUN

**Karabük University
Institute of Graduate Programs
Department of Computer Engineering**

Thesis Advisor:

Assoc. Prof. Dr. Yüksel ÇELİK

Co-Advisor:

Assoc. Prof. Dr. Erdal BAŞARAN

June 2023, 118 pages

A micro-expression is a brief facial muscle movement that occurs involuntarily and uncontrollably as a leakage of genuine emotions when there is an attempt to conceal or suppress emotions. It completely reflects real feelings and occurs for a short duration. Detecting and recognizing these low-intensity expressions that occur only on the part of the face is of vital importance in various fields such as clinical, forensic, national security, customs, communication, education, commercial, and political domains. While the detection rate of these expressions is very low, such as 32%, for individuals who are not specialized in this field, even experts in the area have a

micro-expression identification rate of only 47%, which is not considered to be very high. In cases where micro-expression recognition cannot be performed, vital events such as suicide in clinical patients, smuggling at customs, and unjust judgments in law may be misinterpreted, leading to negative consequences. Micro-expression detection and recognition require an expert's laborious and detailed analysis of video clip frames, which can take hours. The presence of trained professionals in the field of micro-expression is scarce both in the public and private sectors, and even after hours of training, the desired efficiency cannot be achieved. This study utilized artificial intelligence methods to determine emotions in micro-expression video clips. In the two suggested model studies, a compound new dataset was created by combining samples from the publicly available datasets CASME II, SAMM, and SMIC, among the most popular ones. On the other hand, in another proposed model study, only the CASME II dataset was used. Firstly, facial detection, alignment, and cropping operations were performed in the SAMM dataset to extract the facial regions. In the SMIC dataset, the operations involved detecting the index positions of the frames where the expression was most intense, known as the apex frames. Subsequently, the TV-L1 and Farneback optical flow techniques were utilized to obtain micro-expression motion information for each video clip frame sequence. This was achieved by leveraging the differences between the apex frame and the onset frame, which represents the start of the micro-expression. In the extraction of feature maps from the images, experimental studies were conducted using data-driven deep learning methods such as VGG-16, AlexNet, SqueezeNet, MobileNetV2, EfficientNetB0, DenseNet121, DenseNet169, DenseNet201, Xception Convolutional Neural Network (CNN) models, as well as traditional methods such as the Gabor filter. In CNN models, transfer learning was employed using pre-trained network weights from ImageNet. For feature selection, Particle Swarm Optimization (PSO), Recursive Feature Elimination (RFE), and Cross-Validated Recursive Feature Elimination (RFECV) algorithms were utilized in experimental studies. Support Vector Machines (SVMs) with linear, quadratic, finegaussian, and cubic kernels were used in the classification process. As a result, among the three proposed models, the micro-expression emotion recognition model composed of the Xception CNN model, RFECV feature selection algorithm,

and linear SVM classifier achieved a performance accuracy of 92.48% when evaluated with the CASME II dataset, yielding the most successful results.

Key Word :Micro-expression, facial expressions, face detection, optical flow, feature extraction, machine learning, deep learning, feature selection.

Science Code : 92432



TEŞEKKÜR

Bu tez konusunun belirlenmesinde, çalışmanın planlanmasında, araştırma sürecinde, yürütülmesinde ve tamamlanmasında desteğini ve ilgisini esirgemeyen, derin bilgi ve birikimlerinden faydalandığım, yönlendirmeleri ve katkılarıyla tezimi bilimsel temeller üzerinde şekillendiren, akademik olarak çok şey öğrenerek mesleki gelişimime önemli katkılar sağlayan çok değerli sayın danışman hocalarım Doç. Dr. Yüksel ÇELİK ve Doç. Dr. Erdal BAŞARAN' a sonsuz şükranlarımı sunarım.

Tez çalışmam süresince kıymetli değerlendirme ve önerileriyle desteklerini sunan sayın hocalarım Doç. Dr. İlker TÜRKER ve Doç. Dr. Muhammed Kamil TURAN' a teşekkürlerimi sunarım.

Doktora eğitimi süresinde değerli katkılarından dolayı Karabük Üniversitesi Bilgisayar Mühendisliği Bölümü tüm öğretim elemanlarına teşekkürlerimi sunarım.

Yaşamım boyunca maddi manevi yanımda olan, eğitim hayatımda hep destek olan başta anneme, dedeme, rahmetli babanneme ve rahmetli babama sonsuz şükranlarımı sunarım.

İÇİNDEKİLER

	<u>Sayfa</u>
KABUL	ii
ÖZET.....	iv
ABSTRACT.....	vii
TEŞEKKÜR.....	x
İÇİNDEKİLER	xi
ŞEKİLLER DİZİNİ.....	xv
ÇİZELGELER DİZİNİ	xvii
SİMGELER VE KISALTMALAR DİZİNİ	xix
BÖLÜM 1	1
GİRİŞ.....	1
1.1. PROBLEMİN TANIMI	1
1.2. MİKRO İFADELER İLE İLGİLİ ZORLUKLAR	1
1.3. TEZİN AMACI	2
BÖLÜM 2	5
LİTERATÜR ARAŞTIRMASI	5
BÖLÜM 3	12
YÜZ İFADELERİ VE YÜZ HAREKETİ KODLAMA SİSTEMİ	12
3.1. EVRENSEL DUYGULAR	14
3.2. BİLGİSAYARLI GÖRÜ VE YÜZ İFADELERİ	14
3.3. MİKRO İFADELER	16
BÖLÜM 4	18

	<u>Sayfa</u>
MATERYAL VE METOD	18
4.1. VERİ SETLERİ.....	18
4.2. ÖN İŞLEME METODLARI	21
4.2.1. D&C-RoIs.....	21
4.2.2. Yüz Bölgesinin Çıkarılması.....	22
4.2.3. Optik Akış.....	23
4.2.4. Veri Arttırma	26
4.3. ÖZELLİK ÇIKARMA YÖNTEMLERİ	27
4.3.1. Gabor Filtresi Metodu.....	27
4.3.2. Evrişimsel Sinir Ağları	28
4.3.2.1. Evrişimsel Sinir Ağları Katmanları	29
4.3.2.2. Aktivasyon Fonksiyonları.....	33
4.3.2.3. ESA Parametreleri	38
4.3.2.4. ESA Optimizasyon Algoritmaları.....	39
4.3.3. Kullanılan Esa Modelleri.....	43
4.3.3.1. AlexNet.....	43
4.3.3.2. VGGNet	44
4.3.3.3. SqueezeNet	44
4.3.3.4. MobilNet V2	45
4.3.3.5. EfficientNetB0	46
4.3.3.6. DenseNet.....	47
4.3.3.7. Xception.....	48
4.4. ÖZELLİK SEÇİMİ METODLARI	48
4.4.1. Parçacık Sürü Optimizasyonu.....	48
4.4.2. Özyinelemeli Özellik Eleme (ÖÖE) Algoritması.....	50
4.4.3. Çapraz Doğrulama ile Özyinelemeli Özellik Eleme (ÇDÖÖE)	50
4.5. SINIFLANDIRNMA DESTEK VEKTÖR MAKİNELERİ	51
4.6. PERFORMANS ÖLÇÜTLERİ	53
4.6.1. K-Katlı Çapraz Doğrulama.....	53

4.6.2. Karmaşıklık Matrisi	54
4.6.3. Alıcı İşlem Karakteristiği (AİK).....	56
 BÖLÜM 5	58
ÖNERİLEN MİKRO İFADE TANIMA METODLARI	58
5.1. DERİN TRANSFER ÖĞRENME VE PARÇACIK SÜRÜ OPTİMİZASYON ALGORİTMASI İLE MİKRO İFADE TANIMA (DTÖ-PSO).....	58
5.1.1. Veri Seti	60
5.1.2. Ön İşleme.....	61
5.1.3. Özellik Çıkarımı	63
5.1.4. PSO Optimizasyon Algoritması ile Özellik Seçimi	64
5.1.5. Sınıflandırma	65
5.2. DENSENET169 MODELİ İLE ÖZYİNELEMELİ ÖZELLİK ELEME (DensNet169-ÖÖE).....	66
5.2.1. Veri Seti	67
5.2.2. Ön İşleme.....	67
5.2.3. Özellik Çıkarımı	68
5.2.4. ÖÖE Algoritması ile Özellik Seçimi	68
5.2.5. Sınıflandırma	68
5.3. XCEPTIONNET İLE ÇAPRAZ DOĞRULAMALI ÖZYİNELEMELİ ÖZELLİK ELEME TABANLI MODEL (XNET-ÇDÖÖE).....	69
5.3.1. Veri Seti	70
5.3.2. Ön İşleme.....	70
5.3.3. Özellik Çıkarımı	71
5.3.4. Özellik Seçimi	72
5.3.5. Sınıflandırma	72
 BÖLÜM 6	73
DENEYSEL TEST SONUÇLARI VE TARTIŞMA.....	73
6.1. DTÖ-PSO MODELİ İÇİN DENEYSEL TEST SONUÇLARI	73

	<u>Sayfa</u>
6.2. ÖNERİLEN DENSNET169-ÖÖE MODELİ İÇİN PERFORMANS SONUÇLARI	82
6.3. ÖNERİLEN XNET-ÇDÖÖE MODELİ İÇİN PERFORMANS SONUÇLARI	90
6.4. TARTIŞMA.....	97
 BÖLÜM 7	 100
SONUÇ VE ÖNERİLER	100
KAYNAKLAR	103
 ÖZGEÇMİŞ	 118

ŞEKİLLER DİZİNİ

Sayfa

Şekil 3.1. YHKS’de bulunan yüz eylem birimleri [59].	13
Şekil 4.1. CASME-II veri setinde bulunan sınıflara ait tepe kare örnekleri.	19
Şekil 4.2. SAMM veri setinde bulunan sınıflara ait tepe kare örnekleri.	19
Şekil 4.3. SMIC veri setinde bulunan sınıflara ait tepe kare örnekleri.	20
Şekil 4.4. a) Yüzün tespiti, yüzün işaretlenmesi, b) yüzün çerçeveye alınması c) yüzün kırılması.	23
Şekil 4.5. Başlangıç ve tepe kareleri ile optik akış görüntülerinin elde edilmesi.	26
Şekil 4.6. a) RGB tepe kare görüntüsü, b) Gabor Filtreleme işleminden geçmiş tepe kare görüntüsü.	28
Şekil 4.7. Giriş görüntüsü üzerinde evrişim işleminin hesaplanması.	31
Şekil 4.8. Maksimum ve ortalama havuzlama işleminin hesaplanması.	32
Şekil 4.9. Tam bağlı katmanın yapısı ve bağlantıları.	33
Şekil 4.10. Sigmoid fonksiyonunun dalga görünümü.	34
Şekil 4.11. Hiperbolik tanjant fonksiyonunun dalga görünümü.	35
Şekil 4.12. ReLu fonksiyonunun dalga görünümü.	37
Şekil 4.13. Seyreltme katmanın yapısı ve bağlantıları.	38
Şekil 4.14. a) Momentumlu SGD algoritması, b) Momentumsuz SGD algoritması.	41
Şekil 4.15. Ters Çevrilmiş Artık Bloğu’nun yapısı ve katmanları.	46
Şekil 4.16. DVM sınıflandırma işlemi ile örneklerin ayrılması.	52
Şekil 4.17. a) Doğrusal olarak ayrılmayan veri kümesi b) verilerin daha yüksek bir uzay boyutunda yansıtılması c) herhangi bir çekirdek fonksiyonu ile DVM birlikte kullanılarak sınıflandırılması.	53
Şekil 4.18. Beş kat çapraz doğrulamaya ait bir örnek uygulamanın grafiksel gösterimi.	54
Şekil 4.19. İki sınıflı bir veri setinde karmaşıklık matrisinin parametrelerinin elde edilmesi.	55

Şekil 5.1. Mİ sınıflandırma görevi için önerilen DTÖ-PSO modelinin mimari yapısı.	59
Şekil 5.2. Mİ video dizindeki başlangıç bitiş ve tepe karlerlerin konumları.	63
Şekil 5.3. Başlangıç ve tepe karesi kullanılarak optik akışın elde edilmesi.....	63
Şekil 5.4. Mİ sınıflandırma görevi için önerilen DensNet169-ÖÖE modelinin mimari yapısı.....	66
Şekil 5.5. Mİ sınıflandırma görevi için önerilen Xnet-ÇDÖÖE modelinin mimari yapısı.	70
Şekil 6.1. Çapraz doğrulama olmadan elde edilen karmaşıklık matrisi ve auc roc grafığı.	77
Şekil 6.2. Çapraz doğrulama olmadan elde edilen karmaşıklık matrisi ve AİK eğrisi grafığı.	78
Şekil 6.3. Çapraz doğrulamalı modelin karmaşıklık matrisi.....	79
Şekil 6.4. Çapraz doğrulamalı modelin AİK eğrisi grafığı.	80
Şekil 6.5. Modelin karmaşıklık matris tablosu.....	87
Şekil 6.6. Modelin AİK eğrisi grafığı.	88
Şekil 6.7. Xception modeline ait karmaşıklık matrisi.	94
Şekil 6.8. Xception modeline ait AİK eğrisi grafığı.	94
Şekil 6.9. Xception modeline ait öğrenme eğrisi grafığı.	95

ÇİZELGELER DİZİNİ

	<u>Sayfa</u>
Çizelge 3.1. Altı temel duyguya ait eylem birimleri [48].	14
Çizelge 4.1. Spontane Mİ veri setlerine ait genel özellikler	20
Çizelge 5.1. DTÖ-PSOmodelinde kullanılan bileşik veri setindeki örneklerin özellikleri.....	61
Çizelge 5.2. DTÖ-PSOmodelinde ESA modellerinden elde edilen özellikler.	64
Çizelge 5.3. DensNet169-ÖÖE modelinde kullanılan bileşik veri setindeki örneklerin özellikleri.....	67
Çizelge 5.4. Xnet-ÇDÖÖE modelinde kullanılan CASME-II veri setindeki örneklerin özellikleri.....	70
Çizelge 6.1. ESA modelleri kullanılarak sınıflandırılan deneysel ölçümlerin sonuçları	74
Çizelge 6.2. 1000 ESA özellik haritasının DVM çekirdekleriyle sınıflandırılması... 75	
Çizelge 6.3. ESA modellerinden elde edilen 1000'er özelliğin birleştirilip DVM çekirdekleri ile sınıflandırma performansları.	76
Çizelge 6.4. PSO analizinden sonra DVM çekirdekleri ile sınıflandırma sonuçları.. 77	
Çizelge 6.5. PSO analizi ve çapraz doğrulama sonrası DVM çekirdekleri ile sınıflandırma sonuçları	79
Çizelge 6.6. Çapraz doğrulama olmadan ve çapraz doğrulama ile önerilen modelin en iyi performansları	81
Çizelge 6.7. Bütünleşik veri setinin DenseNet modelleri ile sınıflandırılmasının sonuçları.	82

Çizelge 6.8. Arttırılmış bütünleşik veri setinin DenseNet modelleri ve Doğrusal DVM ile sınıflandırma sonuçları.....	83
Çizelge 6.9. Arttırılmış bütünleşik veri setinin DenseNet modelleri doğrusal DVM-ÖÖE ile sınıflandırma sonuçları.....	85
Çizelge 6.10. Zenginleştirilmiş özellik haritalarının DenseNet modelleri ve doğrusal DVM ile sınıflandırma sonuçları	86
Çizelge 6.11. Zenginleştirilmiş özellik haritalarının DenseNet modelleri doğrusal DVM-ÖÖE ile sınıflandırma sonuçları.....	87
Çizelge 6.12. Zenginleştirilmiş ikili özellik haritalarının DenseNet modelleri doğrusal DVM-ÖÖE ile sınıflandırma sonuçları.....	88
Çizelge 6.13. Modellerin arttırılmamış CASME-II veri setine uygulanması ve doğrusal DVM ile sınıflandırma sonuçları.....	91
Çizelge 6.14. Modellerin dengeli arttırılmış CASME-II veri setine uygulanması ve doğrusal DVM ile sınıflandırma sonuçları.....	92
Çizelge 6.15. Filtrelenmiş özellik haritalarının Xception ve Gabor modelleri ve doğrusal DVM ile sınıflandırma sonuçları.....	93
Çizelge 6.16. Önerilen çalışmanın en son modellerle karşılaştırılması	97

SİMGELER VE KISALTMALAR DİZİNİ

SİMGELER

- α_k : lagrange katsayısı
 $H_k()$: toplu normalleştirme
 m_x, m_y : gabor fonksiyonunun merkezi
 x_k : eğitim örnekleri
 y_k : eğitim örneklerinin etiketi
 σ_x^2, σ_y^2 : alıcı alan
 $\nabla F(\theta_i)$: i inci iterasyon için kayıp fonksiyonunun gradyanı
 A : Simetrik matris
 b : vektör
 c_1 : skaler
 c_1, c_2 : ivme katsayıları
 d : global öteleme
 f : filtre matrisinin boyutu
 f_2 : sinyal
 g : çekirdek
 $g_{en_iyi}^i$: sürünün t anında çözüme en çok yaklaşılan durum
 h_1 : ilk kare
 h_2 : diğer kareler
 I_{apex} : tepe karesi
 I_{onset} : başlangıç karesi
 i : giriş matrisinin boyutu
 k : çapraz doğrulama için veri setinin bölünme sabiti
 ken_iyi^i : parçacığın t anında çözüme en yaklaştığı durum
 m, n : çıktı matrisin satır ve sütunu

p :dolgu boyutu

rnd_1, rnd_2 : hız güncellemesinde yeniden oluşturulan rastgele değerler

s : adım sayısı

t : giriş görüntüsü

v_i^t : çekirdek parçacığının t anındaki hızı

X : görüntü tabanlı bir hata kriterini en aza indiren bir eşitsizlik haritası

x,y : evrişime ait matrisin satır ve sütunu

x_i^t : çekirdek parçacığının t anındaki konumu

z :çıktı matrisinin boyutu

B : eylemsizlik katsayısı

γ : yön

δ :faz kayması

η_i : öğrenme oranı, adım boyutu

λ : veri ögesinin ağırlık sabiti

ϕ : görüntü verilerinin aslına uygunluğu

ψ : düzenlileştirme terimi

$k()$: Çekirdek fonksiyonu

KISALTMALAR

ÇDÖÖE: Çapraz Doğrulama ile Özyinelemeli Özellik Eleme

DAE : Derinlemesine Ayrılabilir Evrişim

DN : Doğru Negatif

DP : Doğru Pozitif

DTÖ : Derin Transfer Öğrenme

DVM : Destek Vektör Makinesi

ESA : Evrişimsel Sinir Ağı

EVB : Eulerian Video Büyütme

EYK : En Yakın Komşuluk

Mİ : Mikro İfade

ÖÖE : Özyinelemeli Özellik Eleme

PSO : Parçacık Sürü Optimizasyonu

YHKS : Yüz Hareketi Kodlama Sistemi

YN : Yanlış Negatif

YP : Yanlış Pozitif

BÖLÜM 1

GİRİŞ

1.1. PROBLEMİN TANIMI

Yüz ifadeleri, kişilerin duygu ve niyetinin anlaşılmasında önemli rol oynayan evrensel bir iletişim aracıdır [1,2]. Duyguların yüze yansması, günlük hayatta herkes tarafından aşına olunan, çok sık karşılaşılan yüz ifadeleri olan makro ya da tam ifadelerin yanı sıra sadece belirli koşullar altında meydana gelen mikro ifadeler olarak adlandırılan yüz ifadelerinde de görülmektedir. Makro ifadeler her zaman gerçek duyguları yansıtmayıp, aksi yüz ifadesi uygulanarak gerçek duygular gizlenebilir fakat mikro ifadeler (Mİ) (Micro Expression - ME) gerçek duyguları anlamak için kullanılır [3]. Çünkü Mİ bilinçsiz ve istemsiz olarak meydana gelirler. Mİ'ler yüksek riskli ve güvenli olmadığı düşünülen ortamlarda, insanlar yüz ifadelerini kontrol edip bastırmaya çalıştığı anlarda çok ince ve hızlı meydana gelir [4]. İnsana ait duygularının yapay olmayan, saf ve tamamen gerçek biçimde yansması olan Mİ'lerin tanınması klinik, adli, eğitim, ticari, gümrük, ulusal güvenlik, siyasi gibi birçok alanda yanıltıcı ve istenmeyen durumlardan kaçınılmasını sağlamaktadır. Literatüre göre Mİ'lerin gerçekleşme sürelerinde tam bir fikir birliği olmasa da en kapsamlı aralığı saniyenin 1/25 ile 1/2 arasındaki değişen sürelerde şekillenip (biçimlenip) ortaya çıkıp kaybolduğu araştırmalarda belirtilmektedir [5,6]. Mİ oluşumu sıradan insanlar tarafından bile çok zor anlaşıldığı gibi uzmanlar tarafından bile tanıma oranı %47 olarak belirlenmiştir [7]. Mİ'ler tespit edilip doğru bir şekilde tanındığında intihar meylindeki şizofreni hastalarının, gümrükteki yasa dışı olayların, hukuktaki yanlış yargılamaların, eğitimde dikkat bozukluğu, siyasette yalan söylemlerin gibi hayati önem taşıyan birçok konuda zamanında önüne geçilip tedbirler alma açısından yetkililere ve devlete hayati faydalar sağlayacaktır. Aksi halde tamamen gerçek duyguları gösteren bu ipucu verilerinden mahrum yetkililer günün birinde sonuçları

büyük zarara neden olabilecek yanlış kararlar vermesi muhtemeldir. Günümüzde Mİ tespit ve tanıma kabiliyetini arttırmak, bu alanda uzman ve yeterli kişiler yetiştirmek için eğitim araçları geliştirilse de istenilen ve beklenen oranlarda bir katkı yapmadığı eğitim sonrası yapılan testlerde görülmüştür. Bilgisayar teknolojilerindeki gelişmeler sonrası araştırmacılar tarafından yapay zeka, bilgisayarlı görme, görüntü/video işleme alanlarında Mİ'lerin tespit ve tanınması için çalışmalar yapılmaktadır [8]. Böylece Mİ erken teşhisi sayesinde, intihara eğilimli şizofreni hastalarının, gümrük kaçakçılarının ve yasa dışı göçmenliğin, hatalı adli yargılamaların önlenmesi gibi birçok alanda yetkililerin karar verme süreçlerinde ve iş yüklerinin azalmasında destek sağlayacaktır.

Mİ içeren video klipleri, ifadenin şekillenmeye başladığı kare olan başlangıç (onset), ifadenin en yoğun görüldüğü, pik yaptığı kare olan tepe (apex) ve ifadenin sönümlenip bittiği anı içeren kare olan bitiş (offset) olmak üzere üç önemli video klip karesinden oluşmaktadır [9]. Mİ'lerin kısa süreli ve düşük yoğunluklu olmasından kaynaklanan sorunlarla başa çıkabilecek yöntemler geliştirmenin zorlukları [10] da göz önüne alındığında, bu çalışmada literatürde bulunan diğer bir çok çalışma [11,12,13,14] gibi ifadelerin en yoğun bulunduğu tepe karesine ve başlangıç karesine odaklanmaktadır.

1.2. MİKRO İFADELER İLE İLGİLİ ZORLUKLAR

Mİ tanıma görevinde ifadelerin kısa sürmesi, yoğunluğunun düşük olması gibi zorluklarının yanı sıra farklı birtakım zorluklarda içermektedir. Hemen hemen tüm mevcut veri setleri belirli bir ülkeye veya etnik kökene ait katılımcılardan alınan örnekleri içermektedir. Her ne kadar yüz ifadeleri evrensel olsada duygu aktarımları katılımcıların kültürel geçmişlerinden dolayı aynı duygulara farklı şekillerde veya farklı yoğunlukta tepki verebilirler. Bundan dolayı iyi bir veri seti, örneklerinde olması gerektiği gibi etnik olarak da iyi bir dağılıma sahip olmalıdır [15].

Günümüzdeki veri setleri 30 ila 40 denekten oluşan kısıtlı örneğe, maksimum birkaç yüz, sahip, küçük ölçekli Mİ videolarıdır. Mİ etiketlemenin çok uzun bir süreç aldığı göz önünde bulundurulursa yüksek örneklemeye sahip veri setlerinin oluşturulması

son derece zorlayıcı ve zaman alıcı olacaktır [16]. Mutluluk gibi bazı duyguların ortaya çıkarılması korku, üzüntü, öfke gibi duygulara kıyasla daha kolay olduğu için veri setlerinde ki duygu başına örnek dağılımı ve denek başına örnek dağılımları dengesizdir. Buda Mİ tanıma üzerine yapılan çalışmalarda bazı duygular lehine önyargıya sebebiyet verebilir [17]. İnsan yüz ifadelerinde eylem birimlerin nesnel tanımlayıcısı YHKS kodları, tüm veri setlerinde bulunması gerekirken bazı veri setlerinde verilmemiştir. Bunun yanı sıra gözlemci kodlayıcılarında güvenilirlik raporları veri setlerinde sunulmalıdır [15].

Mİ kayıtlarının yapıldığı stüdyolardaki baş duruşu ve baş hareketlerinden kaynaklı aydınlatma gölge değişimlerine neden olabilmektedir. Buda piksel değişiminin yoğunluğundan etkilenen LBP, HOG gibi özellik tanımlayıcılarının yanlış özellik tahminine neden olup performanslarını olumsuz etkileyebilmektedir [17,18]. Ayrıca stüdyo ortamında elde edilen Mİ video klip örnekleri gerçek hayattaki Mİ örneklerini iyi bir şekilde temsil etmeyebilir. Gerçek hayattan seçilen örneklere sahip Mİ veri setleri uygulama alanında büyük ilerleme sağlayabilir [15]

1.3. TEZİN AMACI

Bu tezin amacı; bir Mİ video klibinin hangi duyguya ait olduğunu etiketlemek için kare kare analiz edip belirlemenin saatlerce sürdüğü göz önüne alınırsa, Mİ tespit ve tanıma görevini yapay zekâ ve makine öğrenmesi yöntemlerinden faydalanılarak sınıflandırmaktır. Mİ tespit görevinde; tepe kare konum indeksleri sunulmayan SMIC veri seti video kare örnekleri için LBP özellik tanımlayıcısından faydalanarak maksimum özellik farkını hesaplayan D&C-RoIs tekniği ile konum indeks numaraları tespit edilmektedir. Ardından önışlem yapılmamış veri setlerinde OpenCV Dlib kütüphanesinde yer alan metotlar kullanılarak yüzün algılanması, yüzün noktalarla işaretlenmesi ve kırılması süreçleri bu iki kareye uygulanarak görüntülerden yüz bölgeleri çıkarılmıştır. İki kare arasındaki piksel değişimini hesaplayan optik akış yöntemi kullanılarak, başlangıç ve tepe kareleri arasındaki piksel farklarından küçük hareketlerin yakalandığı görüntüler elde edilmektedir. Mevcut veri setlerindeki örnek dengesizliği ve eksikliği sorunlarını çözmek için optik akış ile elde edilen görüntüler

veri arttırma teknikleri kullanılarak veri seti örnekleri yapay olarak çoğaltılmıştır. Mİ sınıflandırmak için görüntü özelliklerinin otomatik olarak elde edildiği Mİ görüntülerinde başarı oranının oldukça yüksek olduğu derin öğrenme temelli transfer öğrenme modelleri ile görüntülerin sınıflandırılması çalışmaları yapılmıştır. Buna ilaveten; tez çalışmamızda Mİ'lerin otomatik olarak tespit edilerek görüntülerin sınıflandırılmasına yönelik özgün bir model önerisi geliştirilmiştir. Ayrıca deneysel çalışmalarımızda örnek sayısının azlığından dolayı halka açık olarak sunulan üç doğal Mİ veri seti CASME-II, SAMM ve SMIC bir çok çalışmada [15,16,17] olduğu üzere birleştirilerek bileşik bir veri seti de oluşturulmuştur.

Bu kapsam dahilinde araştırma çalışması olan bu tez çalışması altı bölümden oluşmaktadır. Giriş bölümünde teze ait genel bilgiler verilerek konuya ait genel bir bakış açısı kazandırılmak istenmiştir.

Literatür araştırmasının sunulduğu ikinci bölümde, Mİ tespit ve tanıma yönelik yapay zekâ ve makine öğrenmesi ile yapılan çalışmaların ayrıntılı literatür taramasının özeti verilmiştir.

Üçüncü bölümde, yüz ifadeleri, duygular ve Mİ'nin tanımı ve aralarındaki ilişki hakkında bilgiler sunulmuştur.

Materyal metot olarak adlandırılan dördüncü bölümde, bu tez kapsamında yapılan deneysel araştırmalarda kullanılan veri setlerinin tanımlanması, verileri normalleştirmek için kullanılan teknikler, özellik çıkarmak için kullanılan geleneksel ve veriye dayalı yöntemler ve Mİ görüntülerinin sınıflandırılmasında kullanılan yöntem ve tekniklere ait bilgiler verilmiştir.

Önerilen Mİ tanıma metotlarına değinilen beşinci bölümde tez kapsamında önerilen her model için önileme, özellik çıkarma, özellik seçme ve sınıflandırma adımlarını detaylandıran bilgiler sunulmuştur.

Altıncı bölüm olan bulgular ve tartışma kısmında; yapılan deneysel çalışmaların neticesinde elde edilen analiz sonuçları, tablolar ve şekiller aracılığı ile verilmiştir. Ayrıca çalışmadan elde edilen sonuçlar, literatürde son zamanlarda Mİ alanında yapılan diğer çalışmalar ile karşılaştırılmış ve analiz sonuçları tartışılarak detaylandırılmıştır.

Sonuç ve önerilerin sunulduğu yedinci ve son bölümde; çalışma sonunda, araştırmalardan elde edilen kazanımlar, görüş, düşünce ve öneriler sunulmuştur. Yapılan bu tez çalışmasının literatüre sunduğu katkılar bahsedilmiş ve Mİ alanında gelecekte yapılabilecek araştırmalar üzerine görüş ve düşünceler bildirilmiştir.



BÖLÜM 2

LİTERATÜR ARAŞTIRMASI

Mİ tespit ve tanıma yönelik literatürde bulunan yapay zekâ tabanlı çalışmalar incelendiğinde; Mİ görevi sırasıyla ön işleme, özellik çıkarma ve sınıflandırma süreçlerinden meydana geldiği görülmektedir. Ayrıca özellik çıkarma sürecinde de geleneksel yöntemler ve veriye dayalı derin öğrenme yöntemleri kullanılmaktadır.

Veri setlerindeki farklılıklar ve örnekleme sayısının düşük olması gibi olumsuzluklarının yanı sıra mevcut doğal Mİ veri setlerinin azlığı problemleri araştırmacılar için hala önemli zorlukların başında gelmektedir [22]. Literatürde yapılan çalışmaların birçoğunda, mevcut veri setlerini birleştirilerek elde ettikleri yeni bütünlük veri setleri üzerinde çalışılmıştır. 2018'de 13. Elektrik ve Elektronik Mühendisleri Enstitüsü (Institute of Electrical and Electronics Engineers -IEEE) uluslararası Otomatik Yüz ve Jest Tanıma (Automatic Face and Gesture Recognition-FG) konferansında düzenlenen Yüz Mikro İfadesi Büyük Yarışmasında (Facial Micro-Expression Grand Challenge-MEGC) CASME-II ve SAMM veri setleri birleştirilerek bileşik bir veri seti kullanılmıştır [23]. Aynı şekilde 2019'da aynı yarışmada CASME-II, SAMM ve SMIC veri setleri birleştirilerek bileşik bir veri seti kullanılmıştır [19]. 2019 yılındaki yarışmaya göre bileşik bir veri seti oluşturulsa bile toplam örnek sayısı 500'e ulaşamamaktadır.

Ön işleme süreci, bir sonraki süreç olan özellik çıkarma sürecinin verimini arttırmak amacıyla Mİ görüntülerinin normalleştirilmesidir. Veri setindeki örnek sayısının düşük olmasının yanı sıra, birde Mİ'nin yüzün sadece belli bir bölümünde, hızlı ve düşük yoğunlukta meydana gelmesinden ötürü baş hareketleri, arka plan veya ışıklandırmanın neden olduğu gölgelenmeler geliştirilen modellerin Mİ'yi tespit ve

tanımasını güçleştirmektedir. Mİ video karelerine uygulanacak önışleme teknikleri daha iyi tanıma doğruluğu sağlamada fayda sağlayacaktır [22]. Bu aşamada baş hareketlerinin ve arka plan etkisini kaldırmak amacıyla yüzün dışında kalan bölgeleri görüntüden çıkarmak için yüz bölgesini algılama, hizalama ve kırpma işlemleri uygulanmaktadır [13]. Mİ video klipleri yüksek kare hızında elde edildiği için benzer pozda birçok kare bulunmaktadır. Mİ tespit etmede yardımcı olmayan ve yanlış tahminlere yol açan kareleri kaldırıp tepe kareyi elde etmek için yüz algılama da Liu vd. [24] 66 noktan oluşan DRMF (Discriminative Response Map Fitting) yöntemini kullanmışlardır. Bir diğer çalışmada S.-J. Wang vd. [25] Uzun Süreli Evrişimsel Sinir Ağını Aktarma (TLCNN) adını verdikleri çalışmada görüntüleri ağa beslemeden önce, yüz bölgesi 66 nokta ile işaretlenmiş alanı içine alan bir dikdörtgen ile kırparak yüz alanı elde etmişlerdir. Yao vd. [26] çalışmalarında SEnet ve üç akışlı 3B ESA (SETFNet) modellerini tanıtırken, üç akışlı modele iki göz ve ağız bölgesi girdi olarak verilmektedir. Tüm yüz yerine bu üç yerel özelliğin kullanılması, gereksiz bilgilerden arındırılarak hesaplama maliyetini düşüreceği fikri ile yapılmıştır. Çalışmada bu bölgeleri algılayıp kırpma için DRMF yöntemi kullanılmaktadır. Hesaplama maliyeti ucuz olsa da baş hareketleri sonucu ön yüz açısının kapanması durumunda etkili değildir [27]. Li vd. [10] Mİ yakalama ve tanıma bölümlerinden oluşan çalışmalarında, Mİ tanıma kısmında yüz bölgelerini 68 nokta ile temsil eden aktif biçim modeli (Active Shape Model-ASM) yöntemini kullanarak orijinal görüntülerden yüz bölgelerini çıkarmışlardır. Huang vd. [4] uzamsal ve zamansal yüz temsiline dayalı Mİ çalışmalarında CASME-II ve SMIC veri setlerindeki video kare dizisinden yüz görüntülerini elde etmek için ASM yöntemini tercih etmişlerdir. Xia vd. [28] tarafından yapılan çalışmada yüz bölgeleri hizalanıp kırılmadan önce yüz bölgesini algılamak için ASM algoritması kullanılmıştır. Daha sonra tüm yüz bölgeleri ilk kareye göre Yerel Ağırlıklı Ortalama (Local Weighted Mean-LWM) algoritması kullanılarak hizalanır. Çift Zamansal Ölçekli Evrişimsel Sinir Ağı (DTSCNN) adlı verilen başka bir çalışmada, farklı kare hızlarına sahip Mİ video kliplerine uyum sağlamak için iki akışlı bir ağ modeli oluşturulmuştur. Ağ optik akış görüntü dizileriyle beslenmeden önce önışlem aşamasında aktif biçim modeli ile yüz bölgesi algılanıp, yerel ağırlıklı ortalama algoritması ile hizalanmaktadır [2]. Algılama sürecinden sonra

yüz bölgesinin doğru hizalanmaması durumundan Mİ tanıma performansı olumsuz etkilenebilmektedir [29]. Yüz bölgesini algılama, hizalama ve kırpma süreçlerinden geçen Mİ görüntüleri uzamsal alanda normalleştirilmektedir. Shreve ve Godavarthy yapmış oldukları çalışmada makro ifadelerin ve Mİ'lerin bulunduğu video görüntülerden yüzdeki gerilimden yüz ifadelerinin otomatik tespit edildiği çalışmada yüz bölgesini tespit etmek için Viola-Jones yüz dedektörü kullanılmıştır [30]. Viola Jones, Haar özelliklerinin hesaplanmasının ardından Adaboost algoritması kullanan eğitilmiş bir özellik öğrenme algoritmasıdır. He vd. çalışmalarında, CAS(ME)² ve SAMM veri setlerindeki uzun videolarda ifadeleri tespit ederken yüz bölgesini orijinal resimden çıkarmak için Dlib kütüphanesindeki araç seti kullanmışlardır [31]. Zhao vd. MERSiamC3D adı verdikleri 3B evrişime dayalı modellerinde video kare dizisi yerine, başlangıç, tepe ve bitiş anahtar karelerinde içinde bulunduğu 11 kareden oluşan özet Mİ kare dizilerini çalışmalarında kullanmışlardır. Bu çalışmalarında ön işleme aşamasında OpenCv'ye ait Dlib kütüphanesinin yüz dedektörü fonksiyonu kullanılmıştır [13]. VGG-f ESA modeli ve transfer öğrenmenin (ImageNet) kullanıldığı çalışmada ise yüz bölgesi Haar Cascade dedektörü tarafından algılanıp kırpılmıştır [3].

Literatürdeki bazı uzamsal ve zamansal özelliklerin birlikte çıkarıldığı Mİ çalışmalarında algılama, hizalama ve kırpma süreçlerine ek olarak farklı bazı ön işleme teknikleride uygulanmaktadır. Bunlardan Eulerian Video Büyütme (EVB) tekniği [32] videolardaki ince hareketleri daha belirgin hale getirdiğinden ve renk değişimini güçlendirdiğinden dolayı Mİ çalışmalarında da kullanılmaktadır. Zhao ve Xu sığ evrişimli sinir ağlarını oluşturdukları bileşik veri seti (CASME I, CASME II, CAS(ME)², SMIC, SAMM) örneklerinden başlangıç ve tepe kareden elde edilen optik akış görüntüleri ile besledikleri çalışmalarında Mİ'lerdeki zorluklardan olan ince genlik ve kısa süre sorununu çözmek için EVB yöntemini kullanmışlardır [11]. Cai vd. çalışmalarında kullandıkları Squeeze ve Excitation ağını bir 3B DenseNet ile birleştirmişlerdir. 3B SE-DenseNet adı verdikleri bu modelin SMIC veri setindeki deneylerinde EVB ön işleme tekniği kullanıldığında performansında %7.93 artış ile %92.68'lik bir başarı elde etmişlerdir [33]. Zamansal Enterpolasyon Modeli

(Temporal Interpolasyon Modeli-TIM) de video uzunluklarının yetersiz olduğu durumlarda, videoların uzunluklarını arttırmak için kullanılan bir yöntemdir. Bu yöntem ile video görüntü karelerinin arasına, önceki ve sonraki kareler referans alınarak, yeni kareler eklenmektedir [3,10]. Li vd. SMIC, CASME, CASME II veri setlerinin kullanıldığı 3B akış tabanlı ESA modellerini tanıttıkları bir çalışma yapmışlardır. Bu çalışmalarında gri tonlamalı görüntüler, yatay ve dikey optik görüntüler zamansal enterpolasyon modeli kullanarak 10 ve 15 kare enterpolasyon yapıldıktan sonra modele giriş olarak verilmiştir [34]. Etkili özellik bilgisi sağlayan Tek Yönlü Degrade Histogramı (HSDG) ve LBP-TOP yöntemlerinin birleştirildiği başka bir çalışmada da ön işlem aşamasında eşit olmayan video kare sorununu çözümünde zamansal enterpolasyon modeli kullanılırken, videodaki kas hareketlerini büyütme içinde EVB tekniği kullanılmıştır [35].

İncelenen literatürün hemen hemen tamamında önışlemi tamamlanan video karelerinde geleneksel, veriye dayalı veya hibrit yöntemler ile yapılan çalışmalarda, önemli özellikleri belirlemek [29], bu arada özelliklerin boyutunu azaltıp iyi bir tanıma performansı elde etmek için [13,32] özellik çıkarma işlemine geçilir [22]. Görüntülerin özelliklerini çıkarma aşamasında en çok kullanılan geleneksel yöntemler Optik Akış (Optical Flow-OF), Üç Ortogonal Düzlemin Yerel İkili Deseni (Local Binary Pattern of Three Orthogonal Planes- LBP-TOP) ve Gabor olarak öne çıkmaktadır. İki video karesi arasındaki piksel farkından hareket değişimini algılayan optik akış hemen hemen birçok Mİ çalışmasında sıkça görülmektedir. 11 anahtar kareden oluşan özet bir Mİ kare dizisi oluşturularak yapılan MERSiamC3D adlı çalışmada da 11 karenin optik akış görüntüleri çıkarılırken klasik Farneback yöntemi hızlı olduğu için tercih edilmektedir [13]. Zhou vd. çalışmalarında yatay ve dikey optik akış görüntüleri ile beslenen çift akışlı bir Inception ESA modeli kullanmışlardır. Çalışmalarında optik akış görüntülerini elde etmek için TV-L1 algoritması, diğer geleneksel yöntemlerden daha etkili ve pek çok çalışmadaki başarısı dikkate alınıp seçmişlerdir [37]. Liong vd. tepe karesini ve referans karesi olarak başlangıç çerçevesini kullanarak Mİ özelliklerini çıkarmanın yeterli olduğunu gösterdikleri çalışmalarında yön, büyüklük ve optik gerinim görüntülerini kullanan Çift Ağırlıklı Yönlendirilmiş Optik Akış (Bi-

Weighted Oriented Optical Flow-Bi-WOOF) tekniğini önermişlerdir [12]. Diğer bir yöntem ise, gri görüntülerden doku özelliklerini elde etmek için yerel ikili desenler (Local Binary Pattern-LBP) kullanılır. Yerel ikili desenler görüntüye ait sadece uzamsal özellikleri çıkarırken üç ortogonal düzlemin yerel ikili deseni görüntüye ait hem uzamsal hem de zamansal özellikleri çıkarmaktadır. Wang vd. Mİ video kliplerini ilk iki boyut uzamsal, üçüncü zamansal ve dördüncü ise renk bilgisinin bulunduğu dördüncü dereceden bir tensör olarak ele almışlardır. Çalışmalarında üç katmanlı RGB görüntülerini dört katmanlı renk bileşeni TICS görüntülerine dönüştürmüşlerdir. Çünkü dördüncü katmandaki dinamik doku ve bağımsız renk bileşeni daha fazla doğruluk sağladığını belirtmişlerdir. Sonrasında görüntülerin özelliklerini üç ortogonal düzlemin yerel ikili deseni doku tanımlayıcısı kullanarak çıkarılmıştır [38]. Üç ortogonal düzlemin yerel ikili deseni tanımlayıcısı XY, XT ve YT düzlemlerinde görüntülerin doku özelliklerini çıkartmaktadır. Y. Wang vd. Bu düzlemlerde çok fazla bilginin tekrarlandığını ispatlayıp 12 noktadan kesişen altı benzersiz noktayı çalışmalarında kullanarak yüksek boyuttan kaynaklı hesaplama yükünü normal üç ortogonal düzlemin yerel ikili deseni tanımlayıcısına göre 38 kat düşürmüşlerdir [39]. Lin vd. Mİ görüntülerine ön işlem aşamasında EVB yöntemi ile videodaki küçük Mİ hareketlerini büyütme işlemi gerçekleştirdikten sonra düzlemsel doku özelliklerini ve dokuların hareketini aynı anda hesaba katan uzamsal-zamansal bileşenden oluşan Gabor filtrelerini önermişlerdir. Nihai özellikler boyut küçültmek için uzamsal zamansal maksimum havuzlamadan geçirildikten sonra sınıflandırma aşaması için doğrusal bir DVM'ye girdi olarak verilip ifadeler sınıflandırılmıştır [40]. Zhang vd. Mİ tanıma sistemi adı verdikleri çalışmada ifade özelliğini çıkarmak için gabor dalgacık filtresi kullanırken boyut küçültme için temel bileşen analizi (Principal Components Analysis-PCA) ve doğrusal diskriminant analizi (Linear Discriminant Analysis-LDA) sınıflandırma işleminde de Destek Vektör Makinesi (DVM) kullanılmıştır [41].

Mİ tanıma görevinde, görüntülerin özelliklerini çıkarma aşamasında geleneksel yöntemlerinin yanı sıra veriye dayalı derin öğrenme yöntemleri kullanılmaktadır. Bu yöntemlerde kullanılan ağların çok sayıda parametreyi öğrenmesi gerektiğinden

modelin performansını arttırabilmek için eğitimde kullanılacak örneklerin arttırılması, sınıfların dengeli dağılması gerektiği belirtilmiştir [21]. Örneklemelerin başarılı bir öğrenme için yetersiz olduğu durumlarda problemin üstesinden gelmek için daha önce farklı problemler için başarılı olmuş öğrenme ağırlıklarının yeni duruma aktarılması olan transfer öğrenme yöntemleri kullanılmıştır [37,38,39,40,41]. Bazı çalışmalarda ise veri arttırma tekniği kullanılarak veri setindeki örnekleme sayısını arttırılmıştır [37,38,42,43]. Mİ görüntülerinden derin öğrenme mimarisi kullanılarak hem özellik haritaları çıkarılabilir [37,38] hem de yapılarında bulunan softmax aktivasyon fonksiyonu sayesinde çıkarılan özellikler sınıflandırılabilir [20,44,45]. Özellik çıkarmada kullanılan iki katmandan ilki piksel ilişkilerini koruyarak önceki katmandan gelen özellik haritalarındaki yerel komşuluktan özellik çıkaran evrişim katmanı ve ikincisi önemli bilgileri koruyarak sürekli boyut küçültme işlevini gerçekleştiren havuzlama katmanıdır [4,38]. 2B ESA'lar yalnızca uzamsal boyutlardan 2B özellikleri çıkarırken, 3B-ESA hem uzamsal hem de zamansal boyutlardan özellikleri çıkarmaktadır [4]. Peng vd. CASME ve CASME II veri setlerini kullandıkları çalışmalarında veri setlerindeki video kliplerinin farklı kare hızlarına adapte olabilen biri 64fps diğeri 128 fps iki akışlı Çift Zamansal Ölçekli Konvolüsyonel Sinir Ağı (DTSCNN) adı verdikleri 3B ESA modeli kullanmaktadır. Modelin yapısında aşırı uydurmadan kaçınmak için sığ bir ağ yapısı tasarlanırken, sınıflandırma kısmında doğrusal bir DVM'den yararlanmışlardır [2]. Gan vd. Mİ tanıma üzerine yaptıkları çalışmada Off-ApexNet adı verdikleri modeli tanıtmışlardır. Örneklemeye sayısının yetersizliğinden dolayı iki evrişim iki havuzlama katmanının olduğu sığ bir iki akışlı ESA model önermişlerdir. Model iki kare arasındaki piksel farkından ince hareketleri belirleyen TV-L1 optik akış yöntemine ait yatay ve dikey görüntüleri ile beslenmiş, softmax aktivasyon fonksiyonundan faydalanılarak görüntü özellikleri sınıflandırılmıştır [29]. Wang vd. Mİ video klip karelerinden özellik çıkarmak için derin ESA ve zamansal dizi bilgisini öğrenmek için Uzun Kısa Süreli Bellek (LSTM) birleştirdikleri TLCNN adlı bir model tasarlamışlardır. Küçük örnekleme sahip veri seti üzerinde iyi performans elde edebilmek için ise transfer öğrenme kullanılırken, sınıflandırma aşamasında softmax aktivasyon fonksiyonundan faydalanılmıştır [25]. Sınıflandırma sürecinde ise softmax aktivasyon fonksiyonundan başka kullanılan bir

diğer yöntem ise, makine öğrenmesi için büyük öneme sahip, DVM ve çekirdekleridir [10,12,34,46,47]. Uzun vd. beş farklı ESA modelini kullandıkları çalışmada (VGG16, AlexNet, SqueezeNet, MobilNetV2, EfficientNetB0) tam bağlı katmanlarından 1000'er adet özellik çıkarıp hem verimli özellikleri elde etmek hem de boyut küçültmek için PSO algoritması kullanmışlardır. Sınıflandırma sürecinde doğrusal DVM yanı sıra, Gaussian, kübik ve ikinci dereceden DVM çekirdeklerinin karşılaştırıldığı bir çalışma yürütmüşlerdir [14]. Mikro ifade özelliklerini çıkarmak için Üç Ortogonal Düzlemin Çift Çapraz Desenleri (Dual-cross patterns-DCP-TOP) ve Üç Ortogonal Düzlemin Sıcak Tekerlek Desenleri (Hot Wheel Patterns-HWP-TOP) ikili tanımlayıcılarının karşılaştırılması kullanıldığı çalışmada, HWP-TOP tanımlayıcısı ve düz (smooth) DVM sınıflandırıcısı ile 0.868'lik bir Mİ tanıma doğruluğu elde edilmiştir [44]. Sun vd. önerdikleri modellerinde Openface araç setini kullanarak veri集中的 örnekler için yüz bölgesini çıkarmışlardır. Daha sonra özellik çıkarma görevi için artık ağ ve sınıflandırma görevi için DVM kullandıkları modellerini, FERA-2017 makro ifade veri seti ile ön eğitime tabi tutmuşlardır. Ardından SMIC, CASME, CASME II ve SAMM veri setleri kullanarak yaptıkları deneylerde en yüksek performansı dört sınıfa ayırdıkları SAMM veri seti üzerinde 0.8674 doğruluk ile elde etmişlerdir [53]. Mayya vd. önerdikleri modellerinde ESA ve doğrusal DVM kullanmışlardır. Mİ video dizilerini zamansal enterpolasyon (TIM) kullanarak enterpolasyona tabi tuttukları çalışmalarında, SMIC veri seti için 0.7805 doğruluk performansı elde etmişlerdir [54]. Takalkar vd. üç ortogonal düzlemin yerel ikili deseninden elde ettikleri ve ESA'dan elde ettikleri Mİ özelliklerini birleştirerek softmax ve DVM sınıflandırıcılarında eğitmişlerdir. Çalışmalarında DVM'nin doğrusal, polinom, radyal ve sigmoid gibi dört farklı çekirdeğinin kullanılmıştır. DVM, CASME II, SMIC-NIR, CAS(ME)² veri setleri için softmax sınıflandırma algoritmasından %4 ila %7 oranında üstün olduğunu gözlemlemişlerdir[55].

BÖLÜM 3

YÜZ İFADELERİ VE YÜZ HAREKETİ KODLAMA SİSTEMİ

İnsan duyguları uzun yıllardır birçok araştırmacı için merak konusu olmuştur. İnsan duygularını incelemek için yüz ifadelerindeki değişimler incelenmiştir. Yüz ifadeleri insan iletişimde, sözel olmayan birincil araç olarak kabul edilmektedir. Bununla birlikte, yüz ifadeleri sözel iletişimi yüz ifadeler ile birlikte kullanarak vurgulayıp desteklemekte, duygu iletiminde ve niyet belirtmede de sıkça kullanılmaktadır [56]. Duygu ve niyet tanıma güvenlik, klinik, eğitim, robotik gibi alanlarda uygulama geliştirme açısından önem arz etmektedir. Yüz ifadeleri iletişimde kullanıldıkları için bir mesaj içermektedir. Bu mesajlardaki yargıyı ölçmek ve mesajı ileten işaret araçlarını ölçmek olmak üzere iki tür ölçümden bahsedilmektedir [57]. Yargıdan kasıt edilen mutluluk, üzüntü, sinir gibi mesajın anlamını doğrudan çözmeyi amaçlayan bir yoruma, bir yargıya varmaktır. İşaret temelli yaklaşım ise hiçbir yorum ve çıkarım yapmadan dudak kenarları yukarı doğru gibi mesajı iletmek için kullanılan fiziksel sinyali incelemeyi amaçlar, yorum yapılmaz sadece kodlanır. İletişimde yüz ifadelerinin önemi nedeniyle, yüz ifadelerini ölçmeye yönelik Ekman ve Friesen 1978’ de 44 eylem biriminden oluşan Yüz Hareketi Kodlama Sistemi (YHKS) adı verdikleri bir tanımlayıcı geliştirmişlerdir [58]. YHKS yüz hareket tespitinde en etkili ve yaygın değerlendirme yöntemi olan anatomik ve bir çok kişi tarafından standart ölçü olarak kabul edilen bir sistemdir [59]. Bu sistem yüzün görsel olarak ayırt edilebilir ince özellik değişikliklerini tespit etmek için tasarlanmış, insan gözlemci ve işaret tabanlı bir sistemdir. Videoya kaydedilmiş yüz hareketinin anatomik yapısı analiz edilerek, farklı kas hareketlerinden kaynaklı görünüm farklılıklarını içermektedir. Eylem birimi olarak adlandırılan yüzdeki 30 kas eylemi (18 alt ve 12 üst) ve 14 adet baş duruşu bakış gibi eylem tanımlayıcısı bulunmaktadır. Eylem birimleri bir ya da daha fazla yüz kasının hareketi sonucu belirli yüz yapılarını tanımlayan duyguların yorumlanmasından bağımsız kodlardır. Eylem birim ifadeleri

yoğunlukları minimum yoğunluğu tanımlayan A ve maksimum yoğunluğu tanımlayan E harfler ile birlikte kullanılmaktadır. Şekil 3.1’de yüzün üst ve alt kısmındaki hareketleri tanımlayan 30 eylem birimi gösterilmektedir. Bu eylem birimleri içerisinde üst kısımda yer alan gözler, göz kapakları, kaşlar ve elmacık kemiklerinin hareketleri incelenirken, alt kısımda ise burun, ağız, çene, dudaklar, yanaklar ve nazolabial denilen burun deliklerinin yan kısmından başlayıp dudak köşelerine kadar ilerleyen çizginin hareketleri incelenmiştir.

ÜST YÜZ EYLEM BİRİMLERİ					
AU 1	AU 2	AU 4	AU 5	AU 6	AU 7
					
göz kapağı düşmüş	göz kapakları kısık	gözler kapalı	gözler kısık	gözler kırık	tek göz kırık
ALT YÜZ EYLEM BİRİMLERİ					
AU 9	AU 10	AU 11	AU 12	AU 13	AU 14
					
					
					

Şekil 3.1. YHKS’de bulunan yüz eylem birimleri [59].

3.1. EVRENSEL DUYGULAR

Paul Ekman'a göre evrensel olarak tanımlanmış altı temel (birincil) duygu mutluluk, öfke, korku, iğrenme, üzüntü ve şaşkınlık yüz ifadeleri aracılığı ile iletilirler. Bu duygular evrensel ve kültürlerden bağımsız olarak meydana gelmektedirler [59]. Bunların anlamlandırılması mesaj yargılama vasıtası ile ölçülmektedir [56]. Duygulara ait yüz ifadeleri birden çok kas hareketinin kombinasyonundan oluştuğu için, YHKS ile ifade edildiğinde birden fazla eylem birimleri ile tanımlanmaktadır. Çizelge 3.1'de evrensel altı temel duygu ve bunlara ait eylem birimleri verilmiştir.

Çizelge 3.1. Altı temel duyguya ait eylem birimleri [56].

Duygular	Eylem Birimleri
Mutluluk	6, 12, 25
Öfke	4, 5, 7, 10, 17, 22-26
Korku	1, 2, 4, 5, 20, 25, 26, 27
İğrenme	9, 10, 16, 17, 25, 26
Üzüntü	1, 4, 6, 11, 15, 17
Şaşkınlık	1, 2, 5, 26, 27

3.2. BİLGİSAYARLI GÖRÜ VE YÜZ İFADELERİ

YHKS alanında uzman yetiştirmenin 100 saatten fazla eğitim süresi alması, her bir mikro ifade klibinin bir dakikasının incelenip etiketlenmesi için gerekli sürenin bir saatten fazla olması gibi yüksek maliyetler YHKS analizinin yaygınlaşmamasındaki en büyük dezavantaj olarak görülmektedir [56]. Bundan dolayı bilgisayarlı görme, makine öğrenmesi gibi yapay zekâ alanlarında yapılan çalışmalar bu zorlukların üstesinden gelmek için araştırmacılara çalışma imkânı sunmaktadır. Makine öğrenmesi ve bilgisayarlı görü araştırmalarında kullanılmak üzere oluşturulan veri setlerinde kayıt sırasında öğrenmeyi zorlaştıracak durumlar şunlardır; kafa hareketi, duruş, konuşma, sahne karmaşıklığı ve arka plan, deneyler arasındaki bireysel farklılıklar, görüntü alma, çözünürlük, aydınlatma ve gölgelendirme olarak

sıralanabilir. Bu zorlukların üstesinden gelmek için bazı çalışmalarda veriler ön işleme aşamasından geçirildikten sonra kullanılmakta ve bu da öğrenmeyi kolaylaştırmaktadır. Ayrıca doğal olmayan pozlanmış mikro ifade görüntüleri ve doğal mikro ifade görüntüleri zamansal ve yüz eylem birimleri birlikteliklerinde farklılıklar göstermektedir [56]. Pozlanmış yüz görüntüleri istemli hareketlerin kontrol edildiği beynin motor korteks tarafından kontrol edilirken, doğal yüz görüntüleri ise istemsiz hareketlerin kontrol edildiği beynin subkorteks denilen bölgesinden kontrol edilmektedir [52,53].

Bilgisayarlı görü, yüz ifadelerini tespit ve tanıma alanında istenilen düzeye geldiğinde yalan makinesi gibi aldatmaya yönelik durumların tespiti yerine gelecekte kullanılabilecektir. Aldatmaya yönelik sözel olmayan ipuçları, öncelikle insanın düşünce yapısında yani bilişsel davranışında belirirken sonrasında ise beklenmeyen durumlarda kontrolün kaybedilip duygusal davranışlardan sızmasıyla sözel olmayan ipuçları verdiği kuramsallaştırılmıştır [62]. Her ne kadar bastırılmış duygular vücut hareketleri, yüz ifadesi veya ses tonu gibi sözsüz olarak farklı yollarla dışa vurulduğu belirtilse de sözsüz iletişimin çoğu yüz ifadeleriyle gerçekleşmektedir [63]. Duyguların bastırılıp kontrol edilmeye çalışılması kültürel ya da toplumsal nezaket kuralları olarak küçük yaşta öğrenilse de bu ifadelerin sızıntısı otomatik olarak “güvenilir” kaslar olarak belirtilen yüz kasları vasıtası ile ya da Mİ’ler yoluyla gerçekleşmektedir. Bastırılmış duyguların sızıntı şeklinde yüz ifadesi olarak iki şekilde meydana gelmektedirler. Bunlardan ilki olan ince ifadeler ilişkili kas sisteminin yalnızca bir kısmının istemsizce kullanılması sonucu ortaya çıkmaktadır. İkinci olarak yüzde görülen sızıntı şeklinde ortaya çıkan ifadelerden Mİ ise çok kısa zamanda meydana gelip ilişkili kas sisteminin tamamında görülmektedir. Fakat o kadar kısa zaman diliminde gerçekleşir ki YHKS eğitimi ile doğru kodlansa da, gözlemcilerin çoğu bu ifadeleri algılayamamaktadır [64]. Bundan dolayı bilgisayarlı görü yüz ifadelerini ve gerçek duyguları anlamada bir çok alanda önemli bir yer alacaktır.

3.3. MİKRO İFADELER

Yüz ifadeleri, duyguları yansıtan bir araç olarak ifade edilmektedir [65]. Yapılan araştırmalar yüzde meydana gelen temel duygu ifadelerinin cinsiyetler ve tüm kültürler için evrensel olduğu belirlenmiştir [66]. Duygunun üretimi, ifadesi, deneyimi veya tanınması gibi durumlarda otonom sinir sisteminin rolü kritiktir [67]. Fakat bazı psikolojik araştırmalar, kişiler gerçek duygularını bilinçli şekilde bastırdıklarında, yüzünde otonom sistemi tarafından kontrol edilemeyen Mİ'lerin oluştuğunu göstermektedir[55,60].

Haggard ve Isaacs tarafından 1966'da psikoterapi seanslarının filmlerini taranırken keşfedilmiş ve onlara mikro anlık yüz ifadeleri adını verdi. Ardından 1969'da Ekman ve Friesen tarafından psikiyatri hastaları üzerinde video kayıt yaparak deneyler yapıldı. Hastalardan biri video boyunca mutlu görünse de ağır çekimde (1/12s), hastada acı duygusu fark edildi. Sonraki seanslarda hastanın itirafıyla intihar etme planını gizlemek için yalan söylediği belirlendi. Ekman ve Friesen tarafından ağır çekimde deneklerin fark edilen bu ve diğer duygulara Mİ adı verilmiştir.

Ekman'a göre yüzde meydana gelen altı temel duygu ifadesi olan korku, hor görme, şaşırma, tikslenme, üzüntü, öfke, mutluluk gibi farklı duyguları içeren Mİ'lerin ortak özellikleri, düşük yoğunluklu, kısa zamanlı, istemsiz, riskli koşullar altında ve kontrolsüz meydana gelen bastırılmış duygulardır. Mİ'ler ayrıca gerçek duyguları gösteren doğal ve tamamlanmamış parçalı hareketlerdir [69].

Mİ'ler sınır güvenliği, kolluk kuvvetleri, siyasi, eğitim, klinik gibi alanlarda kullanılması oldukça önem arz eden bir alan olarak görülmektedir. Mİ tespit ve tanıma konusunun önemi hakkında örnekler verilecek olursa, 1938'de Hitlerin İngiltere başbakanına Çekoslovakya'yı işgal etmeyeceği yalanını söylemesi veya psikopatların şartlı tahliyesinin diğer mahkumlara göre 2,5 kat fazla olması bu alanda elde edilecek bilgilerin ne kadar önemli olduğunu göstermektedir [70]. Mİ'ler çok hızlı meydana geldiklerinden dolayı insanlar tarafından kolayca gözden kaçırılabilir [39]. Mİ'lerin eğitim almadan, uzmanlaşmadan tespit edilmesinin ve tanınmasının çok zor

olduđu, başarı oranlarının düşük olduđu belirtilmiştir [66]. Beş Mİ'yi tanıma oranı eğitimsiz insanlar için %32 iken, eğitimli uzman kişiler tarafından tanıma oranı ise %47 olarak tespit edilmiştir [7]. Buda verimli seviyenin çok altındadır [10].

Mİ tanıma doğruluđu yeteneđini belirlemek için yapılmış olan ilk test olarak bilinen Japon ve Kafkas Kısa Duygu Tanıma Testi (JACBART) daha sonra Mİ tanınmasını Mikro ifade eğitim aracına internet üzerinde eğitim veren bir eğitim aracına dönüştürüldü [69]. Mikro ifade eğitim aracı duygu tanımda eksiklikleri olan klinik hastalarının eğitimi için de kullanılmıştır. Zaman ve maliyet açısından verimli olarak görülen mikro ifade eğitim aracına çeşitli iş alanlarında insanların Mİ algılama yeteneklerini geliştirmek için kullanılmıştır [71].

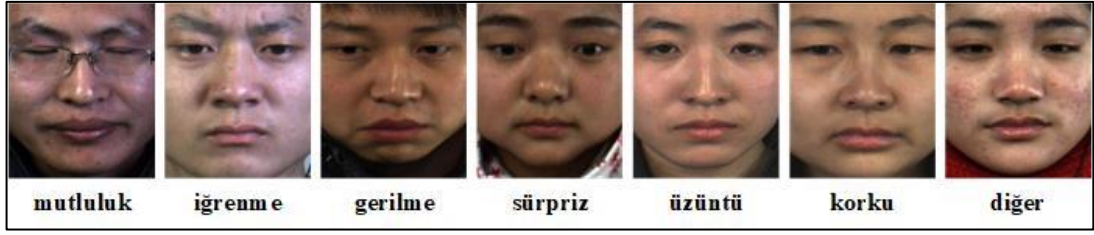
BÖLÜM 4

MATERYAL VE METOD

4.1. VERİ SETLERİ

Otomatik Mİ tanıma sistemlerinin geliştirilmesi için yeterli eğitim örneklerine ve yüksek kalite ile kaydedilmiş veri setlerine ihtiyaç duyulmaktadır. Mİ'lerin belirlenmesi için bu tez çalışması kapsamında önerilen modellerde kullanılan CASME-II, SAMM, SMIC veri setleri kamuya açık, en popüler ve en yeni veri setleridir.

CASME-II [72] veri setinde, 26 katılımcıdan oluşan 247 doğal Mİ örneği bulunmaktadır. Örnekler avi video klip, avi'den dönüştürülmüş jpg fotoğraf kare dizisi ve yüz bölgesi çıkarılarak ön işlem uygulanmış jpg fotoğraf kare dizisi formatında kullanıcılara sunulmaktadır. Bu örneklerle ait kayıtlar 200 fps ve 280x340 piksel yüz çözünürlüğüne ve uygun aydınlatma ortamına sahiptir. Mİ'ler YHKS kılavuzunda bulunan eylem birimlerine göre etiketlenmişlerdir. Ayrıca anahtar kare olarak adlandırılan başlangıç, tepe ve bitiş karelerinin konum indeksleri excel formatında kullanıcılara sunulmuştur. Şekil 4.1'de CASME-II veri setindeki “cropped” klasörü altında bulunan, önışlemden geçirilmiş, tüm sınıflara ait tepe kare örnekleri sunulmuştur.



Şekil 4.1. CASME-II veri setinde bulunan sınıflara ait tepe kare örnekleri.

SAMM [73] veri seti ise 16 erkek ve 16 kadından oluşan 32 katılımcı bulunmaktadır. 17 farklı etnik kökene sahip bu katılımcı grubunun yaşları 19 ila 57 arasında değişmektedir. Kayıtlar 200 kare hızına, 1040x1088 görüntü çözünürlüğüne ve bu görüntülerde bulunan yüz bölgeleri de yaklaşık olarak 400x400 çözünürlüğüne sahiptir. Örnekler yüz bölgesi çıkarılmamış, ön işlemsiz fotoğraf çerçeveleri (.jpg) formatında kullanıcılara sunulmaktadır. Mİ örnekleri YHKS kılavuzunda bulunan eylem birimlerine göre etiketlenmiştir. Anahtar karelerinin konum indeksleri excel formatında kullanıcılara sunulmuştur. Şekil 4.2’de SAMM veri setindeki tüm sınıflara ait tepe kare örnekleri sunulmuştur.



Şekil 4.2. SAMM veri setinde bulunan sınıflara ait tepe kare örnekleri.

SMIC [74] veri setinde yaş ortalaması 28.1, kadın ve 10 erkekten oluşan toplam 16 katılımcıdan, bunların sekizi kafkasyalı sekizi asyalı, elde edilen 164 tane spontane

Mİ örneği kaydedilmiştir. Kayıtlar 100 fps ve 640x480 çözünürlüklü yüksek hızlı (HS) kamera ile toplanmıştır. Ayrıca SMIC veri setinde 8'er katılımcıdan elde edilen 25 fps ve 640x480 çözünürlüklü, normal görüş kamerası (VIS) ve yakın kızıl ötesi kamerası (NIR) ile elde edilmiş video klip kayıtları da bulunmaktadır. Her iki kayıt türünde de 71 adet Mİ video klip yer almaktadır. Örnekler yüz bölgeleri çıkarılmış ve önilem yapılmamış ham fotoğraf kareleri (.jpg) formatında kullanıcılara sunulmaktadır. Anahtar karelerinin konum indeksleri kullanıcılara sunulmamıştır. Şekil 4.3'de SMIC veri setindeki "SMIC_all_cropped" klasörü altında bulunan, önilemden geçirilmiş, tüm sınıflara ait tepe kare örnekleri sunulmuştur. Ayrıca Çizelge 4.1'de bu veri setlerine ait özellikler gösterilmektedir.



Şekil 4.3. SMIC veri setinde bulunan sınıflara ait tepe kare örnekleri.

Çizelge 4.1. Spontane Mİ veri setlerine ait genel özellikler

Veri Seti	CASME II	SMIC-HS	SAMM
Katılımcı	24	16	28
Örnekler	154	157	133
Negatif	94	66	92
Pozitif	32	51	26
Sürpriz	28	40	15
Kare Hızları (fps)	200	100	200
Çözünürlük	640x480	640x480	2040x1088

Yüz Çözünürlüğü	280x340	190x230	400x400
Duygular	3	3	3
Sponta/Pozlanmış	Spontan	Spontan	Spontan
Etnik Köken	1	3	13

4.2. ÖN İŞLEME METODLARI

Bu bölümde tepe kare indeks numarası verilmemiş SMIC veri setinde tepe karenin tespiti yapılmıştır. Ardından SAMM veri setindeki örneklerde herhangi bir ön işlem aşamasından geçirilmemiş hali kullanıcılara sunulmadığı için yüz bölgelesinin çıkarılması işlemleri uygulanmıştır. Böylece tüm veri setlerindeki örneklerin arka plan gürültülerinden arındırılmış sadece yüz bölgelerinin bulunduğu görüntü verileri, başlangıç ve tepe kareleri elde edilmiştir. Başlangıç ve tepe kare görüntülerinden faydalanılarak optik akış görüntüleri elde edilmiş daha sonra veri setlerindeki örnekleme yetersizliğinden dolayı optik akış görüntüleri veri artırma tekniği ile yapay olarak artırılmıştır.

4.2.1. D&C-RoIs

Önerilen deneysel çalışmalarda Mİ video kare dizilerinden sadece en yüksek yoğunlukta Mİ içeren [75] “tepe kare” leri ve referans olarak “başlangıç kare” leri ince hareket değişimini tespit etmek için kullanılmıştır. Bu karelerin indeks konumları CASME II ve SAMM veri setlerinde bildirilirken, SMIC veri setinde bildirilmemiştir. SMIC veri setindeki tepe kareleri bulmak için yaklaşık değerini otomatik olarak bulan ilgili bölgeyi böl ve fethet tekniği D&C-RoIs (Divide and Conquer-Region of Interest) [76] geliştirilmiş ve iyi performansından dolayı Mİ çalışmalarında kullanılmıştır [16,25,73]. D&C-RoI yöntemi, özellik tanımlayıcı LBP ve böl ve fethet tekniklerini içerir. LBP, bir video dizisindeki her karenin yüz bölgelerinin özelliklerini hesaplar. Ardından başlangıç karesi referans alınarak her bir çerçevenin özellik farkı hesaplanır. Böl ve fethet algoritması, maksimum farka sahip tepe karelerin konumunu (dizin) belirler. Her çerçevedeki RoI bölgeleri hesaplanıp, ilk kare ve diğer kareler arasındaki korelasyon LBP tanımlayıcı ile hesaplanarak tepe karesi elde edilmiştir [76]. LBP

özelliklerinin farkına dayalı D&C-RoIs yönteminde kullanılan matematiksel eşitlik denklem 4.1’de verilmiştir [20].

$$d = \frac{\sum_{i=1}^B h_{1i} x h_{2i}}{\sqrt{\sum_{i=1}^B h_{1i}^2 x \sum_{i=1}^B h_{2i}^2}} \quad (4.1)$$

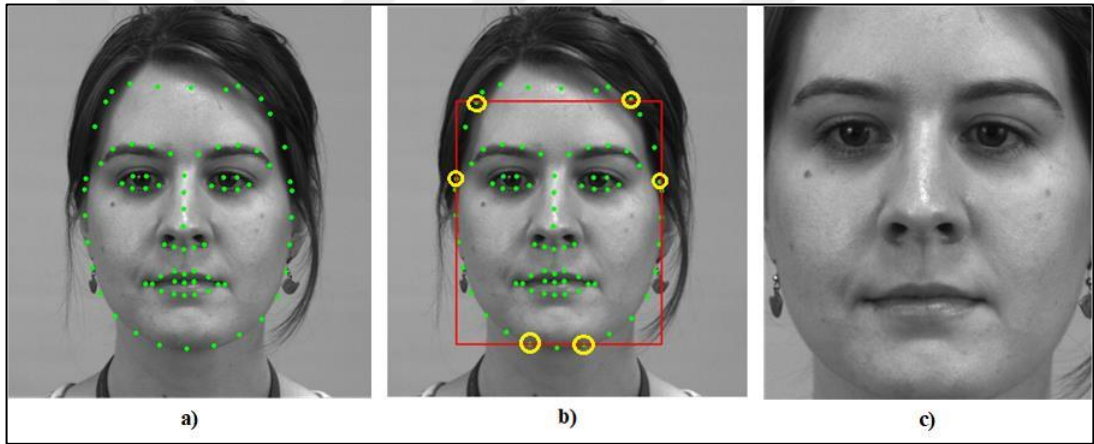
Burada ilk kare (h1) ile diğer karelerin (h2) korelasyonları alınmaktadır. B ise h1 ve h2 histogramlarındaki kutu sayısıdır. LBP özelliklerinin fark oranı (1-d) daha sonra Mİ tanımada en etkili üç RoI kaş, göz ve ağız arasında karşılaştırılır ve en yüksek fark oranına sahip RoI seçilir. Son olarak, maksimum yüz kas değişiklikleri ile tepe kare aramak için bir böl ve yönet stratejisi uygulanır [16,72]. Böl ve fethet stratejisine göre, bir video klip kare dizisi alt dizilere bölünür. Çerçevelerin korelasyon katsayıları her bir alt dizide toplanır. Toplamı en yüksek olan indeks tutulur ve geri kalanlar atılır. Maksimum değere sahip çerçeve bulunana kadar devam eder [12].

4.2.2. Yüz Bölgesinin Çıkarılması

CASME-II ve SMIC örnekleri önışlem yapılmış biçimleri veri setinde bulundurulduklarından bunlara herhangi bir önışlem yapmaya gerek duyulmamıştır. Fakat SAMM örnekleri herhangi bir önışleme aşamasından geçmediğinden dolayı görüntülerin optik akış özelliklerini çıkarmadan önce SAMM veri setinde ön işlem gerçekleştirilmiştir. SAMM veri setlerindeki örneklere OpenCv’nin Dlib kütüphanelerinde bulunan fonksiyonlar kullanılarak sırasıyla yüz algılama, hizalama, kesme ve kaydetme işlemleri uygulanmıştır. Dlib kütüphanesindeki yüz algılama fonksiyonu CNN tabanlı bir maksimum kenar boşluğu nesne dedektörü kullanmaktadır.

Mİ ifadeler yüzün belirli bir yerinde meydana geldiği için bu bölgelerin dışındaki alanlar (saç, boyun, arka plan, kulak vs.) Mİ tanıma görevinde performans kaybına neden olabilir. Bundan dolayı yüzün belirli bölgelerine odaklanmak ve alakasız bölgelerden kurtulmak için ön işleme görevi yapılmaktadır. Python Dlib

kütüphanesinde bulunan dedektör fonksiyonu kullanılarak algılanan yüz bölgesi ve aynı kütüphanede bulunan 81 nokta ile yüz bölgesini tahmin eden eğitilmiş model ile tespit edilmiştir. Yüz bölgeleri, Dlib kütüphanesinde bulunan dedektör fonksiyonu ile algılandıktan sonra, yine Dlib kütüphanesinde bulunan, 81 nokta ile yüz bölgesini tahmin eden eğitilmiş model kullanılmıştır. Şekil 4.4’de gösterilen yüz bölgesi, üstte alnın sağ ve sol kısımlarında çember içindeki noktaların ortalamasına denk gelen y_1 konumu, altta çenenin alt kısmında çember içindeki sağ ve sol noktaların ortalamasına ait y_2 konumu, sağ ve sol gözün dışa doğru bölgesinde bulunan çemberler içindeki noktalar x_1 ve x_2 konumlarına göre bir dikdörtgen çerçeve oluşturulduktan sonra yüz bölgesi kesilmiş ve kaydedilmiştir. Şekil 4.4’de yüzün tespiti, yüzün işaretlenmesi, yüzün çerçeveye alınması ve yüzün kırılması süreçleri gösterilmiştir.



Şekil 4.4. a) Yüzün tespiti, yüzün işaretlenmesi, b) yüzün çerçeveye alınması c) yüzün kırılması.

4.2.3. Optik Akış

Optik akış, iki kare arasındaki ince hareketleri, küçük değişimleri piksel değişimine göre yakalayabilen ve görüntüdeki önemli özellikleri içeren bir yöntemdir [36]. Bu yöntem hareket tanıma, hareket algılama, nesne takibi gibi alanlarda kullanılmaktadır. Farneback [78] ve Total Variation-L1 (TV-L1) [79] tekniği her noktanın optik akışını tespit eden yoğunluk tabanlı bir yöntemdir. Farneback yöntemi hızlı hesaplama,

piramidal ayrıştırma ve yüz hareketlerinin analizinde nispeten daha az hata üretmesi gibi avantajlarından dolayı tercih edilmektedir [80]. Bundan dolayı Mİ çalışmalarında da kullanılmaktadır [13,77]. Farneback optik akış, ikinci dereceden polinom ile polinom genişlemesini kullanarak her pikselin yaklaşık komşuluklarını tahmin etmektedir. Polinom genişletme ortogonal koordinat sisteminin iki boyutlu sinyal uzayının temel bir fonksiyon uzayına $(1, x, y, x^2, y^2, xy)$ dönüştürülmesidir. Böylece görüntü ve görüntü alanı $r_1 + r_2x + r_3y + r_4x^2 + r_5y^2 + r_6xy$ altı katsayı ile ikinci dereceden denklemler ile temsil edilebilir [82]. Farneback optik akışa ait matematiksel denklemler Eşitlik 4.2 - 4.6 arasında verilmiştir [78].

$$(x, y)A \begin{pmatrix} x \\ y \end{pmatrix} + b^T \begin{pmatrix} x \\ y \end{pmatrix} + c$$

$$f_1(x) = x^T A_1 x + b_1^T x + c_1 \quad (4.2)$$

Burada $f(x)$, yerel bir koordinat sistemindeki x 'nın gri piksel koordinat vektörüdür,

$A = \begin{bmatrix} r_4 & \frac{r_6}{2} \\ \frac{r_6}{2} & r_5 \end{bmatrix}$ simetrik bir matristir, $b = [r_2 \ r_3]$ bir vektördür ve $c_1 = r_1$ bir skalerdir. Yeni

bir f_2 sinyali, global bir öteleme d kullanılarak oluşturulur.

$$f_2(x) = f_1(x - d) = (x - d)^T A_1 (x - d) + b_1^T (x - d) + c_1$$

$$= x^T A_1 x + (b_1 - 2A_1 d)^T x + d^T A_1 d - b_1^T d + c_1 \quad (4.3)$$

$$= x^T A_2 x + b_2^T x + c_2$$

Eşitlik 4.4 de ikinci dereceden polinomlardaki katsayılar eşitlenirse Eşitlik 4.5 ve Eşitlik 4.6 eşitlikleri elde edilir.

$$A_1 = A_2 \quad (4.4)$$

$$b_2 = b_1 - 2A_1 d \quad (4.5)$$

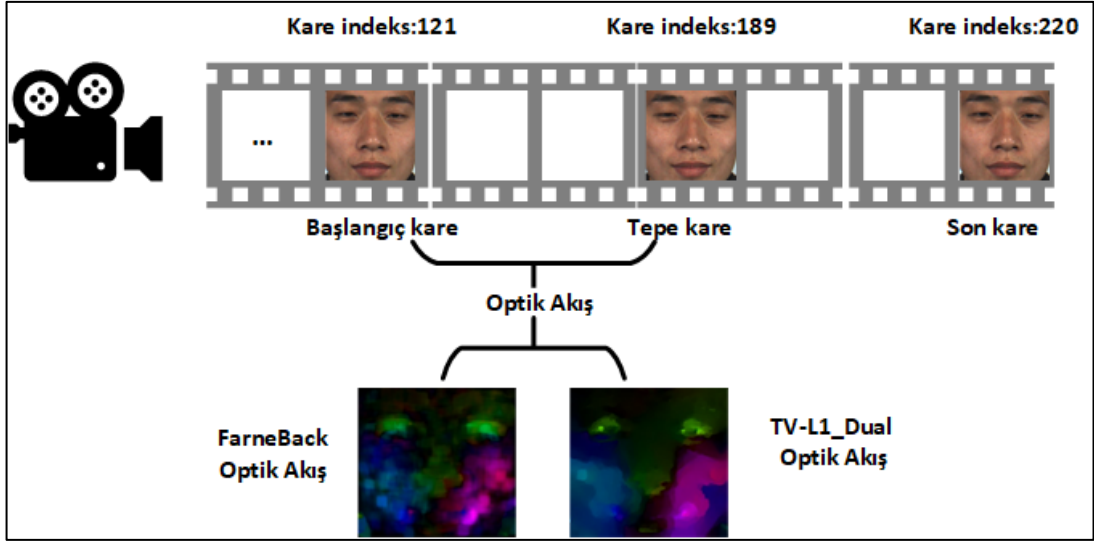
$$d = -\frac{1}{2} A_1^{-1} (b_2 - b_1) \quad (4.6)$$

TV-L1 yöntemi görüntülerin süreksizliğini tahmin etme ve kenar özelliklerini koruma, L1 normunu kullanma gibi avantajlarından dolayı tercih edilmiştir [11,20]. L1 formu sayesinde nesnelerin hareketlerinden kaynaklı parlaklığının değişmezliği ile gölgeler, vurgular, aydınlatma gibi gürültüler ihmal edilmektedir [25,32]. Ayrıca diğer bazı geleneksel optik akış yöntemlerinden daha verimli ve sağlam olduğunu kanıtlamıştır [83]. Bundan dolayı bir çok Mİ çalışmasında kullanılmıştır [11,20,25,32]. Şekil 4.5’ de bir video dizisinde başlangıç, tepe, bitiş kareleri ve bu karelerden optik akış görüntülerinin elde edilmesi süreçleri gösterilmektedir. Ayrıca iki görüntü arasındaki yerel türevlere dayanan bir yaklaşıma sahip olan yöntemin matematiksel ifadesi Eşitlik 4.7 ve 4.8 arasında verilmiştir [79].

$$E = \int_{\Omega} \left\{ \lambda \phi \left(I_{onset}(x) - I_{apex}(x + u(x)) \right) + \psi(u, \nabla u, \dots) \right\} dx \quad (4. 7)$$

Burada optik akış için kullanılan iki kare I_{onset} ve I_{apex} : $(\Omega \subseteq \mathbb{R}^2) \rightarrow \mathbb{R}$ verilmiştir. X , bir düzenleştirme kuvveti ile birlikte görüntü tabanlı bir hata kriterini en aza indiren bir eşitsizlik haritasıdır. ϕ , $(I_{onset}(x) - I_{apex}(x + u(x)))$ formülünde, görüntü verilerinin aslına uygunluğu, $\psi(u, \nabla u, \dots)$ önceki şekli tetikleyen düzenleştirme terimini belirtir, λ veri ögesinin ağırlık sabitidir. $\phi(x) = |x|$ ve $\psi(\nabla u) = |\nabla u|$ denklemler L1 veri ceza terimini ve toplam değişimi düzenlemek için seçilir [79].

$$E = \int_{\Omega} \left\{ \lambda |I_{onset}(x) - I_{apex}(x + u(x))| + |\nabla u| \right\} dx \quad (4. 8)$$



Şekil 4.5. Başlangıç ve tepe kareleri ile optik akış görüntülerinin elde edilmesi.

4.2.4. Veri Arttırma

Olumsuz duyguların olumlu ve sürpriz duygularla karşılaştırıldığında üç veri setinde de daha çok örneklemeye sahip olduğu Çizelge 4.1’de görülmektedir. Veri setlerindeki sınıf dengesizliği, DVM sınıflandırıcısının örnekleme fazla olan sınıfın lehinde sınıflandırma yapmaktadır [84]. Bu çalışmada da kullanılan Mİ veri setlerindeki örnekleme dengesizliğinden dolayı veri arttırma tekniği kullanılmıştır. Böylece dengeli örneklemeye sahip bir veri seti oluşturarak, kullanılacak DVM’nin daha doğru sınıflandırma yaptığı sonuçlar hedeflenmektedir. Bu tez kapsamında yapılan birinci deneysel çalışmada kullanılan bileşik veri setinde bulunan sınıflarda negatif örnekleme 266, pozitif örnekleme 112 ve sürpriz örnekleme 88 adettir. Yapılan ikinci deneysel çalışmada kullanılan bileşik veri setinde bulunan sınıflarda negatif örnekleme 266, pozitif örnekleme 109, sürpriz örnekleme 83 adettir. En fazla veri örneklemesine sahip negatif sınıf referans alınarak, diğer sınıflardaki veriler arttırılmıştır. Veri arttırımı sonunda ilk model için her sınıf eşitlenerek 266 adet Farneback optik akış görüntü verisine, ikinci modelde ise her sınıf 252’şer adet Farneback ve TV-L1 optik akış görüntüsüne sahip olmuştur. Deneysel çalışma üçte ise sadece CASME II veri seti kullanılıp, burada bulunan 94 negatif, 32 pozitif ve 28 sürpriz örnekleme sayısı, veri

arttırma tekniği sonrası her biri 94 örneklemeye sahip üç sınıf haline getirilmiştir. Önerilen üçüncü model çalışmasında tek veri setinin kullanılması, diğer iki deneysel çalışma ile karşılaştırılması hedeflenmektedir. Düşük sayıda örneklemeye sahip sınıflardaki veriler, 90^0 , 180^0 , 270^0 derece döndürülerek yapay olarak arttırılmıştır.

4.3. ÖZELLİK ÇIKARMA YÖNTEMLERİ

4.3.1. Gabor Filtresi Metodu

Gabor filtresi, görüntü işlemede ve bilgisayar görüşünde kullanılan kenar algılama, doku analizi, özellik vektörü çıkarma gibi görevleri yapabilen doğrusal bir hesaplama yöntemidir. Gabor filtrelerinin çalışma prensibi, memelilerin görsel sisteminde bulunan korteks hücrelerinin fizyolojik çalışmasına dayanmaktadır [18,80].

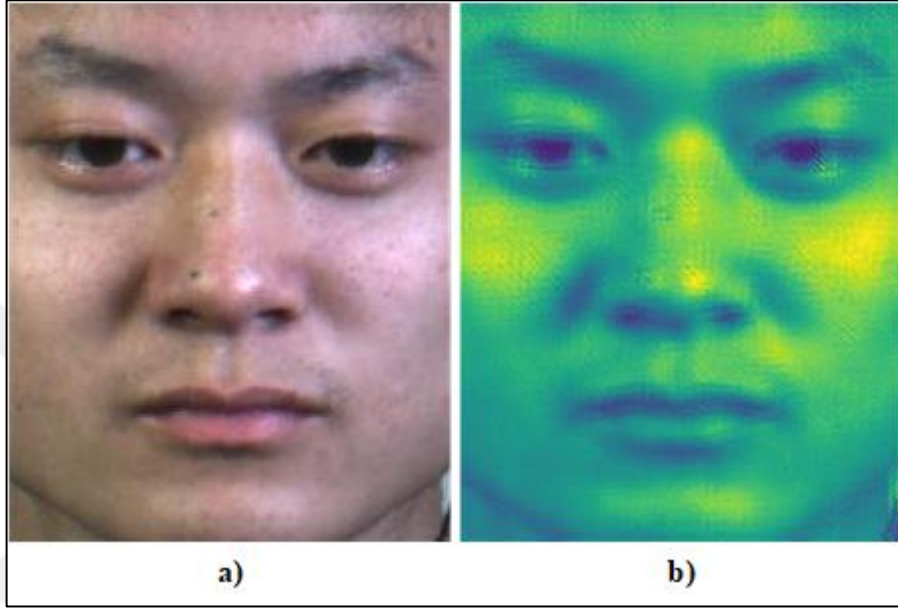
Gabor filtreleri bir görüntüdeki kenar ve doku yapısındaki değişim noktalarını algılamakta, ilgili filtreler bu uzamsal konumlarda ayırt edici özelliğe sahiptir. Gabor filtreleri nesne, yüz ve yüz ifadesi tanıma için kullanılmaktadır [81,82]. Ayrıca yüze ait özellik haritaları çıkarılırken yüzün tamamında ya da bir kısmında [88] kullanılabilir. Yüz hareketlerini sınıflandırmada iyi performans veren gabor, aydınlatma ve döndürmelere karşı dayanıklı farklı frekans ve yönlere sahip filtre bankaları kullanan piksel tabanlı popüler bir yöntemdir [18,81]. Belirli bir bant frekansın geçişine izin veren bir gabor filtresi bir sinüs dalgası ve bir Gauss zarf fonksiyonunun birleşiminden oluşmaktadır. Gabor filtresine ait denklemler Eşitlik 4.9 - 4.11'de verilmiştir [89].

$$Gabor(x, y) = \exp\left(-\frac{x'^2}{2\sigma_x^2}\right) \exp\left(-\frac{y'^2}{2\sigma_y^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \delta\right) \quad (4.9)$$

$$x' = (x - m_x) \cos \gamma - (y - m_y) \sin \gamma, \quad (4.10)$$

$$y' = (x - m_x) \sin \gamma + (y - m_y) \cos \gamma, \quad (4.11)$$

Burada δ parametresi bir faz kaymasını, γ parametresi yönü, σ_x^2 ve σ_y^2 parametreleri alıcı alanı, m_x ve m_y görüntü koordinatlarında gabor fonksiyonunun merkezini belirtir. Şekil 4.6'in sol tarafında CASME II veri setinde bulunan katılımcı birden elde edilen mutluluk duygusuna ait RGB tepe karesi görüntüsü ve sağ tarafta aynı görüntünün Gabor filtreleme işleminden geçirilmiş biçimi gösterilmektedir.



Şekil 4.6. a) RGB tepe kare görüntüsü, b) Gabor Filtreleme işleminden geçmiş tepe kare görüntüsü

4.3.2. Evrişimsel Sinir Ağları

Evrişimsel Sinir Ağı (ESA) hayvanların görsel sinir sisteminden esinlenerek modellenmeye çalışılan bir derin öğrenme mimarisidir [90]. İlk olarak 1962 yılında Hubel ve Wiesel hayvanların görsel korteksi üzerindeki alıcı alanda ışığı algılaması sonucu keşfedilmiştir [91]. 1980'de Fukushima [92] Evrişimli sinir ağının en eski öncüsü kabul edilebilecek, nöronların öğrenme yoluyla görsel kalıpları tanınması üzerine modellenen çok katmanlı neocognitron'u önerdi [93]. 1990'da LeCun ve arkadaşları [94] tarafından çok katmanlı, geri yayılım algoritması ile eğitilebilir, LeNet-5 isimli modern bir yapay sinir ağı tanıtıldı. Mnist el yazısıyla yazılan

rakamları tanıma uygulamasında kullanılan LenNet-5 modelinde evrişimsel katmanlar, havuzlama (alt örnekleme) tam bağlı katmanların kullanıldığı bir mimariye sahiptir. Daha sonraki yıllarda evrişimsel sinir ağlarındaki gelişmeler hızla ilerleyerek 2012’de Alex Krizhevsky ve diğerleri tarafından AlexNet [95] ve peşi sıra VGGNet [96] ve GoogleNet [97], ResNet [98] gibi son teknoloji daha karmaşık hesaplamalar yapabilen, daha derin evrişim katmanları kullanan modeller geliştirilmiştir.

Çok katmanlı ileri beslemeli bir yapay sinir ağı (YSA) olan ESA’ların yapısında bulunan katmanlardaki nöronlar girdinin yükseklik, genişlik ve derinlik boyutuna göre göre düzenlenmiştir. Ayrıca katmanlarda bulunan nöronlar kendisinden önceki katmanın sadece küçük bir bölgesine bağlanmaktadır [94,95]. Evrişimsel sinir ağları nesne tanıma, nesne algılama, görüntü segmentasyonu, görüntü sınıflandırma gibi görevlerde başarılı performanslar göstermektedir.

Sonraki bölümlerde ESA ağlarının mimarisinde kullanılan giriş katmanı, evrişim katmanı, havuzlama katmanı, tam bağlı katman ve çıkış katmanı sırasıyla tanıtılmıştır.

4.3.2.1. Evrişimsel Sinir Ağları Katmanları

Giriş Katmanı (Input Layer)

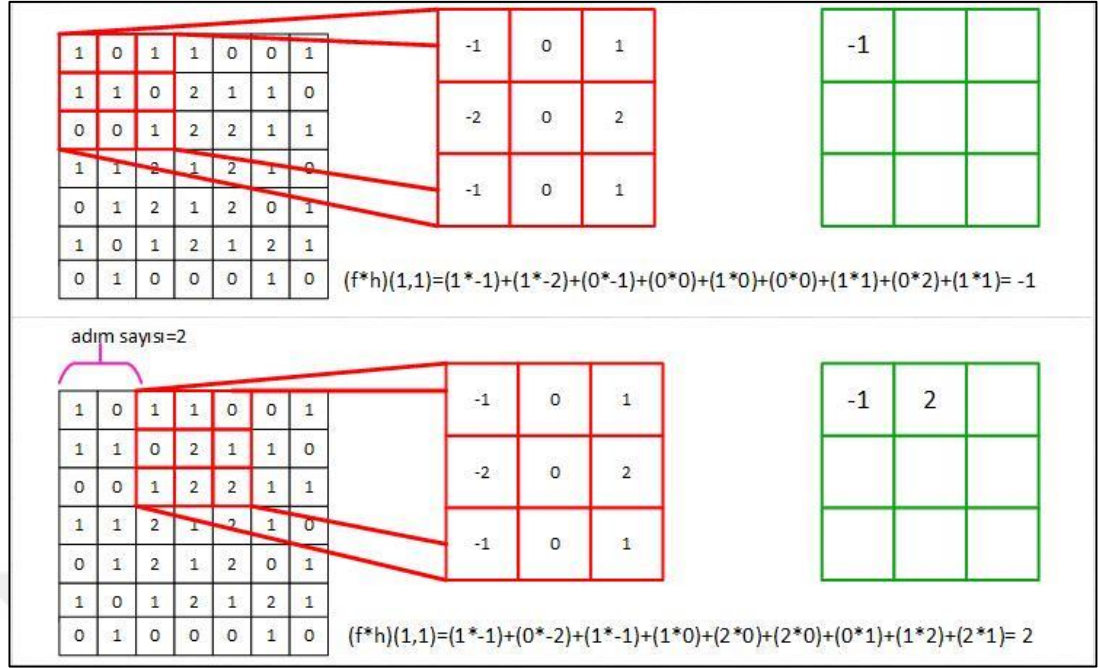
ESA’da işlemler matrisler üzerinden gerçekleştirildiğinden, giriş katmanları görüntüye ait piksellerin matris değerini tutmaktadır [99]. Bu piksel değerleri, görüntünün en ve boy uzunluğundaki boyutlarda matrislerle ifade edilmektedir. Burda görüntünün renk kanallarıda göz önüne alındığında renkli (RGB) bir görüntünün matris formatına çevrilmesiyle $[En \times Boy \times 3]$ şeklide bir matris oluşmaktadır. Veri setindeki görüntülerin boyutları model ağın hızını, eğitim süresini, donanım ihtiyacını belirleyecektir [100].

Evirişim Katmanı (Convolutinal Layer)

Çekirdek (kernel) katmanı olarak da adlandırılan bu katmanda filtre adı verilen (nxn) iki boyutlu matrisler kullanılarak görüntünün kenar, köşe gibi özellik haritaları öğrenilmektedir [101]. Tercih edilecek filtre sayısı ve boyutu, kullanılacak veri setine ve model ağı göre değişim gösterecek olup, aynı zamanda ağı eğitim sürecini ve başarı performansını doğrudan etkileyecektir [100]. Filtreler görüntü üzerinde soldan sağa, yukarıdan aşağıya resim boyunca kaydırılarak görüntünün özellik haritaları oluşturulmaktadır. Adım kaydırma (stride) parametre değişkeni değerine göre filtre, resim üzerinde adımlayarak sola ve aşağıya doğru tüm pikseller işlem görene kadar hareket etmektedir. Görüntüye ait piksel değerlerinin üzerine denk gelen filtre değeri çarpılarak toplanmakta ve görüntünün boyutu pikseller arasındaki ilişki korunarak düşürülmektedir. Ayrıca Filtrenin derinliği de uygulandığı katmanın derinliği ile aynı olmak zorundadır. İlk evirişim katmanında düşük seviyeli özellikler öğrenilirken, sonraki evirişim katmanlarında daha yüksek seviyeli, daha karmaşık özellikler öğrenilmektedir. Bu katmanın çıktı matrisinin derinliği filtre sayısına eşit olacaktır. Bir ağ üzerinde farklı boyut ve sayıda filtreler kullanmak etkin ve zayıf düzeydeki özelliklerin öğrenilmesinde fayda sağlayacaktır [100]. Bu katmanda yapılan evirişim işleminin matematiksel denklemi Eşitlik 4.12’de gösterilmektedir [102].

$$(t * g)[m, n] = \sum_x \sum_y g[x, y] \cdot t(m - x, n - y) \quad (4.12)$$

Denklemde giriş görüntüsü t ve çekirdek g, çıktı matrisinin satır ve sütun indisleri sırasıyla m ve n, evirişim işleminden sonra oluşacak matrisin satır ve sütunu sırasıyla x ve y olarak belirtilmiştir. Şekil 4.7’de giriş görüntüsü üzerinde evrişim işleminin hesaplama süreci gösterilmektedir.



Şekil 4.7. Giriş görüntüsü üzerinde evrişim işleminin hesaplanması.

Evrişim işleminde çıkış matrisinin boyutunu korumak ya da yeniden boyutlandırmak için giriş matrisine fazladan sıfır pikseller eklenebilmektedir. Amacı sınır bilgilerini korumak ve giriş çözünürlüğünü sabit tutmaktır. Bu piksel ekleme işlemine dolgu (padding) denmektedir. Çıkış matrisinin boyutu giriş matrisinin boyutu ile aynı olması isteniyorsa, evrişim işlemi gerçekleştirildikten sonra dolgu işlemi gerçekleştirilir. Çıkış matrisinin boyutları dolgu işlemi gerçekleştirildikten sonra Eşitlik 4.13' de verilmiştir [103].

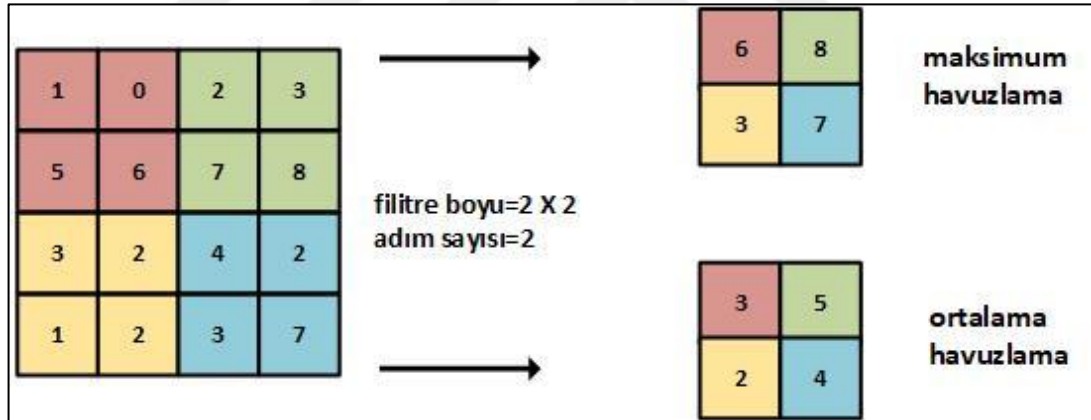
$$z = \left(\frac{i - f + 2p}{s} \right) + 1 * \left(\frac{i - f + 2p}{s} \right) + 1 \quad (4.13)$$

Burada i, giriş matrisinin boyutu f, filtre matrisinin boyutu s, adım sayısı p dolgu boyutunu, z ise çıkış matrisinin boyutunu göstermektedir.

Havuzlama Katmanı (Pooling Layer)

Havuzlama katmanında görüntüye ait en önemli özelliklerin korunarak görüntünün bir alt örnekleme elde edilmektedir [93]. Havuzlama katmanında öğrenme işlemi gerçekleşmemektedir. Havuzlama katmanının amacı verinin boyutunu düşürerek, parametre sayısını azaltmakta böylece hesaplama maliyetini ve hesaplama karmaşıklığını azaltmaktır. Bunun bir dezavantajı bazı bilgilerin kaybolması sonucu performansın azalmasına neden olmaktadır [101]. Burada derinlik (kanal sayısı) sabit tutulurken, indirgeme işlemi en ve boy üzerinde gerçekleştirilir.

Derin öğrenme mimarilerinde en fazla kullanılan maksimum havuzlamada filtreler en büyük piksel değerlerini seçerken, ortalama havuzlama tekniğinde filtre içerisindeki ortalama değeri seçilmektedir. Şekil 4.8’de bir görüntü matrisinin maksimum ve ortalama havuzlama işlemini gösterilmektedir.

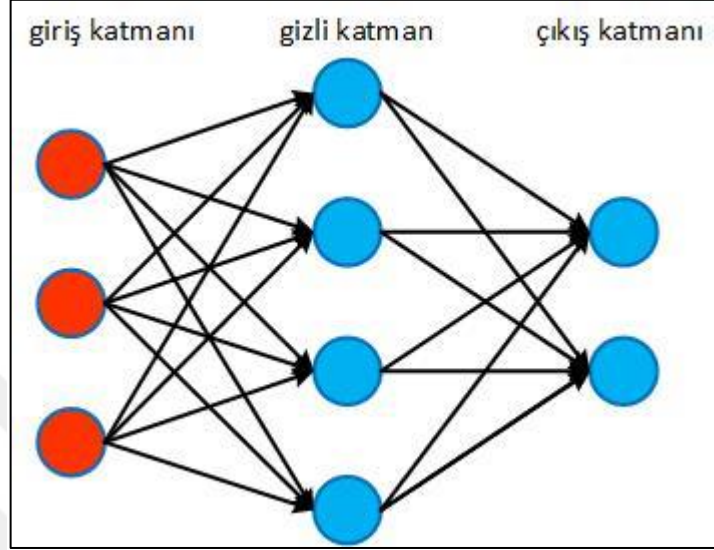


Şekil 4.8. Maksimum ve ortalama havuzlama işleminin hesaplanması.

Tam Bağlı Katmanı (Fully Connected Layer)

Bu katmanda önceki katmanlardan gelen özellik haritası düzleştirilerek tek boyutlu vektör haline getirilir. Bu tek boyutlu vektör, her çıkış nöronunun bir sonraki katmanın tüm giriş nöronlarına bağlı olduğu, yapay sinir ağlarından oluşan, tam bağlı katmanın girdisi olarak kullanılmaktadır. Bu katmanın amacı, önceki katmanlardan elde edilen

önemli özelliklerin skorlarına göre sınıflandırmayı öğrenmektir. Bu katmanda global örüntü kalıpları öğrenilmektedir. Şekil 4.9’da tam bağlı katmanın yapısını ve bağlantılarını temsil eden bir görüntü sunulmuştur.



Şekil 4.9. Tam bağlı katmanın yapısı ve bağlantıları.

4.3.2.2. Aktivasyon Fonksiyonları

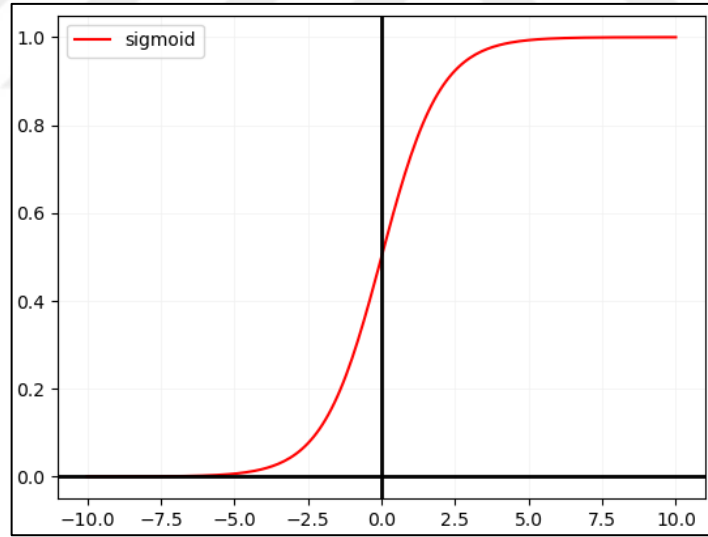
Doğrusal olmayan ses, görüntü, video gibi daha karmaşık gerçek dünya verilerinin yapay sinir ağları tarafından öğrenilebilmesi, zengin bir hipotez uzayı sunabilmesi için doğrusal olmayan bir aktivasyon fonksiyonuna ihtiyaç duyulmaktadır [104]. Aktivasyon fonksiyonu, giriş sinyalleri üzerinde yapılan doğrusal olmayan dönüşümler uygulanarak elde edilen anlamlı bir çıkış sinyalinin bir sonraki nörona girdi olarak aktarılmasıdır. Aktivasyon fonksiyonu kullanılmayan bir yapay sinir ağının çıktısının tek dereceli doğrusal bir fonksiyondan farkı olmayacaktır. Aktivasyon fonksiyonundan iyi bir performans elde edebilmek için seçilen fonksiyonun probleme, veriye ve sisteme uygun olması gerekmektedir. Aktivasyon fonksiyonu seçimi ağın performansı üzerinde önemli bir etkisi olmakla birlikte farklı aktivasyon fonksiyonları aynı ağın farklı bölümlerinde kullanılabilir [105]. Aktivasyon fonksiyonu doğrusal olmayan, sürekli yani türevlenebilir bir fonksiyon

olması demek ağırlıkların geri yayılımı yoluyla güncellenerek öğrenmenin etkin olarak gerçekleştiği anlamına gelmektedir [106].

Sigmoid

Doğrusal olmayan aktivasyon fonksiyonlarından biri olan sigmoid fonksiyonu, girişten bağımsız olarak $(-\infty, +\infty)$ her zaman 0.0 ile 1.0 aralığında çıkış değeri vermektedir. Girdi değeri ne kadar büyükse 1.0'a yaklaşırsa, küçüldükçe 0.0'a doğru yaklaşmaktadır. Şekil 4.10'daki grafikte görüldüğü üzere X eksenini sağ tarafında 1'e yaklaşırsa, sol tarafında 0'a doğru yaklaşmaktadır. Bu yüzden ikili sınıflandırma problemi için ideal bir aktivasyon fonksiyonudur. Sigmoid fonksiyonunun matematiksel denklemi Eşitlik 4.14'de verilmiştir [104].

$$Sigmoid(x) = \frac{1}{1 + e^{-x}} \quad (4.14)$$



Şekil 4.10. Sigmoid fonksiyonunun dalga görünümü.

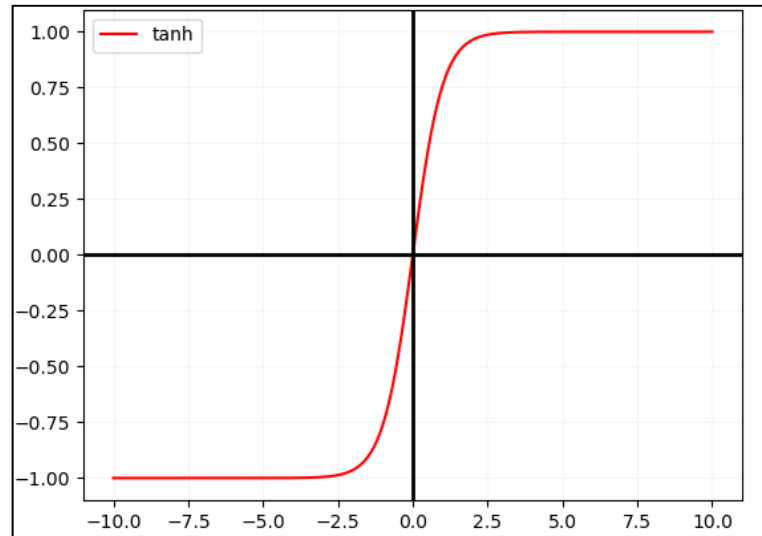
Sigmoid fonksiyonunun dezavantajlarından ilki kaybolan gradyan sorunu, ikincisi sigmoid çıktılar sıfır merkezli olmamasıdır [99,102]. Sigmoid aktivasyon fonksiyonu

sınıflandırmada kullanıldığından çıktı katmanında kullanılmakta önceki gizli katmanlarda kullanılmamaktadır [108].

Hiperbolik Tanjant

Hiperbolik tanjant (Tanh) fonksiyonu -1 ve 1 aralığında çıktı değerleri veren doğrusal olmayan aktivasyon fonksiyonudur. Sigmoid fonksiyonunda bahsedilen sıfır merkezli olmama sorununu ortadan kaldırdığından sigmoid fonksiyona tercih edilmektedir [99,102]. Orjine göre simetrik olmasından dolayı sigmoid fonksiyonunun aksine, bir sonraki katmana girdi olarak farklı çıkış işaretleri ile beslenecektir. Tanh, türevlendiğinde daha fazla değerler alabiliyor olmasından dolayı hızlı öğrenme ve geniş aralıklı sınıflandırma yeteneği bulunan verimli bir aktivasyon fonksiyonudur., Şekil 4.11’da bu fonksiyona ait grafiksel gösterimi sunulurken, Eşitlik 4.15 de tanh fonksiyonuna ait matematiksel denklemler verilmiştir [104].

$$Tanh(x) = 2Sigmoid(2x) - 1 \quad veya \quad Tanh(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}} \quad (4.15)$$



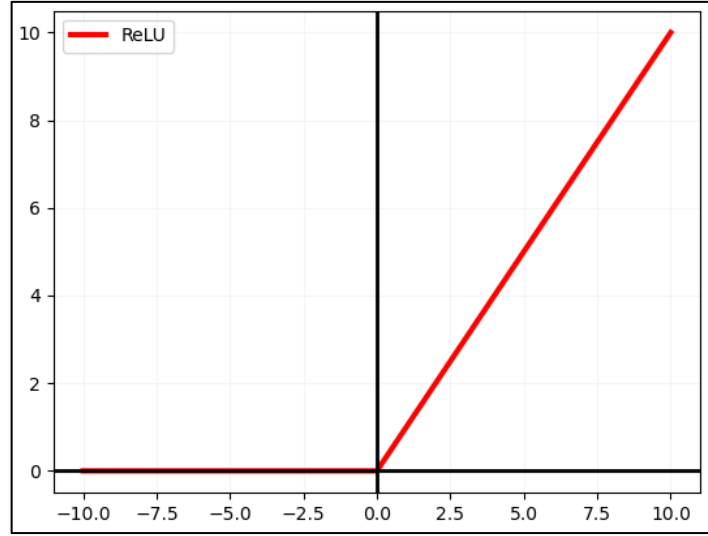
Şekil 4.11. Hiperbolik tanjant fonksiyonunun dalga görünümü.

ReLU (Rectified Lineer Unit)

ReLU aktivasyon fonksiyonu, pozitif girişler için çıkış girdi değerlerine eşit iken, negatif girişler için ise her zaman sıfır çıkış değerini veren doğrusal olmayan bir fonksiyondur. Bu yönüyle yarı doğrusal fonksiyon gibi görülsede negatif değerler için doğrusal olmadığından parçalı doğrusal fonksiyon olarak da tanımlanmaktadır. ReLU diğer fonksiyonlardan daha etkilidir çünkü tüm nöronlar aynı anda aktive edilmediğinden, bir seferde belirli sayıda nöron aktive edilir. Buda ağı hızlı çalışmasını yanı sıra daha az ek yük getirecektir [108]. Eğitim işlemi sırasında yakınsamayı hızlandıran ReLU fonksiyonu geleneksel sigmoid ve tanh fonksiyonundan daha iyi çözüm sunmanın yanı sıra sistem performansında daha başarılı olduğu görülmüştür [104,105]. Ayrıca ReLU aktivasyon fonksiyonuna sahip derin evrişimli sinir ağlarının, ön eğitim almadan bile verimli bir şekilde eğitilebilmektedir [90]. Daha hızlı öğrenme, büyük veri kümeleri üzerinde eğitilmiş büyük modellerin performansı üzerinde büyük bir etkiye sahip olan ReLU fonksiyonunun matematiksel eşitliği Denklem 4.16'da verilmiştir [95].

$$Relu(x) = \max(0, x) \quad (4. 16)$$

Relu fonksiyonunun dezavantajları olarak Tanh ve sigmoid'e göre maliyetli olması ve pozitif değerleri değişiklik yapmadan olduğu gibi yansıtmasıdır [103,106]. Şekil 4.12'de ReLU fonksiyonunun grafiksel gösterimi sunulmuştur.



Şekil 4.12. ReLu fonksiyonunun dalga görünümü.

ReLu fonksiyonun bazı nöronların sıfırlanıp etkinleştirilmemesinin avantaj ve dezavantajları bulunmaktadır. Optimizasyon sırasında nöronlar aktif olmadığından dolayı gradyan sıfır olacak bu da kaybolan gradyanlara sebep olacaktır. Bu gradyanlar aktif olmadığı için gradyan tabanlı bir optimizasyon algoritması başlangıçta hiçbir zaman etkinleştirilmeyen bir birimin ağırlıklarını güncelleyemeyerek, öğrenme kaybolan gradyanlar nedeniyle yavaşlabilecektir. Ayrıca Relu fonksiyonun kısmi doğrusallığı nedeniyle hesaplama maliyeti daha verimli olacaktır [99,107].

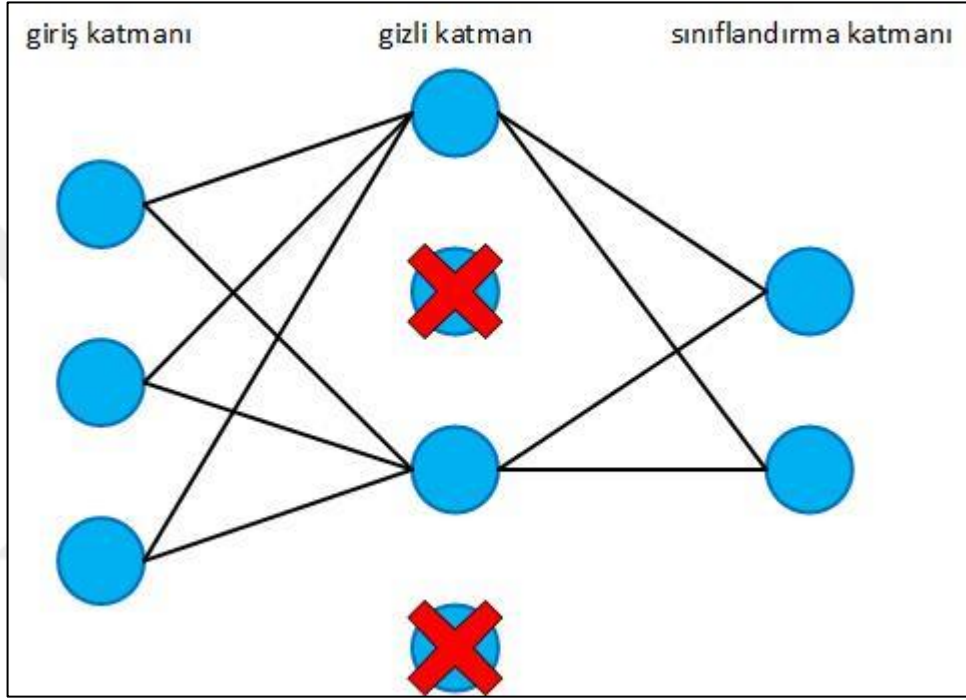
Softmax

[0,1] aralığında, çıktıların toplamı 1 olacak şekilde olasılıksal dağılımlara sahip çıktı değeri veren softmax fonksiyonu, sigmoid fonksiyonundan farklı olarak herhangi bir sayıda çok boyutlu sınıflandırma problemleri için kullanılmaktadır [104]. Olasılıksal dağılım değeri en yüksek olan nöron, tahmin edilen sınıfı temsil edecektir.

Seyreltme Katmanı (Dropout)

Derin öğrenme sırasında ağırlık ezber yapmasını, overfitting (aşırı uyum) olmasını önlemek için bazı verilerin belirli oranlarda unutturulması gerekmektedir. Belirli

oranda rastgele seçilen bazı nöronların eğitim aşaması sırasında kaldırılıp kullanılmaması durumudur. Seyreltme işlemi, bir çok farklı ağı eğitip modellerin ortalaması alınmasına benzemekte olup, aynı zamanda test setindeki hatayı azaltmanın verimli bir yoludur [113]. Dezavantajı olarak, her eğitim sırasında farklı nöronlar aktif olacağı için nihai gradyanları hesaplamak eğitim süresinin uzamasına neden olacaktır [114]. Şekil 4.13’de seyreltme katmanını temsil eden bir görüntü sunulmuştur.



Şekil 4.13. Seyreltme katmanının yapısı ve bağlantıları.

4.3.2.3. ESA Parametreleri

Mini-Batch Değeri

Modelin eğitiminde, veri setindeki tüm örnekleri aynı anda kullanmak zaman ve bellek gerektiren maliyetli bir durum olacağı için, veri setinin küçük gruplara ayrılarak paçalar halinde modelin eğitilmesi durumuna mini-batch denilmektedir. Mini batch parametre güncellemesinin gerçekleştiği her adımda ağı eğitmek için kullanılacak

örnek sayısı olarak adlandırılmaktadır [115]. Mini batch için yaygın olarak 2^n kullanılmaktadır [102].

Epoch

Modelin eğitimi sırasında, eğitim setindeki tüm örneklerin ağa bir kere verilmesine bir epoch (devir) olarak adlandırılmaktadır. Epoch sayısı ise ağın eğitim esnasında tüm veri setini işleme sayısını veren bir hiperparametredir [115]. Her epoch da ağırlıklar güncellenerek model için en uygun ağırlık değerleri bulunmaya çalışılmaktadır. Belli bir noktaya kadar epoch sayısı arttıkça modelin başarı performansı da artacaktır, daha sonra epoch sayısı artsa bile performanstaki artışlar çok küçük olacağı için eğitim bu noktadan sonra sonlandırılabilir [102].

Öğrenme Oranı

Ağın global minimum değere ulaşmak için parametrelerinin güncelleme hızı olarak tanımlanmaktadır [102]. Öğrenme oranındaki çok küçük değerler takılıp kalabilecek uzun bir eğitim süreci ile sonuçlanabilirken, çok büyük bir değerler ağın çok hızlı öğrenmesine ve minimum değerın atlanıp kararsız bir eğitim sürecine neden olabilmektedir. Bundan dolayı öğrenme oranını belirlemek eğitim hızını ve ağın doğruluğunu etkileyecektir [116]. Bazı zaman sabit olarak belirlenen bu parametre, bazen de ağ içerisinde güncellenerek minimuma yaklaştırmaya çalışılmaktadır.

4.3.2.4. ESA Optimizasyon Algoritmaları

Optimizasyon, hedef sistemin performansını olabildiğince verimli hale getirmektir. Bu anlamda optimizasyon algoritmalarında amaç model ağı global minimum değere ulaştırmak yani hata oranını en aza indirmektir [117]. Makine öğrenmesinde kullanılan optimizasyon algoritmaları birbirlerinin eksik yönlerini gidermek için geliştirilmiştir. En yaygın ve popüler olan algoritmalar Stochastic Gradient Descent (SGD), Adagrad, Adadelat, RMSProp ve Adam optimizasyon algoritmalarıdır.

Aşağıda bu algoritmaların problemlere karşı sundukları çözümler ve çalışma teknikleri anlatılmaktadır.

Gradyan Azalma (Gradient Descent)

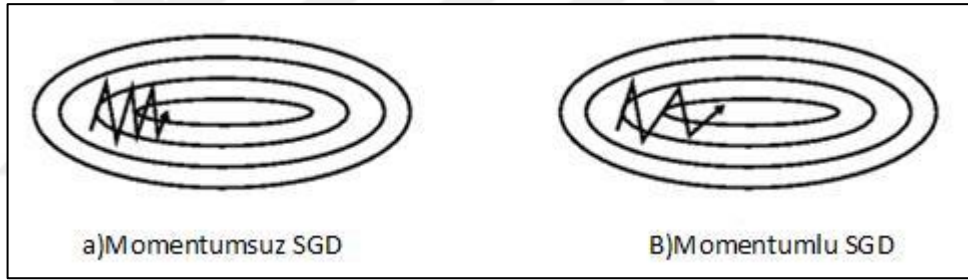
Derin öğrenmede ağıın eğitimi, çıkış katmanından önceki katmanlara doğru geri yayılım yapılması ile ağırlıklar güncellenerek hata oranının azaltılması yoluyla yapılmaktadır. Geri yayılım esnasında bazı katmanlar güçlendirilebilir buna gradyan patlaması, bazı katmanlar zayıflatılıp kaybolabilir buna gradyan kaybolması denmektedir [117].

Gradient azalma sinir ağılarını eğitiminde yerel minimuma erişmek için birinci dereceden yinelemeli bir optimizasyon algoritmasıdır. Kayıp fonksiyonun gradyan hesaplamasında, veri miktarına göre üç farklı gradyan azalma optimizasyon algoritması bulunmaktadır. Bunlar toplu gradyan inişi, SGD ve mini batch gradyan inişi optimizasyon algoritmalarıdır. Toplu gradyan yada vanilya gradyan olarak adlandırılan algoritmada, eğitim veri kümesindeki tüm örneklerin hata miktarı hesaplandıktan sonra yinelenen bir algoritmadır [118]. Bundan dolayı gradyan azalma süreci uzun sürebilmekte ve büyük veri seti üzerinde ölçekleme yapılamayıp, tüm veri seti bellekte bulunması sorunundan dolayı büyük veri setleri için yavaş ve uygun değildir [112,114]. Büyük veri setleri için daha kullanışlı olan SGD her yineleme için tüm veri seti yerine rastgele örnek kümeleri seçilir [118]. Stokastik gradyan algortması, toplu gradyan algoritmasına kıyasla daha hızlı olsa da kayıp fonksiyonu güncellenmesinde fazla dalgalanmalara da neden olabilir [120]. Mini batch gradyan azalma algoritması stokastik ve toplu grandyan inişi algortmalarının bir karışımı olarak, eğitim setini küçük kümelere bölerek her küme için güncelleme gerçekleştirmektedir [113,115]. Basit formdaki gradyan azalmasının matematiksel eşitliği Denklem 4.17’de verilmiştir [121].

$$Q_{i+1} = Q_i - \eta_i \nabla F(Q_t) \quad (4.17)$$

Eşitlikte η i öğrenme oranı ya da adım boyutu, $\nabla F(\theta_i)$ i inci iterasyon için kayıp fonksiyonunun gradyanıdır.

Momentum : Büyük ağ modellerinin eğitiminde sadece basit bir gradyan azalması kullanılması, çok fazla çalışma zaman gerektiren bir sürece dönüşebilmektedir [118]. Öğrenme sürecini hızlandırmak, karşılaşılan salınımları sönümlenmek için momentum katsayısı kullanılmaktadır [115,117]. Bu yöntemin amacı aynı yönü işaret eden gradyanların boyutları için momentumu artırıp öğrenmeyi hızlandırmak, gradyanlar yön değiştirdiğinde momentumu azaltıp öğrenme hızı yavaşlatmaktır [114,118]. Momentum katsayı terimi genellikle 0,9 değerine ayarlanmaktadır. Şekil 4.14'de momentumlu kullanılan SGD ve momentumsuz SGD gösterilmektedir. Momentum kullanılmadığında daha fazla salınım ya da dalgalanmaların olduğu görülmektedir [119].



Şekil 4.14. a) Momentumlu SGD algoritması, b) Momentumsuz SGD algoritması.

Adagrad (Adaptive Gradient)

Her parametre için kendine ait öğrenme oranını (adım boyutu) otomatik olarak hesaplayarak güncelleyen uyarlanabilir bir optimizasyon algoritmasıdır. Küçük gradyanlar artan öğrenme oranıyla, büyük gradyanlar da azalan öğrenme oranıyla dinamik olarak güncellenmektedir [112,118]. Bu uyarlanabilir öğrenme oranına sahip algoritmanın avantajı seyrek verilerin öğrenmeye katılımını arttıracaktır [119]. Bu algortimanın dezavantajı çok büyük gradyanlar için kullanılan devamlı azalan öğrenme oranı belli bir süre sonra sıfıra yaklaşacağı için model öğrenmeyi durdurma

eğilimi gösterecek veya ilerlemenin çok fazla yavaşlaması sonucu optimumu bulamaması anlamına gelecektir [123].

Adadelta

Adagrad optimizasyon algoritmasında, fonksiyonun minimumunu yakalamak için monoton olarak azalan öğrenme oranının sıfıra'a yaklaşmasından kaynaklı problemini çözmek için tasarlanmış bir mimariye sahip optimizasyon algoritmasıdır [119]. Bunu geçmiş tüm kare gradyanları biriktirmek yerine sınırlandırma yapılarak, erken gradyanların unutulup son gözlemlenen kısmi gradyanlar kullanarak yapmaktadır[106,112].

Sadece birinci dereceden bilgilere dayanan, boyut başına öğrenme oranı sağlayan, öğrenme katsayısının manuel ayarlanmasına gerek olmayan dinamik, sağlam bir öğrenme oranı yöntemi olan bir optimizasyon algoritmasıdır. SGD ile karşılaştırıldığında çok düşük hesaplama maliyeti bulundurmaktadır [124].

Karekök Ortalama Yayılımı (RMSProp)

RMSprop, Adadelta'ya benzer şekilde Adagrad optimizasyon algoritmasının monoton azalan öğrenme oranı problemine çözüm sunmak için geliştirilmiş bir optimizasyon algoritmasıdır [119]. Geçmişteki gradyanların hesaplanmasında tüm gradyanlar yerine sabit boyutlu bir pencere ile sınırlayarak, daha eski gradyanları güncelleme dışında tutmaktadır [112,118].

Uyarlanabilir Moment Tahmini (Adam)

Günümüzde sinir ağı eğitimi için en popüler optimizasyon algoritmalarından bir diğeri olan Adam, RMSProp ve momentum yöntemlerinin birleşimidir [114,116]. Algoritma verimli hesaplama, düşük bellek gereksinimi gibi avantajlarının yanı sıra büyük veri setine sahip ve yüksek parametrelili problemlerin çözümü için de uygundur. Hiper

parametreleri sezgisel yorumlara sahip olduğu için çok az ayarlama gerektirmektedir. [125].

4.3.3. Kullanılan Esa Modelleri

Derin öğrenme modellerinde zamandan tasarruf etmek veya daha başarılı performans elde etmek için transfer öğrenme yöntemi derin öğrenme çalışmalarında kullanılmaktadır. Büyük veri setlerinde günlerce, haftalarca süren eğitimler sonucu eğitilen ağırlıkların bir başka veri seti içinde kullanılabildiği, öğrenmenin daha hızlı olduğu ve aktarıldığı duruma transfer öğrenme denilmektedir. Transfer öğrenme, küçük veri setleri için aşırı öğrenme durumundan korumakta, daha büyük veri setlerinde eğitildiği için modelin veri seti üzerindeki performansını arttıracaktır. Eğitilmiş bir modelin, ikinci bir görevde önemli ince ayar yapıldıktan sonra özelliklerin aktarılıp, yeni model için başlangıç noktası olarak yeniden kullanıldığında genelleme performansını iyileştireceği bulunmuştur [126].

4.3.3.1. AlexNet

ILSVRC-2010 ve ILSVRC-2012 yarışmasını görüntü sınıflandırma görevlerinde kazanan ilk ESA model olan AlexNet'in eğitimi sırasında $227 \times 227 \times 3$ sabit boyutlu görüntü girdileri kullanılmakta olup ilk beş katmanı evrişim ve son üç katmanı tam bağlı olmak üzere toplam sekiz katmandan oluşan bir mimari ile tasarlanmıştır. Son bağlı katmanın çıktısı 1000 sınıflı bir problemi sınıflandırmak için softmax aktivasyon fonksiyonu kullanılmaktadır. Birinci, ikinci ve beşinci evrişim katmanlarından sonra maksimum havuzlama kullanılmaktadır. ReLU doğrusal olmama durumu, her evrişimli ve tam bağlantılı katmanın çıktısına uygulanır. Tam bağlı katmanların her biri 4096 nörona sahiptir. Bırakma tekniği ile tamamen bağlı ilk iki katmanda kullanılarak, gizli nöronun çıktısını 0.5 olasılıkla sıfırlama ile her seferinde sinir ağının farklı bir mimariyi örneklemesi sağlanır. Buda aşırı uyumu sorununu önlemektedir. Ayrıca artık yaygın olmayan normalleştirme katmanları kullanılmıştır [95].

4.3.3.2. VGGNet

ILSVRC-2014 yarışmasında görüntü sınıflandırma görevini kazanan VGGNet derin öğrenme ağında 11, 13, 16 ve 19 katmanlı altı farklı modele (A, A-LRN, B, C, D, E) sahiptir. VGGNet'in modellerinin eğitiminde $224 \times 224 \times 3$ sabit boyutlu görüntü girdileri ve her evrişim katmanlarında aynı 3×3 filitreler kullanılmaktadır. Her ağ modelinde ortak, 5 adet maksimum havuzlama ve 4096, 4096, 1000 nörondan oluşan üç tam bağlı katman ve 0.5 bırakma yer almaktadır. Son bağlı katmanın çıktısı 1000 sınıflı bir problemi sınıflandırmak için softmax aktivasyon fonksiyonu tercih edilmiştir. Tüm gizli katmanlarda doğrusal olmayan ReLU fonksiyonu kullanılmıştır [96].

4.3.3.3. SqueezeNet

Bu ESA mimarisinde hedeflenen daha az parametreyle eşdeğer doğruluk sağlamaktır. Öyle ki bu mimari AlexNetde bulunan parametre sayısının 1/50 sini kullanarak ImageNet veri setinde aynı doğruluğu elde etmiştir. Ayrıca SqueezeNet modelinin boyutu sıkıştırma teknikleriyle 0.5 MB altına düşürülebilir ki bu AlexNet ile karşılaştırıldığında 510 kat daha küçük olduğu anlamına gelmektedir [127].

SqueezeNet fire (yangın) adı verilen modüller ve üç önemli strateji üzerine tasarlanmıştır. Fire modülü içerisinde boyut küçültme, sıkıştırma işlemi yapan Squeeze adı verilen ve genişletme işlemi yapan Expand adı verilen katmanlar bulunmaktadır. Squeeze katmanında yalnızca kanallar arasında özellik yakalayan noktasal 1×1 evrişim filitreleri bulundurulurken, çıktısı 1×1 ve 3×3 evrişim filitreleri bulunan Expand katmanını beslemektedir. Diğer yandan ilk stratejik önermeleri çoğu 3×3 filitreler yerine 1×1 filitreler kullanılarak parametre sayısını 9 kat düşürülmüştür. İkinci strateji ise 3×3 evrişimlerden oluşan bir katman için toplam parametre miktarı (giriş kanalı sayısı) $\times$ (filtre sayısı) $\times$ (3×3) şeklindedir. Bu katmanlara giden giriş kanallarının sayısı, sıkıştırma katmanları ile derinliği azaltılıp darboğaz katmanı olarak 1×1 filtrelerden kullanılarak, 3×3 filtrelere düşürmektir. Üçüncüsü sınıflama doğruluğunu artırma üzerine yapılmış bir stratejidir. Daha yüksek sınıflandırma

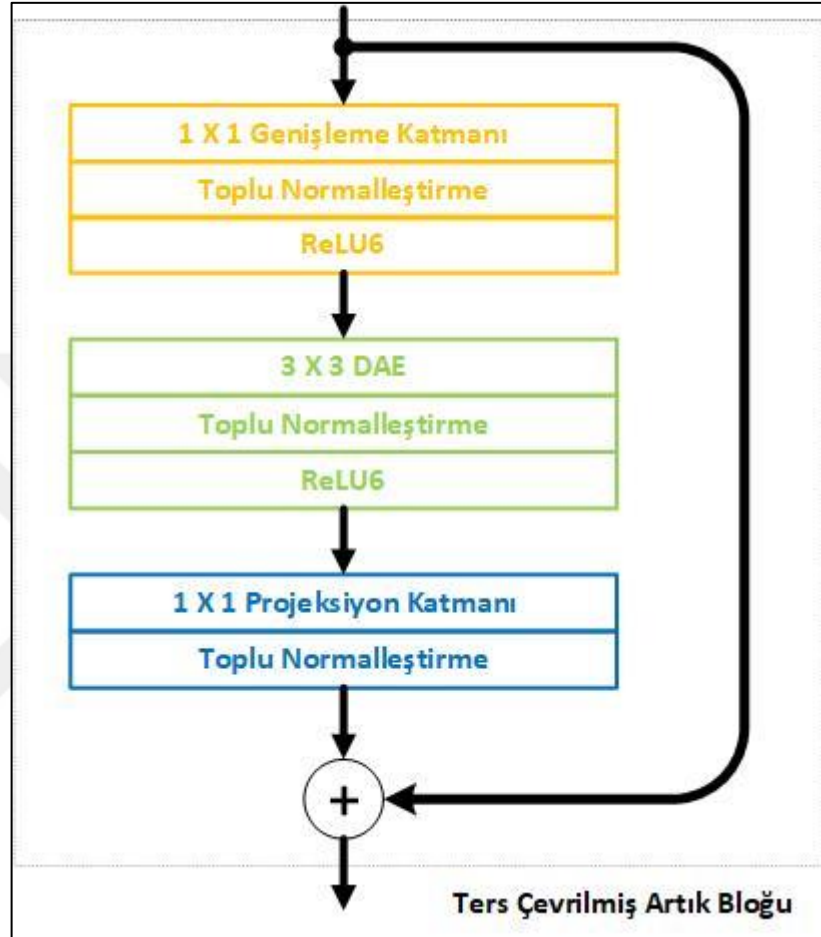
doğruluğu elde edilen gecikmeli alt örnekleme sayesinde [128] büyük özellik haritaları elde edilerek özellik kaybını azaltılmaktadır. Daha geç katmanlarda özellik haritalarının boyutları küçültülmekte buda doğruluğu arttırmaktadır [129]. Bunun içinde filtre adım sayısı düşük tutularak daha büyük aktivasyon haritaları elde edilmektedir.

4.3.3.4. MobilNet V2

MobilNetV1’de ağında hesaplamaları azaltmak için pahalı standart evrişimler yerine daha ucuz olan, Xception [130] modelinde de kullanılan, derinlemesine ayrılabilir evrişimler (DAE) kullanan bir model geliştirilmiştir. MobilNetV2 ise MobilNetV1’den gelen fikirleri temel alıp, ek olarak tersine çevrilmiş artık (residual) bağlantılar (MBConv) ve genişletme/projeksiyon adı verilen katmanların eklenmesi üzerine tasarlanmış bir modeldir. MobilNetV1’den gelen DAE ve giriş kanallarının doğrusal kombinasyonlarını hesaplayarak yeni özellikler oluşturan 1x1 noktasal evrişim katmanı MobilNetV2’de de kullanılmaktadır. MobilNetV2 modelinde kullanılan 1x1 noktasal evrişimleri V1’den farklı olarak sadece genişletme yapmayıp, kanal sayısı hem genişletilip hem de azaltılmaktadır [131].

Şekil 4.15’de gösterilen MobilNetV2 mimarisinde kullanılan ters çevrilmiş artık bloğu ResNet’in aksine kanal sayısı dar, geniş, dar değil daha az parametre kullanılan ters bir mantık geliştirerek geniş, dar, geniş şeklinde tasarlanmıştır. İlk katman, 1x1 evrişim filtresi uygulayarak kanal sayısını arttıran Expansion (genişleme) katmanıdır. Bu katmanda, genişleme faktörü adı verilen ayarlanabilen bir hiperparametre tanımlanmıştır. İkinci katman DAE katmanıdır. Burada iki boyutlu evrişim, her kanal için ayrı tek bir filitreleme kullanılıp daha sonra 3B tensörü için tekrar birleştirildikten sonra 1x1 noktasal bir evrişim kullanılarak doğrusal bir çıktı elde edilmektedir [132]. Üçüncü projeksiyon isimli katman, 1x1 evrişim kullanılarak kanal sayısını azaltıp, giriş verisini çok daha düşük boyuta sahip bir tensöre dönüştürdüğü katmandır. ResNet’de kullanılan artık blokları MobilNertV2 de yalnızca bloğa giren kanal sayısı bloktan çıkan kanal sayısı ile aynı olduğunda kullanılır. Ayrıca bu modelde sıfır ve negatif giriş değerler için çıkışı sıfır veren, sıfır ve altı arasındaki giriş değerler için

mevcut değeri veren, altı ve üzeri giriş değerleri için çıkışında altı değerini veren $\min(\max(x, 0), 6)$ fonksiyonunun karşılığı ReLU6 aktivasyon fonksiyonu tercih edilmektedir [131].



Şekil 4.15. Ters Çevrilmiş Artık Bloğu'nun yapısı ve katmanları.

4.3.3.5. EfficientNetB0

EfficientNet makalesinin en büyük iki katkısı, bileşik ölçekleme ve sinir mimarisi araması NAS (Neural Architecture Search) kullanılarak bir modelin tasarlanması olarak kabul edilmektedir. Geleneksel manuel ölçekleme yöntemiyle ağın derinliği, genişliği veya çözünürlüğünden herhangi birini arttırmak konvolüsyon ağlarında ağın doğruluğu belli bir noktaya kadar arttırsa da belirli bir noktadan sonra doğruluk

doygunluğa ulaşacağı için, bu işlem doğruluk verimliliği arttırmayacaktır. Tüm bu üç parametrenin (derinlik, genişlik, çözünürlük) eşit olarak ölçekleyip dengelemenin daha iyi performansa yol açabileceğini düşünülüp yeni bir bileşik ölçekleme yöntemi, model ölçeklemeyi önermişlerdir. Ağı performansını arttıran bir başka yöntemde EfficientNet modeli, sinir mimarisi araması (Neural Architecture Search-NAS) tarafından geliştirilmiştir. Bu teknik eğitim hızını ve parametre verimliliğini birlikte optimize edip, bir sinir ağı için en iyi mimariyi otomatik olarak tasarlamaktadır. Ayrıca MobilNetV2’de ters çevrilmiş darboğaz evrişimi kullanılmaktadır. Bu bloklar ayrıca daha derin ağlar için ReLU’den daha iyi çalışma eğiliminde olan doğrusal ve sigmoid aktivasyonun çarpımı swish aktivasyon fonksiyonu kullanılmaktadır [133].

$$Swish(x) = x * sigmoid(x) \quad (4.18)$$

4.3.3.6. DenseNet

DenseNet özellik haritalarının boyutunu değiştiren yoğun bağlantıların olduğu, yoğun bloklardan oluşmaktadır. DenseNet, ResNet’de katmanlar arasında bağlantıyı sağlayan identity functiona benzer bir yapı geliştirmiştir. DenseNet, ResNet in aksine bağlantıları bir önceki katmandan değil, kendinden önceki tüm katmanlardan yapılmaktadır. Böylece kendinden önceki tüm özellik haritaları birleştirilir. Modelin yapısında toplu normalleştirme, doğrultulmuş doğrusal birim (RELU) ve 3x3 konvolüsyon işlemlerini birleştiren bileşik fonksiyonu bulunmaktadır. Bloklar arasında geçiş ise 1x1 evrişim ve 2x2 ortalama havuzlama katmanları gerçekleştirilir. DenseNet modelindeki bağlantı eşitliğini gösteren ifade Eşitlik 4.19’da verilmiştir [134].

$$x_k = H_k([x_0, x_1, \dots, x_{l-1}]) \quad (4.19)$$

Eşitlikte bulunan $[x_0, x_1, \dots, x_{l-1}]$ ifadesi katmanlardan gelen özellik haritalarının birleşimini ifade etmektedir. Eşitlikte ki k değeri katmanın indeksini, $H_k()$ fonksiyonu ise toplu normalleştirme, RELU ve konvolüsyon işlemlerini içeren bileşik fonksiyonu ifade etmektedir.

4.3.3.7. Xception

Google tarafından 2017 tarihinde sunulan Extreme Inception [130] anlamına gelen bu ağ, Inception modülü yerine modifiye edilmiş derinlemesine ayrılabilir evrişim katmanlarını temel alan bir evrişim mimari modelidir. Modelin temel fikri uzamsal ve derinliksel olarak özellik haritalarını ayrıştırmak ve ek olarak residual bağlantılarla özellik haritalarını güçlendirmektir. Burada ayrılabilir evrişim ağlarında önce derinlemesine evrişim yapıp daha sonra noktasal evrişim yapılırken, Xception modelinin kullandığı değiştirilmiş derinlemesine ayrılabilir evrişim ağında ise tersi kullanılmaktadır.

4.4. ÖZELLİK SEÇİMİ METODLARI

4.4.1. Parçacık Sürü Optimizasyonu

Parçacık Sürü Optimizasyonu (Particle Swarm Optimization-PSO) analizi, kuş sürülerinin ihtiyacı karşılayacak en iyi hedefe ulaşmak için yaptıkları davranışlarını, matematiksel düzleme yerleştirilmesidir [135]. Dr. Kennedy ve Dr. Eberhart tarafından 1995 yılında geliştirilmiş bir optimizasyon algoritmasıdır [136]. Sürülerin yiyecek ararken hem bireysel hem de sürü halindeki hareketleri birbirlerini etkileyerek amaçlarına daha kısa zamanda ulaştıkları gözlemlenmiştir. PSO analizinde amaç en optimum sürede en iyi çözümü bulmaktır. PSO algoritmaları, hızlı, düşük maliyetli, güçlü arama özelliği, hesaplama basitliği gibi yeteneklerinden dolayı sınıflandırma problemlerinde daha fazla kullanılır hale getirmektedir [137]. Bu algoritma yüz ifadelerinin tanınmasında da DVM ile etkili bir biçimde kullanılmıştır [135]. PSO hesaplamasında kullanılan uygunluk fonksiyonu Denklem 4.20 ve 4.21’de ifade edilmiştir [138].

$$v_i^{t+1} = \beta * v_i^t + c_1 * \text{rnd}_1 * (k_{en_iyi}^i - x_i^t) + c_2 * \text{rnd}_2 * (g_{en_iyi}^i - x_i^t) \quad (4.20)$$

$$x_i^{t+1} = x_i^t + v_i^{t+1} \quad (4.21)$$

Denklem 4.20 ve 4.21'deki ifadelerde, i çekirdek parçacık indeksi, β eylemsizlik katsayısı, x_i^t çekirdek parçacığının t anındaki konumu, v_i^t çekirdek parçacığının t anındaki hızı, rnd_1 ve rnd_2 her hız güncellemesinde yeniden oluşturulan rastgele değerler, kişisel en iyi anlamına gelen $k_{en_iyi}^i$ parçacığının t anında çözüme en yaklaştığı durum, global en iyi anlamına gelen $g_{en_iyi}^i$ sürünün t anında çözüme en çok yaklaşılan durum, $c1$ ve $c2$ ise ivme katsayılarıdır.

PSO yapı itibari ile süreklilik içerirken, özellik seçimi ise ayrık alan uzayında gerçekleşen bir ikili optimizasyon problemidir. Bundsn dolayı Kennedy ve Eberhart [139] ayrık alanlardaki problemlerin çözümü için PSO algoritmasının ikili versiyonunu geliştirmişlerdir. İkili PSO, hız değerini sürekli uzaydan ikili uzaya dönüştürmek, hız değerini $[0,1]$ aralığına eşleştirmek için bir sigmoid transfer fonksiyonunu kullanmaktadır. İkili PSO algoritmasında kullanılan hız değerini dönüştürmek ve parçacığının yeni konumunu elde etmek için Denklem 4.22 ve 4.23 uygulanmaktadır [140].

$$v_i(t) = sig(v_i(t)) = \frac{1}{1 + e^{-v_i(t)}} \quad (4.22)$$

$$x_i(t+1) = \begin{cases} 1, & \text{eğer } rand < v_i(t+1) \\ 0, & \text{değilse} \end{cases} \quad (4.23)$$

Burada $rand$ değeri $[0,1]$ arasında rastgele bir sayıdır. Özellik problemini bir optimizasyon problemi olarak formüle etmek için iki temel konu dikkate alınmaktadır. İlki çözüm temsilidir. Çünkü özellik seçimi, ikili bir optimizasyon problemi olduğundan dolayı çözümü ikili bir vektörle temsil edilmelidir. Burada, 1 değerleri ilgili özelliğin seçildiğini, aksi takdirde seçilmediğini göstermektedir. Çözüm boyutu, her veri kümesindeki özelliklerin sayısıdır. İkinci konu ise, sarmalayıcı yaklaşım benimsendiği için, seçilen özelliklerin sayısı kadar sınıflandırma doğruluğu da dikkate alınmaktadır. Uygunluk fonksiyonu ile hem bilgilendirici özelliklerin sayısı azaltmak ve yüksek doğruluğu elde etmek arasında denge sağlanmaktadır [141].

4.4.2. Özyinelemeli Özellik Eleme (ÖÖE) Algoritması

Özellik seçimi, makine öğrenmesi uygulamalarında verilerin öğrenme performansını iyileştirmesi hedeflenerek yaralı bilgilerin kaybetmeden, optimum ve sağlam özelliklerin seçilmesi, alakasız ve öğrenmeye katkı sağlamayan özelliklerin dışarıda tutulması, boyutluluğu azaltılması işlemidir [142]. Alakasız özelliklerin öğrenmede kullanılması modelin performansını olumsuz etkileyebilmektedir [143]. Bu boyutsallığın azaltılması hesaplama maliyetini ve eğitim süresini azaltırken veri kaybına neden olmaktadır [142].

ÖÖE algoritması genellikle DVM sınıflandırıcısı ile birlikte kullanıldığından, özelliklerin önem derecelerine göre sıralarken [144], DVM sınıflandırıcısının ağırlıklarından faydalanmaktadır [145]. Özellik seçiminde bir özellik, özellik alt kümelerinde de değerlendirildiği için özyinelemeli bir yaklaşım benimsenmektedir [144]. Sınıflandırmada etkisi en az etkisi olan özellikler her yinelemede kaldırılarak sınıflandırma için en uygun özellikler seçilmektedir. Doğrusal DVM için bir ağırlık vektörü eşitliği Denklem 4.24’de belirtilmiştir [146].

$$W_V = \sum_{k \in SV} y_k \alpha_k x_k \quad (4.24)$$

Burada α_k , lagrange katsayısı, y_k , x_k eğitim örneklerinin etiketidir.

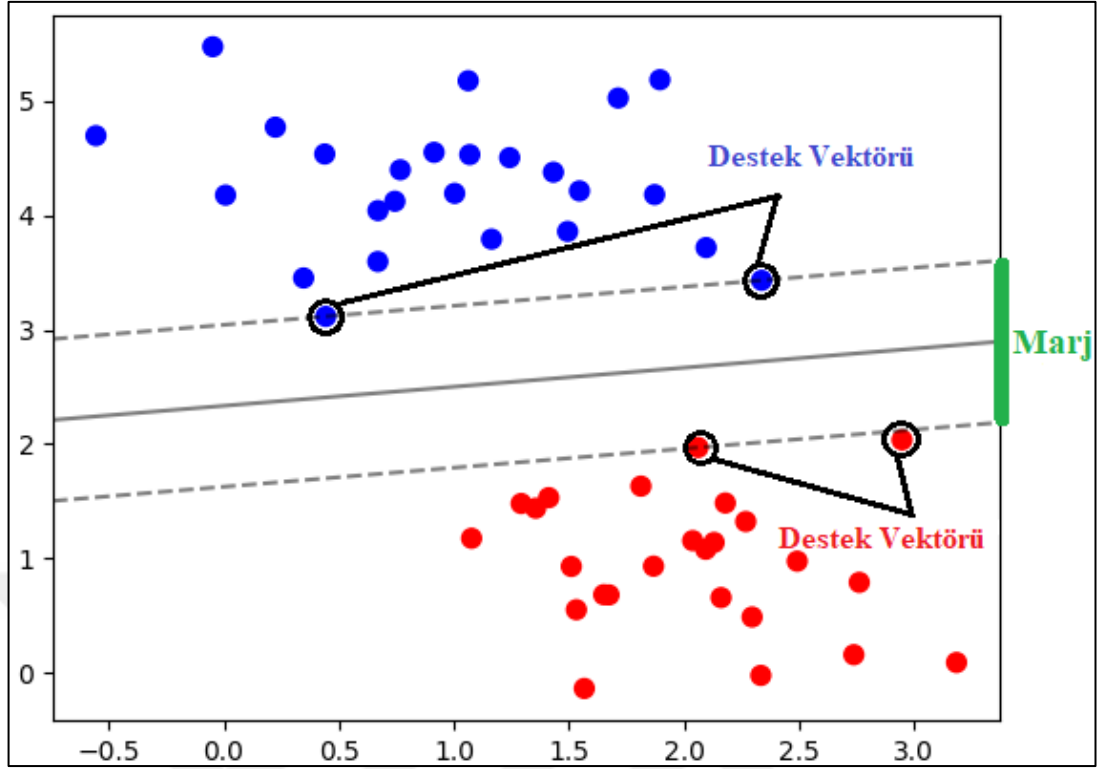
4.4.3. Çapraz Doğrulama ile Özyinelemeli Özellik Eleme (ÇDÖÖE)

Tez kapsamında önerilen XceptionNet modelinde ÖÖE algoritmasının bir güncellenmiş versiyonu olan çapraz doğrulama ile özyinelemeli özellik eleme özellik seçim algoritması kullanılmıştır. ÇDÖÖE, özellik haritalarının sınıflandırmadaki önemlerine göre otomatik çapraz doğrulama kullanarak sıralamaktadır. Ardından algoritma sıralanmış özellik haritaları içerisinde Mİ sınıflandırmasında en alakasız özellikleri birer birer kaldırmaktadır. Bu algoritma adından da anlaşılacağı üzere özellik haritalarının sıralarken özyinelemeli bir işlev kullanıp sınıflandırmada en

yüksek puanı verecek minimum sayıdaki özellikleri belirleyip seçmektedir. Böylece sınıflandırının performansını artmasına yardımcı olmaktadır. [140,141].

4.5. SINIFLANDIRNMA DESTEK VEKTÖR MAKINELERİ

Destek vektör makinesi (DVM) [149], etkili sınıflandırma ve regresyonda analizinde kullanılan bir yöntemdir [39,143]. DVM, iki sınıf arasındaki maksimum marjlı doğrunun denklemini elde ederek sınıflandıran denetimli öğrenme yöntemlerinden biridir [151]. Sınıfların birbirine en yakın olduğu sınır noktaları destek noktaları olarak adlandırılmaktadır. Destek konumları, sınıfları birbirinden ayıran optimum hiper düzlemi elde etmeye yarayan karar noktaları olarak bilinmektedir. Bu noktaların oluşturduğu boşluğu tanımlayan vektörlere de destek vektörleri denmektedir. Destek Vektörleri ve marj aralığının grafiksel temsili Şekil 4.16’de sunulmuştur. Ayrıca DVM’nin matematiksel olarak Eşitliği 4.25’de verilmiştir [151].

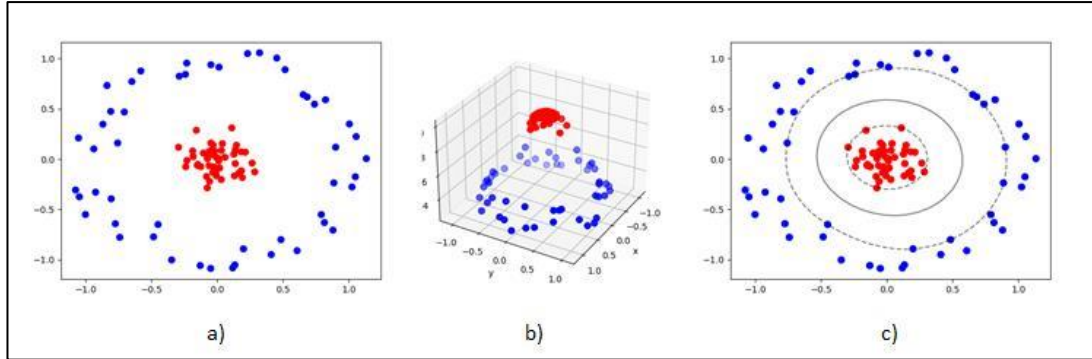


Şekil 4.16. DVM sınıflandırma işlemi ile örneklerin ayrılması.

$$\begin{cases} \min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \\ y_i(w_t x_i + b) \geq 1 - \xi_i, i = 1, 2, \dots, n, \\ \xi_i \geq 0 \end{cases} \quad (4.25)$$

Burada eğitim veri seti doğrusal olarak ayıramadığı durumlarda, daha yüksek bir uzaya eşlemenin maliyetli bir durum olduğu göz önüne alınırsa, doğrusal ayrılabilirdiği daha yüksek bir uzay boyutunda simüle etmek için çekirdek hilesi yöntemi kullanılır [152]. Veriler aynalanarak farklı çekirdek fonksiyonları ile sınıflara ayrılabilir [151]. Çekirdek hilesi adı verilen işlemlerin 2 boyutlu ve 3 boyutlu düzlemlerdeki işleminin görüntüsü Şekil 4.17’de gösterilmiştir. Çekirdek hilesi işlemini hesaplamada kullanılan genel fonksiyon denklemi Eşitlik 4.26’da verilmiştir [153].

$$k(x_i, x_j) = \phi(x_i)^T \phi(x_j) \quad (4.26)$$



Şekil 4.17. a) Doğrusal olarak ayrılmayan veri kümesi b) verilerin daha yüksek bir uzay boyutunda yansıtılması c) herhangi bir çekirdek fonksiyonu ile DVM birlikte kullanılarak sınıflandırılması.

DVM istatistiksel öğrenme, veri optimizasyonunun modellenmesi, nesne algılamanın yanı sıra Mİ tanıma görevlerinde de başarılı bulunduğundan dolayı araştırmalarda kullanılmışlardır [10,34,46].

4.6. PERFORMANS ÖLÇÜTLERİ

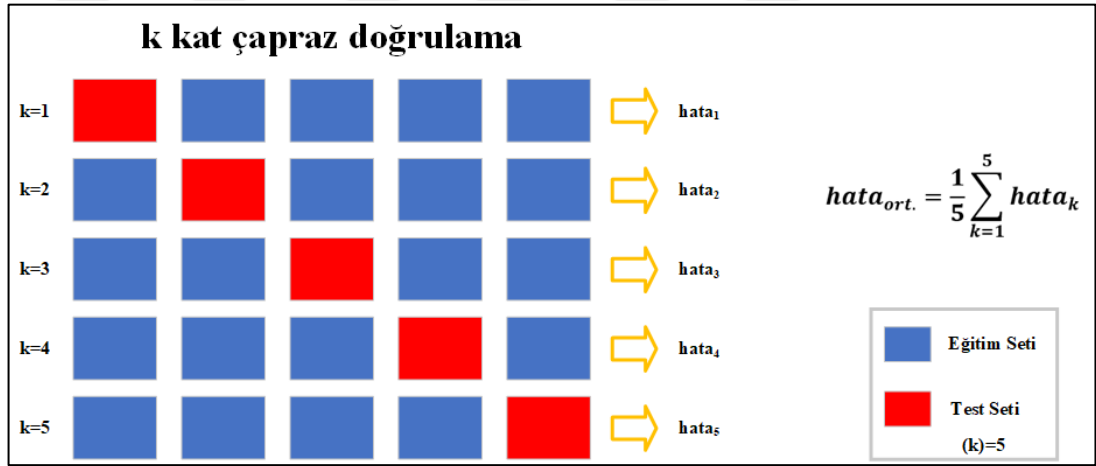
Bu tez kapsamında sınıflandırma için derin öğrenme modellerinin yapısında bulunan ve başlık 4.5.1.4’değınilen softmax fonksiyonu, doğrusal DVM ve farklı DVM çekirdekleri deneylerde kullanılmıştır.

4.6.1. K-Katlı Çapraz Doğrulama

Makine öğrenmesinde kullanılmak üzere seçilen modellere ait sınıflandırma algoritmalarının performansı değerlendirilirken gerçek hata olarak da adlandırılan tahmin hatası adı verilen bir değerlendirme ölçütü bulunmaktadır. Bir çok gerçek dünya probleminde bu hata ölçütünün değeri tam olarak hesaplanamaz ve gerçek değeri bilinemez olsa da verilerden tahmin edilmektedir [154]. Tahmin etme işlemi için faydalanılan en popüler [155] ve yaygın teknik olarak çapraz doğrulama tercih

edilmektedir [156]. Bu teknikte bir veri setinin k eşit alt parçaya rastgele bölünüp birini dışarda bırakarak (k-1) parçası ise modeli öğrenmek için kullanılan eğitim seti, bir parçasını ise modeli değerlendirmek için kullanılan test seti olarak kullanılmaktadır [157]. Veri seti rastgele bölündüğünden dolayı özellikler homojen olarak gruplara ayrılmasa bile, her bir bir parça sırayla test seti geri kalan parçalar eğitim seti olduğunda hem bölünmeden kaynaklanan sapma ve hatalar asgariye indirilir hem de aşırı uyum (overfitting) önlenmiş olur.

Ayrıca bu tez kapsamı içinde önerilen modellerde 5 ve 10 kat çapraz doğrulama uygulanmış olup, Şekil 4.18’de beş kat çapraz doğrulamanın nasıl uygulandığının bir görseli sunulmuştur.



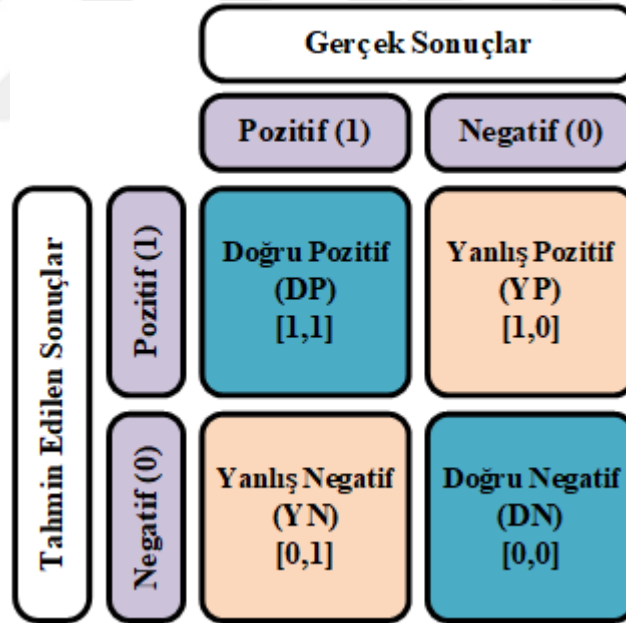
Şekil 4.18. Beş kat çapraz doğrulamaya ait bir örnek uygulamanın grafiksel gösterimi.

4.6.2. Karmaşıklık Matrisi

Örüntü tanımada yaygın olarak kullanılan karmaşıklık matrisi, eğitilen ağın performansını değerlendirmek [158] ve hangi sınıfların daha kolay hangi sınıfların daha zor ayırt edildiğini belirlemek ve kullanılmaktadır [159]. Veri setlerine genellikle çapraz doğrulama tekniği uygulandığında yapılan sınıflandırma işleminin geneline ait ortalama doğruluğu ve ortalama hata bilgisi elde edilmektedir. Karmaşıklık matrisi, yalnızca doğruluk ve hatadan çok daha ayrıntılı bilgi

sağlamaktadır [153,154]. Veri setinde bulunan her sınıfın model tarafından tanınma performansı hakkında hususi bilgiler karmaşıklık matrisi tarafından verilmektedir. Bunun yanı sıra hangi sınıfların hangi sınıflarla karıştırıldığı, hangi sınıfların tama yakın hangi sınıfta düşük tanıma yapıldığı, bilgilerini veren bir matristir. Karmaşıklık matrisinin köşegenleri sınıflandırıcının her sınıf için doğru tahminlerini, köşegen dışında kalan bölgeler ise tanıma hatasını temsil etmektedir [162].

Karmaşıklık matrisinde kullanılan parametreler, doğru Pozitif (DP), Yanlış Pozitif (YP), Yanlış Negatif (YN) ve Doğru Negatif (DN)'dir. Gerçek değerlerle tahminler karşılaştırıldığında pozitif bir sınıf pozitif olarak tahmin edilirse DP, negatif bir sınıf negatif olarak tahmin edilirse DN, pozitif bir sınıf negatif olarak tahmin edilirse YP, negatif bir sınıf pozitif olarak tahmin edilirse YN olarak adlandırılmaktadır. Şekil 4.19'da iki sınıflı bir veri setinde karmaşıklık matrisinin parametrelerinin nasıl elde edildiğinin görseli sunulmuştur.



Şekil 4.19. İki sınıflı bir veri setinde karmaşıklık matrisinin parametrelerinin elde edilmesi.

Bu parametrelerden faydalanılarak bir takım ölçüt birimleri bulunmaktadır. Bunlar doğruluk, hassasiyet (duyarlılık), özgüllük, kesinlik ve F-skor ölçüm birimleridir. Doğruluk, yapılan tahminlerin kaç tanesinin doğru olduğunu veren ölçümdür. Doğru tahminlerin toplam tahminlere oranıdır. Sınıflandırıcının doğru tahmin sıklığını verir. Dengeli sınıflarda doğruluk geçerli bir değerlendirmedir. Duyarlılık, gerçek pozitif oranı verir, pozitif sınıfın ne kadar iyi tahmin edildiğini, özgüllük ise negatif sınıfın ne kadar iyi tahmin edildiğini özetler. Kesinlik, pozitif olarak tahmin edilen örneklerin, kaç tanesinin gerçekte pozitif olduğunu veren bir değerlendirmedir. F-skor ölçümü kesinlik ve duyarlılık değerlerinin harmonik ortalamadır. Kesinlik ve duyarlılık arasında bir denge sağlar. Bu ölçümlerin matematiksel denklemleri Eşitlik 4.27 ve Eşitlik 4.31 arasında verilmiştir [163].

$$\text{Doğruluk} = \frac{DP + DN}{DP + YP + DN + YN} \quad (4.27)$$

$$\text{Duyarlılık} = \frac{DP}{DP + YN} \quad (4.28)$$

$$\text{Özgüllük} = \frac{DN}{YP + DN} \quad (4.29)$$

$$\text{Kesinlik} = \frac{DP}{DP + YP} \quad (4.30)$$

$$F_{\text{skor}} = \frac{2 * \text{Kesinlik} * \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} \quad (4.31)$$

4.6.3. Alıcı İşlem Karakteristiği (AİK)

Makine öğrenmesinde sınıflandırıcıların performansının değerlendirilmesinde son yıllarda yaygın olarak kullanılan Alıcı İşletim Karakteristiği (Receiver Operating Characteristic) eğrileri isabetli ve yanlış tahminleme arasındaki oranı görsel olarak göstermektedir [164]. Modele ait ölçülen duyarlılık (hassasiyet) ve özgüllük değerleri öncelikle bir tablo şekline getirildikten sonra x ekseninde (1-özgüllük) ve y ekseninde (duyarlılık) değerleri olacak biçimde oluşturulan (x,y) noktalarının analitik grafikteki konumu ile üretilen eğrisel çizgiye AİK eğrisi denmektedir [165]. Yani yanlış pozitif oranları ile gerçek pozitif oranlar arasındaki ilişkinin grafiksel

gösterimi AİK eğrisi olarak adlandırılmaktadır. Özetle, duyarlılığın, 1-özgüllüğün bir fonksiyonu olarak grafiksel temsili denilebilir [166].

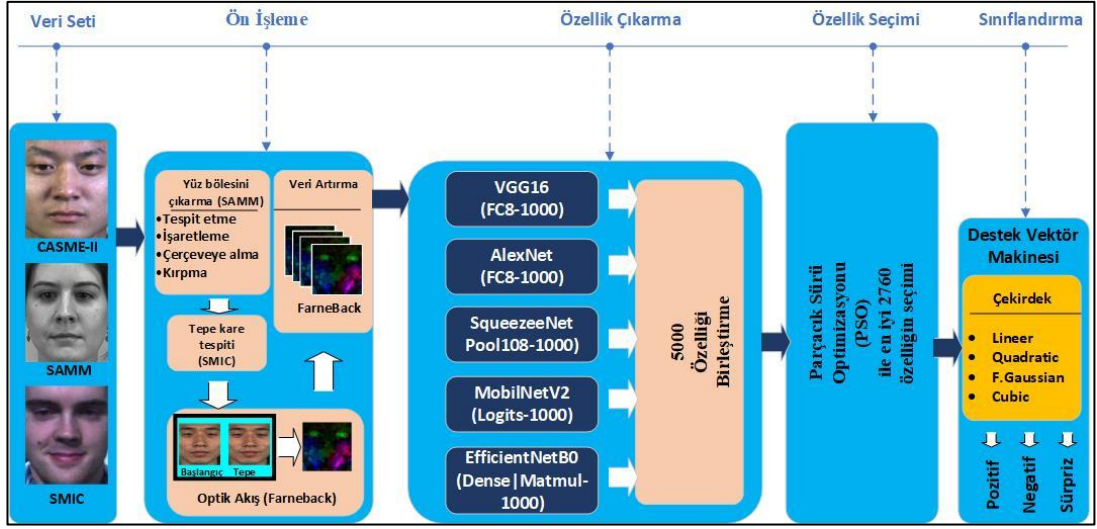
Ayrıca sınıflandırma algoritmalarının doğruluğunu değerlendirmek için AİK eğrisinin altındaki alan da kullanılmaktadır [167]. Bu alan yamuk kuralı kullanılarak hesaplanmaktadır [168]. AİK eğrisinin altındaki alan, sınıflandırıcının rastgele seçilmiş örnekler arasında ne kadar doğru bir ayırım yaptığının değerlendirmesi için kullanılmaktadır [169]. ROC eğrisinin $y=x$ referans doğrusunun üstünde ve sol tarafta bulunan üçgende bulunması beklenmektedir. Modelin 0,5 AİK eğrisinin altındaki alan değeri için ayırt etme yeteneğinin olmadığını, 0,7 ile 0,8 kabul edilebilir, 0,8 ile 0,9 mükemmel olarak ve 0.9 dan büyük değerler olağanüstü kabul edilmektedir [170].

BÖLÜM 5

ÖNERİLEN MİKRO İFADE TANIMA METODLARI

5.1. DERİN TRANSFER ÖĞRENME VE PARÇACIK SÜRÜ OPTİMİZASYON ALGORİTMASI İLE MİKRO İFADE TANIMA (DTÖ-PSO)

Bu tez kapsamında önerilen Derin Transfer Öğrenme ve Parçacık Sürü Optimizasyon Algoritması (DTÖ-PSO) ile Mİ tanıma modelinde tek bir ağı eğitilmesinden elde edilen özellikler yerine beş farklı ağdan özellikler elde edilerek birleştirilmiş, her ağı kendine ait ayırt edici üstünlüklerinden faydalanılarak daha değerli ve farklı özelliklerin elde edilmesi hedeflenmiştir. Bu çalışmada önerilen modelin mimarisinde kullanılan derin transfer modelleri olan VGG16, AlexNet, SqueezeNet, MobilNetV2 ve EfficientNetB0 ağlarının her birinin eğitilmesinden sonra elde edilen 1000 özellik alınmıştır. Derin ağlarından elde edilen bu özelliklerden bir kısmı başarıya olumsuz etki etmektedir. İyi performans sonuçları elde edebilmek için çıkarılan özelliklerin en iyilerinin seçilmesi gerekmektedir. Bu çalışmamızda en iyi özellik seçimi için PSO algoritması kullanılmıştır. PSO'dan elde edilen en iyi özellikler DVM sınıflandırma algoritması kullanılarak yüksek performanslı bir Mİ tanıma modeli geliştirilmiştir. DTÖ-PSO hibrit modelinde önışleme, özellik çıkarma, özellik seçme ve sınıflandırma olmak üzere dört adımdan oluşmaktadır. Modelin yapısı Şekil 5.1' de verilmiştir.



Şekil 5.1. Mİ sınıflandırma görevi için önerilen DTÖ-PSO modelinin mimari yapısı.

DTÖ-PSO modelinde kullanılan ESA modellerinin, özellik seçimi algoritmasının ve sınıflandırma algoritmalarının sebepleri;

- AlexNet mimarisinde, o zamana kadar standart olmuş tanh ve sigmoid yerine fonksiyonu yerine ReLU aktivasyon fonksiyonunu kullanması, aşırı uyum sorununu çözmek için bırakma katmanını kullanması [142], görüntü sınıflandırma görevlerinde etkili sonuçlar vermesi [171], kendinden önceki yöntemlere göre daha iyi performans göstermesi [172] gibi avantajları bulunmaktadır. Derin Öğrenmenin katalizörü olarak kabul edildiği ve görüntü sınıflandırmasında lider mimari olduğunu [173] için modelimizde AlexNet kullanılmıştır.
- VGG16 mimarisi kendinden önceki mimarilerin aksine 3x3 küçük boyutlu evrişim filitreleri kullanarak parametre sayısını düşürmesi, kendinden önceki mimarilerden daha iyi performans göstermesi, önemli ölçüde artan derinlik gibi avantajlarından dolayı modelimizde tercih edilmiştir [96].
- SqueezeNet ImageNet verisetinde AlexNet ile aynı doğruluğu sağlarken AlexNet'in 1/50'si kadar parametre kullanmaktadır. Özellik haritalarının boyutları daha sonraki katmanlarda düşürülerek büyük özellik haritaları elde edilip, doğruluğu arttırmaktadır [129].

- MobilNetV2’de ağında hesaplamaları azaltmak için pahalı standart evrişimler yerine daha ucuz olan, derinlemesine ayrılabilir evrişimler (DAE) kullanan bir model geliştirilmiştir. DAE sadece genişletme yapmayıp, kanal sayısı hem genişletilip hem de azaltılmaktadır [131].
- EfficientNet ile gelen en büyük yenilikler bileşik ölçekleme ve sinir mimarisi araması kullanılmasıdır. Bileşik ölçekleme ile derinlik, genişlik ve çözünürlük parametrelerini eşit oranda ölçekleyip performansını artırmışlardır. Sinir mimarisi araması ile eğitim hızını ve parametre verimliliğini birlikte optimize edip, bir sinir ağı için en iyi mimariyi otomatik olarak tasarlamaktadır. Ayrıca aktivasyon fonksiyonu olarak ReLU’dan daha iyi çalışma eğiliminde olan swish aktivasyon fonksiyonu kullanılmaktadır [133].
- Veri setimizdeki örnekleme sayısının azlığından dolayı aşırı uyum sorununu çözmek ve eğitim süresini düşürmek için modellerimiz transfer öğrenme ile birlikte kullanılmıştır [174].
- PSO en uygun çözümü bulmak için yerel ve küresel arama işlemlerini dengeleyen, verilen zaman aralığında en doğru ve etkin çözümü bulan yapay zekaya dayalı bir algoritmadır [175]. PSO da kullanılan tüm parçalar birbirleriyle iletişimli ve eş güdümlü bir şekilde çalışmaktadır ve bilgi kaybı yoktur [176].
- Makro ve mikro ifadenin yanı sıra diğer sınıflandırma görevleri içinde yaygın olarak kullanılan DVM, doğrusal olarak ayrılmayan sınıfları ayırmak için hiper düzlem kullanılmaktadır [177]. DVM'nin sağlamlık ve genelleme hatalarının düşük olması[178], hesaplama süresindeki avantajlarından dolayı [179] DTÖ-PSO modelinde kullanılmıştır.

5.1.1. Veri Seti

Mevcut veri seti ve örnekleme sayısının sınırlı olması ve sınıflarında bulunan dengesiz örnek dağılımları Mİ tanıma görevi için mevcut zorluklardandır. SAMM veri seti için korku örnekleme sadece 8 tane iken öfke örnekleme 57 tanedir. Bu sebepten ötürü bir çok çalışmada olduğu gibi [17,20,38] bu çalışmada da en yaygın olarak kullanılan üç spontane Mİ veri seti CASME II, SAMM ve SMIC birleştirilerek bileşik veri setleri

oluşturulmuştur. Her üç veri setinde aynı sayıda duygu sınıfı bulunmadığından dolayı en az ve üç duygu sınıfına sahip olan SMIC-HS veri seti, bu çalışmada hedef sınıf sayısı olarak belirlenmiştir. Yedi sınıflı CASME II veri seti ve sekiz sınıflı SAMM veri seti SMIC veri setine eşleştirilmiştir. SMIC veri setindeki örnekler negatif, pozitif ve sürpriz olmak üzere üç duygu sınıfıdır. Bu eşleştirmede diğer iki veri setinde bulunan baskı, kızgınlık, hor görme, iğrenme, korku ve üzüntü sınıfları negatif sınıf altında gruplandırılırken, mutluluk sınıfı pozitif, sürpriz sınıfı isim değiştirmeden aynı isimle bulunmaktadır. Böylece bu üç veri seti birleştirilerek negatif sınıfında 266 adet örnekleme, pozitif sınıfında 112 adet örnekleme ve sürpriz sınıfında 88 adet örnekleme sahip yeni bileşik veri seti elde edilmiştir. Önerilen DTÖ-PSO modeli için kullanılan bileşik veri setine ait özellikler Çizelge 5.1’de sunulmuştur.

Çizelge 5.1. DTÖ-PSO modelinde kullanılan bileşik veri setindeki örneklerin özellikleri.

Birleştirildikten sonraki sınıf isimleri	Pozitif	Negatif	Sürpriz
Birleştirilmeden önceki sınıf isimleri	mutluluk	baskı, kızgınlık, hor görme, iğrenme, korku ve üzüntü	sürpriz
Örnek Sayısı	112	266	88

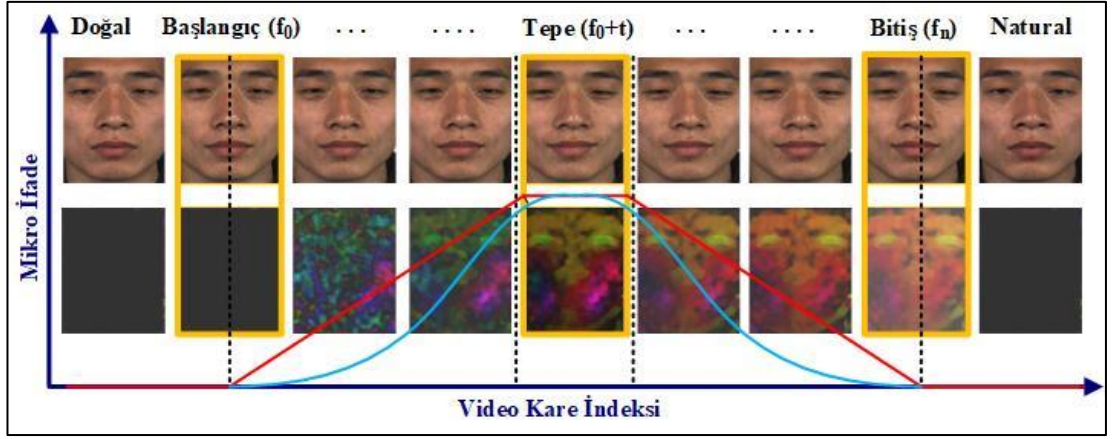
5.1.2. Ön İşleme

Mİ ile yapılan çalışmalarda sekans bazlı (video) ve tepe kare bazlı olmak üzere iki tip çerçeve seçilmiştir [24]. Bu çalışma, son araştırmalarda kullanılan tepe kare tabanlı çerçeveyi kullanarak en yoğun ifade çerçevesine odaklanmaktadır. Tepe tabanlı tasarımda sadece iki anahtar kare dikkate alınmaktadır. Bunlar başlangıç karesi ve tepe karesidir.

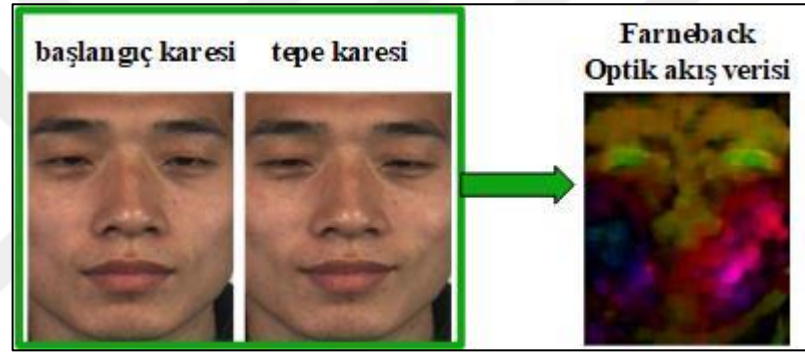
Ön işlemede sürecinde SMIC veri setine ait örneklerin anahtar kareleri (başlangıç, tepe ve bitiş) kullanıcılara sunulmadığından dolayı ilk olarak bölüm 4.2.1’de anlatılan

D&C-RoIs yöntemi ile tepe kare konumu tespit edilmesi işlemi gerçekleştirilmiştir. Kısaca bu yöntemle bir video verisinde bulunan kare dizisindeki görüntülere ait verilerin LBP özellikleri çıkarılmaktadır. Ardından böl ve fethet tekniği ile her karede bulunan yüz bölgeleri RoI bölgelerine bölünmekte ve referans olarak kullanılan başlangıç karesi ile diğer tüm karelerin özellik farkları RoI'ler için hesaplanmaktadır. En yüksek özellik farkının bulunduğu RoI karesi tepe karesinin yaklaşık konumunu vermektedir.

SAMM veri setinde verilen örneklerde ise herhangi bir ön işlem aşamasından geçmediği için bu veri setine ait ham örneklerde ise yüz bölgesinin algılanması, işaretlenmesi, hizalanması ve kırılarak çıkarılması işlemlerinden elde edilen Mİ görüntüleri Bölüm 4.2.2'deki işlemler gerçekleştirilerek yüz bölgesi orijinal görüntülerden çıkarılmış ve yeni görüntü verileri elde edilmiştir. Bir sonraki adımda normalleştirilmiş Mİ görüntülerine Farneback optik akış metodunu uygulamak için her video klip dizisinden referans olarak başlangıç karesi ve duygunun en yoğun olduğu tepe kareleri seçilmiştir. Böylece bu iki anahtar karelerin görüntülerinden Mİ örneklerine ait hareket özelliklerinin elde edildiği yeni bir optik akış görüntüsü elde edilmiştir. Ardından bu optik akış görüntüleri Bölüm 4.2.4'de belirtildiği gibi veri artırma yönteminin döndürme tekniği ile 90^0 , 180^0 , 270^0 derece döndürülerek yapay olarak artırılmıştır. Bu artırma sonucunda tüm sınıflar eşitlenerek en yüksek örneğe sahip olan sınıf negatif sınıfın sayısına eşitlenmiş ve 266 olarak ayarlanmıştır. Şekil 5.2 ve 5.3 'de başlangıç karesi ile diğer kareler kullanıldığında elde edilen optik akışın görüntüleri verilmiştir. Buna göre en yüksek hareket özelliği elde edilebilmesi için başlangıç ve tepe karesinin kullanılması gerekmektedir.



Şekil 5.2. Mİ video dizindeki başlangıç bitiş ve tepe karlerlerin konumları.



Şekil 5.3. Başlangıç ve tepe karesi kullanılarak optik akışın elde edilmesi.

5.1.3. Özellik Çıkarımı

Artırılmış veri setine ESA modelleri uygulanarak her bir modele ait son tam bağlı katmandan görüntülere ait özellik haritaları alınmıştır. Beş farklı ESA modelinden elde edilen 1000' er adet özellikler birleştirilerek 5000 adet özellik elde edilmiştir. DTÖ-PSO modeli için gerçekleştirilen tüm deneylerde modellerin ağırlıkları ImageNet ağırlıklarından faydalanılarak transfer öğrenmesi gerçekleştirilmiştir. ESA modellerinin tamamında yapılan ölçümlerin hiper parametreleri aynı ve maksimum epoch değeri 32, mini batch size değeri 16, Initial Learn Rate değeri 0.001, Validation Frequency değeri 50, Learn Rate Drop Factor değeri 0.1, Learn Rate Drop Period

değeri 16 olarak ayarlanmıştır. ESA modellerinin son bağlı katmanlarının isimleri ve bu katmanlardan elde edilen özellikler Çizelge 5.2’de sunulmuştur.

Çizelge 5.2. DTÖ-PSOmodelinde ESA modellerinden elde edilen özellikler.

Model	Giriş Boyutları	Modele ait Son Bağlı Elde Edilen Katman Özellik
VGG16	224x224	FC8 1000
AlexNet	227x227	FC8 1000
SqueezeNet-	227x227	Pool10 1000
MobileNetV2	224x224	Logits 1000
EfficientNetB0	224x224	dense MatMul 1000

5.1.4. PSO Optimizasyon Algortiması ile Özellik Seçimi

5 farklı ESA elde edilen toplam 5000 adet derin özellikler seçilen optimizasyon algoritmalarına en iyi özellikleri seçmek için verilmiştir. Algoritma sonuçlarına göre 2746 adet derin özellik seçilip DVM sınıflandırmasına giriş verisi olarak verilmiştir.

PSO algoritması ile Mİ tanımaya yönelik görüntülerin içerdiği en iyi özellikler filtrelenerek özellik seçme adımı gerçekleşmiştir. Bu çalışmada PSO optimizasyon algortiması özellik seçiminde kullanılmıştır. Deneyde kullanılan parametreler ise çıkarılan özellik sayısını veren popülasyon boyutu (D) 1000, popülasyon sayısını veren popülasyon boyutu (N) 10, popülasyon matrisi (X) $0 < x_i^t < 1$ aralığında rastgele olarak ayarlanmıştır. $x_i^t > 0,5$ olduğu durumlarda özellik seçimi yapılmış aksi durumlarda özellik seçimi yapılmamıştır. İterasyon sayısı 100 alındığında fonksiyon değerlendirme sayısı değeri $N \times 100 = 1000$, en yakın komşuluk (EYK) sayısal (k) değeri ise 5’tir. PSO parametleri ise bilişsel faktör değeri 2, sosyal faktör değeri 2, atalet ağırlığı değeri ise 0,9 olarak alınmıştır. Uygunluk hesaplama eşitliği Denklem 5.1 ile Denklem 5.3 arasında verilmiştir [180].

$$Hata\ oranı = 1 - Doğruluk \quad (5.1)$$

$$SÖ = \sum_{i=0}^D x_i = 1 \quad (5.2)$$

$$Fitnes\ Function = \alpha * ErrorRate + \beta * (SF/D) \quad (5.3)$$

Burada SÖ toplam seçilen özellikleri temsil ederken, “Doğruluk” seçilen özelliklerin EYK sınıflandırmasında elde edilen doğruluk oranı, $\alpha=0.99$, $\beta=0.01$ olarak alınmıştır.

5.1.5. Sınıflandırma

Veri çapraz doğrulama tekniği kullanılarak veri seti 10 parçaya bölünmüş, 1 parçası test verisi, kalan 9 parçası eğitim verisi olarak ayarlanmıştır ($k=10$). Bu işlem iteratif olarak tüm parçaların hem eğitim hemde test verisi olarak kullanılmasına olanak sağlar. Böylece veri seti üzerindeki her örneklem hem veri seti hem de eğitim seti olarak görev yapar. Buda daha verimli bir sonuç elde edilmesini sağlamaktadır [150].

Son olarak sınıflandırma aşamasında DVM algoritmasının farklı çekirdekleri kullanılarak Mİ sınıflandırılması gerçekleştirilmiştir. Bu çalışmada da DVM ile doğrusal, quadratic, fine gaussian ve cubic çekirdek fonksiyonlarını kullanılmış, Mİ tanıma görevindeki üstünlükleri karşılaştırılmıştır. Çalışmamızda kullanılan çekirdeklerin matematiksel denklemleri Eşitlik 5.4 ve Eşitlik 5.7 arasında verilmiştir [151].

$$\text{Doğrusal DVM} \rightarrow k(x_i, x_j) = x_i^T x_j + c, \quad c \text{ bir sabittir.} \quad (5.4)$$

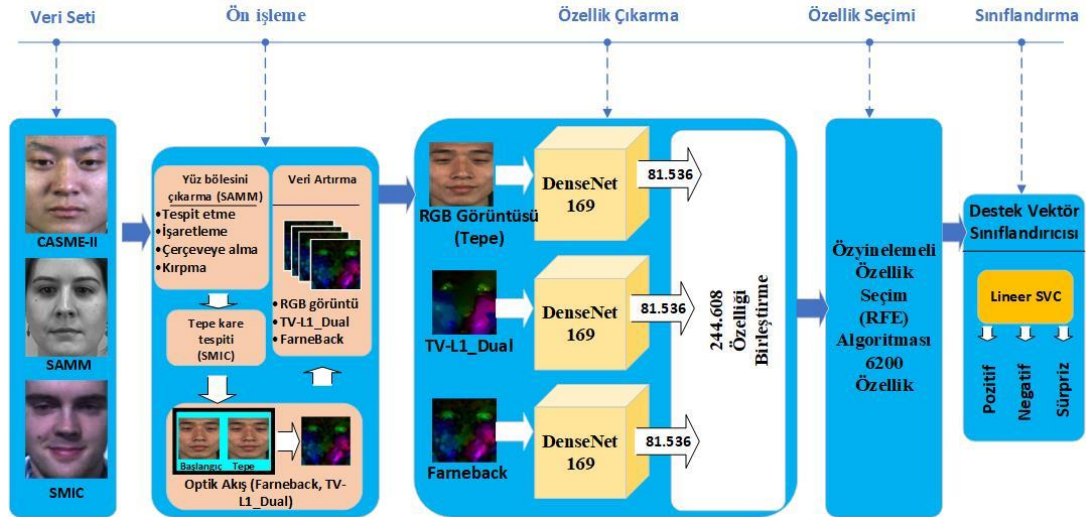
$$\text{Guassian} \rightarrow k(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (5.5)$$

$$\text{Cubic} \rightarrow k(x_i, x_j) = (x_i^T x_j + 1)^3 \quad (5.6)$$

$$\text{Quadratic} \rightarrow k(x_i, x_j) = 1 - \frac{\|x_i - x_j\|^2}{\|x_i - x_j\|^2 + c} \quad (5.7)$$

5.2. DENSENET169 MODELİ İLE ÖZYİNELEMELİ ÖZELLİK ELEME (DensNet169-ÖÖE)

Bu tez kapsamında önerilen DenseNet169 modeli ile özyinelemeli özellik eleme (DensNet169-ÖÖE), ve destek vektör makinaları tabanlı modelinde kullanılan ağ olan DenseNet teknolojisinin önemli avantajlarından biriside diğer son model teknolojilerle kıyaslandığında daha az parametre ve hesaplama maliyeti gerektirmenin yanı sıra parametre artışı ile doğru orantılı bir başarı göstermesidir [134]. Ayrıca modelimizde kullanılan ÖÖE'nin üstünlüğü ise sınıflandırma için kullanılan özelliklerden etkin olmayanları yinelemeli olarak eleyerek en uygun özellik kümesini seçme yeteneğidir [181]. Önerilen modelde, tercih edilen ESA mimarisinin ve özellik seçim algoritmasının bu avantajlarından faydalanılmıştır. Bu model için uygulanan deneylerde DenseNet121, DenseNet169 ve DenseNet201 modelleri Mİ tanıma görevindeki üstünlükleri karşılaştırılmıştır. DensNet169-ÖÖE hibrit modelinde önışleme, özellik çıkarma, özellik seçme ve sınıflandırma olmak üzere dört adımdan oluşmaktadır. Modelin yapısı Şekil 5.4' de verilmiştir.



Şekil 5.4. Mİ sınıflandırma görevi için önerilen DensNet169-ÖÖE modelinin mimari yapısı.

5.2.1. Veri Seti

Mevcut veri seti ve örneğinin sınırlı olması, veri setlerine ait sınıflardaki dengesiz örnek dağılımlarından dolayı en yaygın kullanılan üç spontane Mİ veri seti CASME II, SAMM ve SMIC birleştirilerek önerilen bu modelde de bileşik yeni veri seti kullanılmıştır. Bileşik veri setinde kullanılan örnekler ve sınıflarına ait genel bilgi Çizelge 5.3’de sunulmuştur.

Çizelge 5.3. DensNet169-ÖÖE modelinde kullanılan bileşik veri setindeki örneklerin özellikleri.

Birleştirildikten sonraki sınıf isimleri	Pozitif	Negatif	Sürpriz
Birleştirilmeden önceki sınıf isimleri	mutluluk	baskı, kızgınlık, hor görme, iğrenme, korku ve üzüntü	sürpriz
Örnek Sayısı	109	252	83

Bu modelde DTÖ-PSO modelindeki aynı veri setleri kullanılırken örnek sayısının değişmesinin nedeni, aynı video klipte birden fazla Mİ içeren örneklerden sadece birinin tercih edilmesi olmuştur.

5.2.2. Ön İşleme

Ön işleme aşamasında DTÖ-PSO’da olduğu gibi bileşik veri seti için tepe kare tespiti, yüz bölgesinin çıkarılması, optik akış görüntülerinin elde edilmesi ve veri artırma tekniğinin kullanılması işlemleri benzer olarak DensNet169-ÖÖE modeli için de gerçekleştirilmiştir. DTÖ-PSO modelinden farklı olarak Farneback optik akış görüntülerinin yanı sıra RGB tepe kareleri görüntüleri, TV-L1 görüntüleri ve TV-L1 optik akış görüntülerinin bileşenleri olan TV-L1_yatay ve TV-L1_dikey optik akış görüntüleride kullanılarak, Mİ tanıma görevinde hem kendi aralarında karşılatırılmış hem de özellik çeşitliliği elde edilmesi hedeflenmiştir. Bölüm 4.2.4’de belirtildiği

gibi veri artırma yönteminin döndürme tekniği ile 90^0 , 180^0 , 270^0 derece döndürülerek yapay olarak arttırılmıştır. Bu artırma sonucunda tüm sınıflar eşitlenerek en yüksek örneğe sahip olan sınıf negatif sınıfın sayısına eşitlenmiş ve 252 olarak ayarlanmıştır.

5.2.3. Özellik Çıkarımı

Arttırılmış veri setine en iyi performansı sağlayan DenseNet169 ESA modelleri uygulanarak, modele ait bağlı katmandan hemen önceki katmandan RGB tepe karelerinin görüntülerinin, Farneback ve TV-L1 görüntülerinin özellik haritaları alınmıştır. Bunların her birinden 81536 adet, toplamda 244.608 adet özellik haritası elde edilmiştir.

Deneyisel ölçümlerin tamamında ortak olarak ImageNet ağırlıklarından faydalanılıp, transfer öğrenme gerçekleştirilmiştir. Buna göre giriş parametreleri aynı ve ‘weights’ değeri ‘imagenet’, ‘include_top’ değeri ‘False’, ‘input_shape’ değeri ‘224x224’, olarak ayarlanmıştır. Ayrıca ESA modellerinin özellik çıkarma işleminde batch_size değeri 12, olarak belirlenmiştir. Özellik haritaları sınıflandırma katmanında bulunan küresel ortalama havuzlamadan hemen önceki Relu Aktivasyon işleminden sonra elde edilmiştir.

5.2.4. ÖÖE Algoritması ile Özellik Seçimi

ÖÖE algoritması ile görüntülerin Mİ tanımaya yönelik içerdiği en iyi 6200 özellik filtrelenerek özellik seçimi süreci gerçekleşmiştir. ÖÖE algortması için ‘estimator’ parametresi DVM sınıflandırıcısı, ‘n_features_to_select’ parametresi ‘6200’, ‘step’ parametresi ‘1000’ olarak ayarlanmıştır.

5.2.5. Sınıflandırma

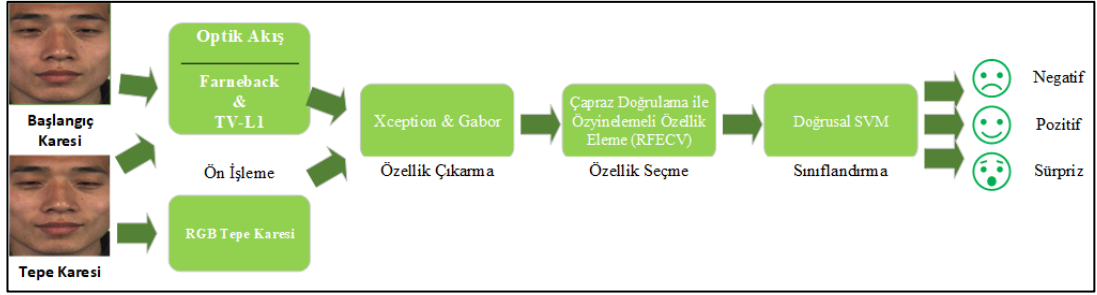
Veri çapraz doğrulama tekniği kullanılarak veri seti 5 parçaya bölünmüş, 1 parçası test verisi, kalan 4 parçası eğitim verisi olarak ayarlanmıştır ($k=5$). Bu işlem veri setindeki

tüm örneklerin hem test verisi hemde eğitim verisi olarak kullanılmasını sağlamaktadır.

Son olarak sınıflandırma aşamasında DVM algoritmasının doğrusal çekirdeği kullanılarak Mİ örnek görüntülerinin sınıflandırılması gerçekleştirilmiştir. DVM ‘kernel’ değeri ‘linear’, ‘C=1.0’, ‘probability’ değeri ‘True’ olarak, RepeatedKfold metoduna ait ‘n_splits’ parametre değeri ‘5’, ‘n_repeats’ parametre değeri ‘5’, ‘random state’ parametre değeri ‘42’ olarak ayarlanmıştır.

5.3. XCEPTIONNET İLE ÇAPRAZ DOĞRULAMALI ÖZYİNELEMELİ ÖZELLİK ELEME TABANLI MODEL (XNET-ÇDÖÖE)

Bu tez kapsamında XceptionNet, Çapraz Doğrulamalı Özyinelemeli Özellik Eleme tabanlı (Xnet-ÇDÖÖE) önerilen Xnet-ÇDÖÖE modelinde kullanılan ağ olan Xception teknolojisinin önemli avantajlarından biriside derinlemesine ayrıştırılabilir evrişimden ve noktasal evrişimden yararlanılarak parametre sayısını büyük oranda azaltıp hesaplama maliyetini azaltmaktadır [130]. ÇDÖÖE’nin ÖÖE’ye göre üstünlüğü ise en uygun özellik sayısını seçmek için çapraz doğrulanmış puanı kullanmasıdır [182]. Bu avantajlarından dolayı önerilen modelde Xception ESA mimarisi ve ÇDÖÖE özellik seçim algoritması tercih edilmiştir. Bu model çalışması için uygulanan deneylerde geleneksel yöntemlerden Gabor yöntemi ve veriye dayalı yöntemlerden ESA modeli olan Xception Mİ tanıma görevi için karşılaştırılmıştır. Bu tez kapsamında önerilen üçüncü modele, mimarisinde kullanılan ESA modelinin ve özellik seçme algoritmasının birleşiminden olan Xnet-ÇDÖÖE adı verilmiştir. Xnet-ÇDÖÖE hibrit modelinde önışleme, özellik çıkarma, özellik seçme ve sınıflandırma olmak üzere dört adımdan oluşmaktadır. Modelin yapısı Şekil 5.5’ de verilmiştir.



Şekil 5.5. Mİ sınıflandırma görevi için önerilen Xnet-ÇDÖÖE modelinin mimari yapısı.

5.3.1. Veri Seti

Önerilen bu çalışma için bileşik veri seti yerine diğerlerinden farklı olarak Mİ tanıma görevinde sadece CASME-II veri seti kullanılmıştır. CASME-II veri setinde kullanılan örnekler ve sınıflarına ait genel bilgi Çizelge 5.4’de sunulmuştur.

Çizelge 5.4. Xnet-ÇDÖÖE modelinde kullanılan CASME-II veri setindeki örneklerin özellikleri.

Modelde kullanılan sınıflar	Pozitif	Negatif	Sürpriz
Model için birleştirilmeden önceki sınıf isimleri	mutluluk	baskı, iğrenme, korku ve üzüntü	sürpriz
Örnek Sayısı	32	94	28

5.3.2. Ön İşleme

Ön işleme aşamasında CASME-II veri setinde anahtar kareler ve arka plandan çıkarılmış yüz bölgesine ait kareler kullanıcılara sunulduğu için önerilen bu modelde optik akış görüntülerinin elde edilmesi ve bu görüntülerin yapay olarak çoğaltılması işlemi gerçekleştirilmiştir. Farneback optik akış görüntülerinin yanı sıra RGB tepe kareleri görüntüleri, TV-L1 görüntüleri ve TV-L1 optik akış görüntülerinin bileşenleri olan TV-L1_yatay ve TV-L1_dikey optik akış görüntüleride kullanılarak, görüntülerin

hareket özellikleri elde edilmiştir. Böylece Mİ tanıma görevinde optik akış görüntüleri ve RGB görüntüleri hem kendi aralarında karşılaştırılmış hem de özellik çeşitliliği elde edilmesi hedeflenmiştir. Ön işleme aşamasının son adımında Bölüm 4.2.4’de belirtildiği gibi veri artırma yönteminin döndürme tekniği ile 90^0 , 180^0 , 270^0 derece döndürülerek veri seti yapay olarak arttırılmıştır. Bu artırma sonucunda tüm sınıflar eşitlenerek en yüksek örneğe sahip olan sınıf negatif sınıfın sayısına eşitlenmiş ve 94 olarak ayarlanmıştır.

5.3.3. Özellik Çıkarımı

Arttırılmış veri setine veriye dayalı yöntem olan Xception derin öğrenme modeli ve geleneksel yöntem olarak bilinen Gabor Filtresi modeli uygulanarak özellik haritaları çıkartılmış ve bu farklı iki yöntem kendi aralarında karşılaştırılmıştır. Xception ESA modelinin tam bağlı katmanlardan hemen önceki katmandan 100352 özellik haritası, Gabor modelinden ise 50176 özellik haritası elde edilmiştir.

Xnet-ÇDÖÖE modeli için gerçekleştirilen tüm deneylerde modellerin ağırlıkları ImageNet ağırlıklarından faydalanılarak transfer öğrenmesi gerçekleştirilmiştir. Ayrıca yapılan deneysel ölçümlerin hepsinde Xception modeline ait parametreler ‘weights’ değeri ‘imagenet’, ‘include_top’ değeri ‘False’, ‘input_shape’ değeri ‘(224,224,3)’, ‘batch_size’ değeri ‘12’, olarak belirlenmiştir. Sınıflandırmada kullanılmak için elde edilen özellik haritaları, küresel ortalama havuzlamadan hemen önceki ‘block14_sepconv2_act’ isimli aktivasyon katmanının çıktısından elde edilmiştir.

Ayrıca Gabor uygulamaları için seçilen parametreler ise, 64 farklı filtre kullanılmıştır. Gabor çekirdek boyutu (5,5) değerini, sigma parametresi bir ve üç değerlerini, theta ve lamda 0^0 , 45^0 , 90^0 , 135^0 açı değerlerini ve gamma 0.05 ve 0.25 değerlerini, faz kayması (psi) sıfır değerini yapılan uygulamalarda kullanılmıştır.

5.3.4. Özellik Seçimi

Ardından Mİ tanımada en az özellik haritası ile maksimum performansı vermesini sağlayacak olan ÇDÖÖE özellik seçim algoritması kullanılmıştır. Özellik seçim algoritması ile Mİ tanımada en verimli özellikler seçilerek, modelin başarısının artırılması hedeflenmiştir. Özellik seçimi algoritmalarından olan ÇDÖÖE algoritması kullanılıp, algoritmada ‘estimator’ parametresi DVM sınıflandırıcısı, ‘step’ parametre değeri ‘1000’, ‘scoring’ parametre değeri ‘accuracy’ olarak ayarlanmıştır.

5.3.5. Sınıflandırma

Veri çapraz doğrulama tekniği kullanılarak veri seti 5 parçaya bölünmüş, 1 parçası test verisi, kalan 4 parçası eğitim verisi olarak ayarlanmıştır (k=5). Bu işlem veri setindeki tüm örneklerin hem test verisi hemde eğitim verisi olarak kullanılmasını sağlamaktadır.

Hem Xception derin öğrenme modelinde hem de Gabor filtreleme modelinde kullanılan sınıflandırma algoritması olan DVM’de ‘kernel’ parametre değeri ‘linear’, ‘C’ parametre değeri ‘1.0’, ‘probability’ değeri ‘true’ olarak ayarlanırken, çapraz doğrulama olarak kullanılan RepeatedKfold algoritmasında ise ‘n_splits’ değeri ‘5’, ‘n_repeats’ değeri ‘5’, ‘random state’ değeri ‘42’ olarak belirlenmiştir.

BÖLÜM 6

DENEYSEL TEST SONUÇLARI VE TARTIŞMA

Önerilen modeller ve yapılan deneysel çalışmada hedeflenen temel maddeler şunlardır: özgün bir model oluşturma, Mİ örneklerinin özelliklerini elde etmek için en başarılı makine öğrenme algoritması seçimi, ayırt edici özellikleri seçmek ve modelin performansını artırmak için özellik seçim algoritmalarını değerlendirme, Mİ veri seti örnekleri için uygun olan derin öğrenme algoritmalarının performansının karşılaştırılmasıdır. Yapılan deneysel çalışmalarda, 12 GB RAM, Intel Corei7-6700HQ 2.60 GHz CPU, NVIDIA GeForce GTX 950M GPU'ya sahip bir bilgisayar, Python programlama dili ve Sklearn, Matplotlib, OpenCV, Tensorflow'a ait keras kütüphaneleri kullanılmıştır.

6.1. DTÖ-PSO MODELİ İÇİN DENEYSEL TEST SONUÇLARI

Bu tez kapsamında önerilen DTÖ-PSO modelin mimarisi, ön işleme, özellik çıkarma, özellik seçme ve sınıflandırma olmak üzere dört adımdan oluşmaktadır. Kullanılan metotların performanslarına ait deneysel test sonuçları sırasıyla verilmiştir.

Önerilen DTÖ-PSO için yapılan ilk deneysel çalışmada, seçilen üç veri setinin birleştirilmesiyle 266 negatif, 112 pozitif, 88 sürpriz örneklemeinden oluşan yeni bileşik veri setinin VGG16, SqueezeNet, MobileNetV2, EfficientNet, AlexNet ESA modellerinin tamamında yapılan ölçümlerin hiper parametreleri aynı olup aşağıdaki parametreler uygulanmıştır:

- Maksimum Tur (Max Epoch) değeri 32,
- Mini Parti (Mini Batch Size) değeri 16,
- Başlangıç Öğrenme Hızı (Initial Learn Rate) değeri 0.001,
- Doğrulama Sıklığı (Validation Frequency) değeri 50,

- Öğrenme Hızı Düşüş Faktörü (Learn Rate Drop Factor) değeri 0.1,
- Öğrenme Hızı düşüş Süresi (Learn Rate Drop Period) değeri 16 olarak ayarlanmıştır. DTÖ-PSO modeli ile yapılan deneysel test sonuçlarında elde edilen performans metrikleri Çizelge 6.1’de verilmiştir.

Çizelge 6.1. ESA modelleri kullanılarak sınıflandırılan deneysel ölçümlerin sonuçları

Model	Veri Seti	Doğruluk	Duyarlılık	Özgüllük	Kesinlik	F1
VGG16		0.7143	0.6626	0.8260	0.6820	0.6656
Squeezenet		0,6857	0,62	0,8027	0,6627	0,6370
MobilenetV2	Kompozit	0.6143	0.4437	0.7221	0.6159	0.4491
EfficientNet		0.6000	0.4613	0.7391	0.5138	0.4686
AlexNet		0.7143	0.6600	0.8150	0.7011	0.6739
VGG16		0.8083	0.8083	0.9042	0.8269	0.8114
Squeezenet	Kompozit	0.7125	0.7125	0.8562	0.7180	0.7137
MobilenetV2	+	0.5583	0.5583	0.7792	0.5568	0.5507
EfficientNet	Arttırılmış	0.6083	0.6083	0.8042	0.6142	0.6082
AlexNet		0.7667	0.7667	0.8833	0.7695	0.7651

Çizelge 6.1’e göre bileşik veri setinde VGG16 ve AlexNet en yüksek olan 0,7143 doğruluk oranları ölçülürken, duyarlılık ve özgüllük ölçümünde de en üstün performansı sırasıyla 0.6626 ve 0.8260 değerleri ile Vgg16 göstermiştir. Özgüllük ve kesinlik ölçümünde VGG16 daha üstünken, kesinlik ve F1 ölçümlerinde AlexNet sırasıyla 0,7011 ve 0,6739 değeriyle en üstün performansı göstermiştir. Arttırılmış veri setinde de aynı hiper parametreler kullanılarak elde edilen sonuçlarda VGG16 modeli hem bileşik veri setindeki kendi değerlerine göre performansını yükseltmiş hem de diğer modellere üstünlük kurmuştur. Arttırılmış veri setinde VGG16 modeli ile doğruluk 0.8083, duyarlılık 0.8083, özgüllük 0.9042, kesinlik 0.8269 ve f1 0.8114 ölçümleri elde ederek diğer modellere Mİ tanıma görevinde üstünlük sağlamıştır.

Bir sonraki deneysel çalışmada, ESA modellerinde sınıflandırma katmanından hemen önceki son tam bağlı katmanları olan; VGG16->FC8, AlexNet->FC8, SqueezeNet->Pool10, MobileNetV2->Logits, EfficientNetB0->MatMul katmanlarına ait 1000 er adet özellik vektörleri alınmıştır. ESA modellerinin son tam bağlı katmanlarından elde

edilen 1000'er özellik vektörleri DVM'nin doğrusal, quadratic, finegaussian, cubic çekirdekleri kullanılarak sınıflandırılmıştır. Elde edilen performans sonuçları Çizelge 6.2'de verilmiştir.

Çizelge 6.2. 1000 ESA özellik haritasının DVM çekirdekleriyle sınıflandırılması.

Model	Çekirdek	Doğruluk	Duyarlılık	Özgüllük	Kesinlik	F1
VGG16	Quadratic	0.6653	0.6659	0.8325	0.6619	0.6633
Squeezenet		0.7866	0.7860	0.8932	0.8011	0.7858
MobilenetV2		0.6904	0.6906	0.8452	0.6894	0.6876
EfficientNet		0.7782	0.7784	0.8890	0.7806	0.7769
AlexNet		0.6987	0.6989	0.8493	0.6988	0.6988
VGG16	Doğrusal	0.5565	0.5570	0.7781	0.5599	0.5568
Squeezenet		0.6234	0.6230	0.8116	0.6443	0.6256
MobilenetV2		0.5858	0.5861	0.7929	0.5840	0.5841
EfficientNet		0.6653	0.6652	0.8325	0.6697	0.6663
AlexNet		0.5816	0.5820	0.7907	0.5816	0.5817
VGG16	Finegaussian	0.6234	0.6224	0.8113	0.8235	0.6146
Squeezenet		0.6485	0.6499	0.8250	0.8058	0.6463
MobilenetV2		0.4268	0.4292	0.7146	0.7886	0.3360
EfficientNet		0.6360	0.6348	0.8176	0.8263	0.6257
AlexNet		0.5816	0.5806	0.7904	0.8045	0.5674
VGG16	Cubic	0.7143	0.7146	0.8571	0.7108	0.7118
Squeezenet		0.8033	0.8027	0.9016	0.8151	0.8012
MobilenetV2		0.7448	0.7449	0.8724	0.7434	0.7430
EfficientNet		0.7741	0.7742	0.8869	0.7796	0.7732
AlexNet		0.7322	0.7324	0.8662	0.7405	0.7327

Çizelge 6.2 incelendiğinde kendi içinde kullanılan DVM'nin çekirdekleri Mİ tanıma görevinde üstün oldukları gözlemlenmiştir. Quadratic çekirdeğinde, Squeezenet 0.7866 ile doğruluk değeri, 0.7860 ile duyarlılık değeri, 0.8932 ile özgüllük değeri, 0.8011 ile kesinlik ve 0.7858 ile F1 değeri ile diğer modellerden bu ölçümlerde üstündür. Doğrusal çekirdek kullanılan DVM ile sınıflandırmada, EfficientNet tüm

ölçümlerde performans olarak diğerlerinden daha başarılı sonuçlar elde edilmiştir. Doğruluk değeri 0.6653, duyarlılık değeri 0.6652, özgüllük değeri 0.8325, kesinlik değeri 0.6697 ve F1 değeri 0.6663 elde edilmiştir. Finegaussian çekirdeği kullanılarak sınıflandırılan veri setinin en yüksek doğruluk oranını Squeezenet 0,6485 değeri ile en yüksek değere sahip modeldir. Duyarlılık, özgüllük ve F1 değerlerindedir diğer modellerden üstünken, kesinlik değerinde VGG16 0.8263 ile en başarılı model olmuştur. SqueezeNet kullanılarak alınan 1000 özellik haritasının Cubic DVM çekirdeği ile sınıflandırılmasında en yüksek doğruluk, duyarlılık, özgüllük ve F1 değerini veren çekirdek olmuştur. Doğruluk 0.8033 oranı, duyarlılık 0.8027 oranı, özgüllük 0,9016 oranı, kesinlik 0,8051 oranı ve F1 0.8012 oranı elde edilmiştir.

DTÖ-PSO için yapılan bir sonraki deneysel çalışmada, VGG16, SqueezeNet, MobilNetV2, EfficientNet ve AlexNet ESA modellerinin her birinin son tam bağlı katmanlarından elde edilen 1000'er özellik vektörleri birleştirilerek oluşturulan 5000 özellik, DVM çekirdekleri ile sınıflandırılmıştır. Deneysel test sonuçlarında elde edilen performans metrikleri Çizelge 6.3'de verilmiştir.

Çizelge 6.3. ESA modellerinden elde edilen 1000'er özelliğin birleştirilip DVM çekirdekleri ile sınıflandırma performansları.

Çekirdek	Doğruluk	Duyarlılık	Özgüllük	Kesinlik	F1
Quadratic	0.7908	0.7911	0.8955	0.7946	0.7897
Doğrusal	0.7029	0.7029	0.8516	0.7101	0.7053
Finegaussian	0.7280	0.7278	0.8637	0.8506	0.7306
Cubic	0.8117	0.8120	0.9060	0.8159	0.8114

Çizelge 6.3'de finegaussian çekirdeği ile elde edilen kesinlik değeri 0.8506 ile en başarılı iken, diğer değerlerin tamamında cubic çekirdeği en iyi neticeleri vermiştir. Cubic çekirdeği ile yapılan sınıflandırmada doğruluk oranı 0.8117, duyarlılık oranı 0.8120, özgüllük oranı 0.906 ve f1 oranı 0.8114 elde edilmiştir. Bu sonuçlar neticesinde, bir önceki deneysel çalışma sonuçlarına göre (Çizelge 6.2) az da olsa bir iyileşme olduğu gözlenmektedir.

Bir sonraki deneysel çalışmada, bir önceki deneysel çalışmada elde edilen 5000 özellik üzerinde PSO algoritması kullanılarak Mİ'yi tanımaya yönelik daha etkin 2760 adet ayırt edici özellik seçilmiştir. Seçilen bu ayırt edici özellikler, farklı DVM çekirdekleri ile sınıflandırma işlemi deneysel test sonuçlarından elde edilen performans metrikleri Çizelge 6.4'de verilmiştir.

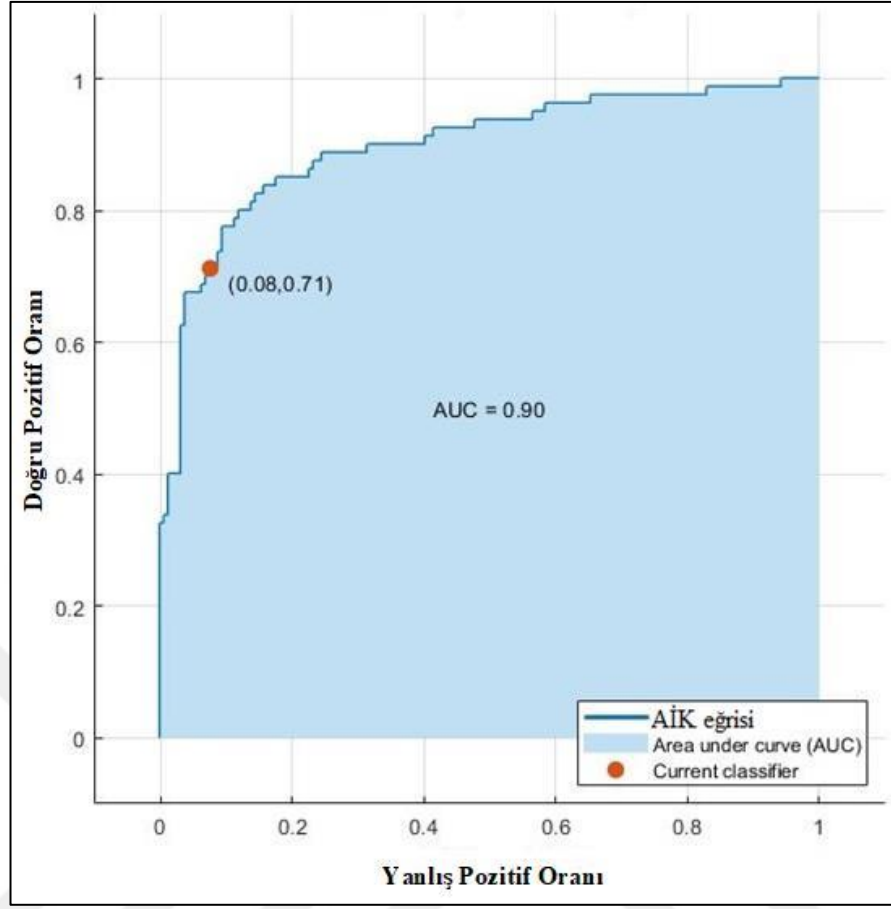
Çizelge 6.4. PSO analizinden sonra DVM çekirdekleri ile sınıflandırma sonuçları.

Çekirdek	Doğruluk	Duyarlılık	Özgüllük	Kesinlik	F1
Quadratictric	0.8243	0.8245	0.9121	0.8275	0.8232
Doğrusal	0.7322	0.7324	0.8660	0.7445	0.7321
Finegaussian	0.6820	0.6813	0.8407	0.8376	0.6833
Cubic	0.8159	0.8162	0.9079	0.8202	0.8140

Buna göre en iyi sonucu veren quadratic çekirdeğinden, doğruluk değeri 0.8243, duyarlılık değeri 0.8245, özgüllük değeri 0.9121 ve F1 değeri 0.8232 elde edilmiştir. En başarılı kesinlik değeri sonucu ise finegaussian çekirdeğinden alındığı gözlemlenmiştir. Çizelge 6.3 ile kıyaslandığında kesinlik ölçümü hariç diğer ölçümlerin performansında iyileşme gözlenmiştir. Quadratic çekirdeğine ait karmaşıklık matrisi ve AİK eğrisi grafiği Şekil 6.1 ve 6.2'de verilmiştir.

Sürpriz	Negatif	Pozitif	Sürpriz
	57	15	8
	8	71	1
Sürpriz	4	6	69
	Negatif	Pozitif	Sürpriz

Şekil 6.1. Çapraz doğrulama olmadan elde edilen karmaşıklık matrisi ve auc roc grafiği.



Şekil 6.2. Çapraz doğrulama olmadan elde edilen karmaşıklık matrisi ve AİK eğrisi grafiği.

DTÖ-PSO modeli için yapılan son deneysel çalışmamızda veri çapraz doğrulama tekniği kullanılarak veri seti 10 parçaya bölünmüş, bir parçası test verisi, kalan dokuz parçası eğitim verisi olarak ayarlanmıştır ($k=10$). Bu işlem sonunda sırasıyla her bir parçanın test diğer kalan dokuz parçanın ise eğitim verisi olarak kullanılmasına olanak sağlar. Böylece veri seti üzerindeki her örneklem hem veri seti hem de eğitim seti olarak görev yapar. Buda daha verimli bir sonuç elde edilmesini sağlamaktadır [150]. Çizelge 6.5’de güçlü özellik seçimi sonrası elde edilen 2760 özellik verisinin, çapraz doğrulama işlemi ile DVM çekirdeklerin sınıflandırma performanslar listelenmiştir.

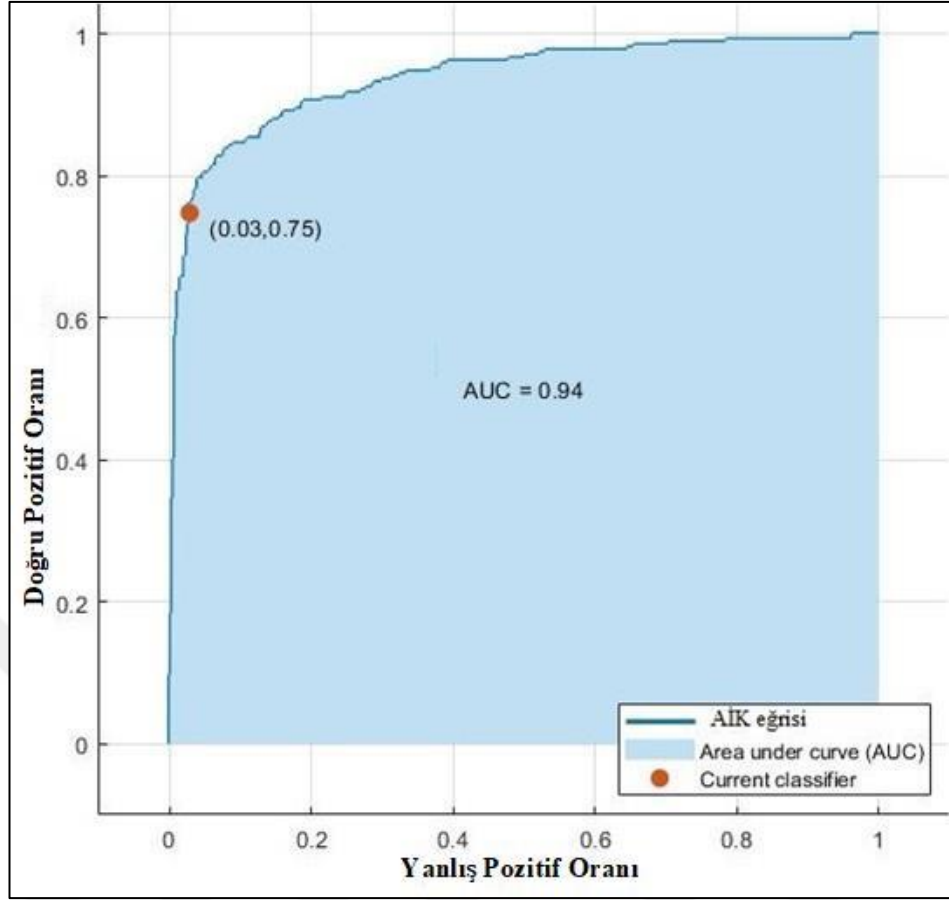
Çizelge 6.5. PSO analizi ve çapraz doğrulama sonrası DVM çekirdekleri ile sınıflandırma sonuçları

Çekirdek	Doğruluk	Duyarlılık	Özgüllük	Kesinlik	F1
Quadratic	0.8484	0.8484	0.9242	0.8508	0.8469
Doğrusal	0.7306	0.7306	0.8653	0.7323	0.7299
Finegaussian	0.6729	0.6729	0.8365	0.7526	0.6712
Cubic	0.8784	0.8784	0.9392	0.8839	0.8765

Buna göre Cubic çekirdeği en başarılı performansa sahip olmanın yanı sıra deneyler içinde en yüksek performans değerleri elde edilmiştir. Elde edilen sonuçlar sırasıyla doğruluk 0.8784, duyarlılık 0,8784, özgüllük 0.9392, kesinlik 0.8839 ve F1 ölçümü 0.8765’dir. Deneyler içinde en başarılı performansı veren ve önerilen modelimiz için seçtiğimiz cubic çekirdeğine ait karmaşıklık matrisi ve AİK eğrisi grafiği Şekil 6.3 ve 6.4’de verilmiştir.

Sürpriz	Negatif	Pozitif	Sürpriz
	199	48	19
	10	247	9
Pozitif	5	6	255
Negatif			

Şekil 6.3. Çapraz doğrulamalı modelin karmaşıklık matrisi.



Şekil 6.4. Çapraz doğrulamalı modelin AİK eğrisi grafiği.

Tez kapsamında önerilen DTÖ-PSO model çalışması değerlendirildiğinde, birleştirilmiş veri seti üzerinde ESA modelleri kullanılarak yapılan sınıflandırmada doğruluk oranlarına ait F1 skorların %72 oranını geçemediği görülmüştür. Performansı artırmak için veri arttırma tekniği ile veri setindeki örnekler yapay olarak çoğaltılıp sınıflandırıldığında başarı oranlarının arttığı gözlemlenmiştir. Artan oranlar incelendiğinde; %80 leri geçen doğruluk oranı, %81 geçen F1 ölçüm oranı elde edilirken, ayrıca özgüllük ölçümlerde de %90 ları bulan değerler elde edildiği Çizelge 6.1’de sunulmuştur. Buradan veri arttırma tekniğinin Mİ tanıma performansını arttırmada, birçok ESA modelinde olumlu etkiler yapacağı sonucuna varılabilir. ESA modellerinin son bağlı katmanlarından alınan 1000 özellik haritaları ile yapılan Çizelge 6.2’deki dört farklı DVM çekirdeğin kullanıldığı sınıflandırmada, genel olarak Çizelge 6.1 temel alındığında, deney sonuçlarının performansın çok üstüne

çıkamamıştır. Çizelge 6.3’de her bir ESA modelinin son bağlı katmanından elde edilen 1000’er özelliğin birleştirilmesi ile elde edilen 5000 özellik ile yapılan sınıflandırma sonuçları ise performansın kısmen üstüne çıkmıştır. Bunun nedeni farklı sınıflandırma tekniklerinden fazlasına ihtiyaç duyulması sonucuna varılabilir. Nitekim bir sonraki aşamada en iyi özelliklerin seçimi için PSO algoritması kullanılarak yapılan deney ile quadratic ve cubic çekirdeklerinden elde edilen sonuçlar Çizelge 6.4’da gözle görülür bir artışa sebep olmuştur. En yüksek değerlere ulaşılan doğruluk ve F1 ölçülerinde sırasıyla 0.8243 ve 0.8232 değerleri elde edilmiştir. Böylece özellik seçimi yapılarak elde edilen özelliklerin, Mİ tanıma performansını arttırmada olumlu etkisi olduğu gözlemlenmiştir. Ayrıca PSO algoritması kullanılarak en iyi özellik seçimi, çapraz doğrulama tekniği ve cubic DVM ile yapılan sınıflandırma deneyinin tüm metrik değerleri için Mİ tanıma performansında önemli gelişmeler olduğu Çizelge 6.5’de gösterilmiştir. Daha somut olarak değerlendirildiğinde, Çizelge 6.1’deki kompozit veri setindeki en iyi ölçüm değerleriyle kıyaslandığında, performans artışı doğruluk, duyarlılık, özgüllük, kesinlik ve F1 ölçümleri sırasıyla %16.41, %21.58, %11.32, %18.28 ve %20.26 olarak ölçülmüştür. Deneylerimizde elde ettiğimiz en iyi sonuçlar Çizelge 6.6 da gösterilmiştir.

Çizelge 6.6. Çapraz doğrulama olmadan ve çapraz doğrulama ile önerilen modelin en iyi performansları

Çekirdek	Doğruluk	Duyarlılık	Özüllük	Kesinlik	F1
Quadratic/çapraz doğrulamasız	0.8243	0.8245	0.9121	0.8275	0.8232
Cubic/çapraz doğrulamalı	0.8784	0.8784	0.9392	0.8839	0.8765

S.-T. Liong vd.’nin [183] yapmış oldukları araştırmadaki Tablo 3’de farklı optik akış yöntemlerinin Mİ tanımadaki başarısı sunulmuştur. Farneback yöntemi kullanılarak üretilen tanıma sonuçları TV-L1 ile birlikte farklı blok boyutlarında birbirlerine üstünlük sağlayan en iyi iki yöntem olarak verilmiştir. Bunun yanı sıra yüz hareketlerinin analizinin hesaplanmasında diğerlerine oranla daha hızlı olan Faneback yöntemi için bir diğer avantaj olarak olarak gösterilmiştir [80]. Modifiye edilmiş PSO

türevleri ve farklı DVM çekirdeklerinin Mİ tanıma doğruluğunu arttırmada daha olumlu sonuçlar verebileceği göz önünde bulundurulmalıdır.

6.2. ÖNERİLEN DENSENET169-ÖÖE MODELİ İÇİN PERFORMANS SONUÇLARI

Bu tez kapsamında önerilen DensNet169-ÖÖE modelin mimarisi, önışleme, özellik çıkarma, özellik seçme ve sınıflandırma olmak üzere dört adımdan oluşmaktadır. Kullanılan metotların performanslarına ait deneysel test sonuçları sırasıyla verilmiştir.

Önerilen bu model için yapılan ilk deneysel çalışmada, seçilen üç veri setinin birleştirilmesiyle 252 negatif, 109 pozitif, 83 sürpriz örneklemeinden oluşan yeni bileşik veri setinin, RGB görüntülerinin tepe kareleri ve iki farklı optik akış özellikleri DenseNet121, DenseNet169, DenseNet201 modellerine girdi olarak verilmiştir. Buna göre iki boyutlu görüntü verileri DenseNet121'den 50176 adet özellik, DenseNet169'den 81536 adet özellik, DenseNet201'den 94080 adet özellik kullanılarak sınıflandırılmıştır. Bu üç ESA modellerinden alınan özelliklerin doğrusal DVM ile sınıflandırma performanslarının sonuçları Çizelge 6.7'de verilmiştir.

Çizelge 6.7. Bütünleşik veri setinin DenseNet modelleri ile sınıflandırılmasının sonuçları.

Model	Özellik Çıkarma	Doğruluk	Duyarlılık	Özgüllük	Kesinlik	F1
DenseNet121	RGB tepe	0.6054	0.5212	0.7606	0.5310	0.5253
DenseNet169	kare	0.6320	0.5536	0.7768	0.5653	0.5585
DenseNet201	görüntüleri	0.6329	0.5573	0.7786	0.5664	0.5612
DenseNet121		0.6248	0.5500	0.7750	0.5898	0.5648
DenseNet169	Farneback	0.6239	0.5359	0.7679	0.5826	0.5515
DenseNet201		0.6396	0.5497	0.7748	0.6044	0.5676
DenseNet121		0.6653	0.5926	0.7963	0.6335	0.6083
DenseNet169	TVL1_Dual	0.6838	0.6184	0.8092	0.6615	0.6353
DenseNet201		0.6806	0.6267	0.8133	0.6504	0.6369

DenseNet121		0.5847	0.4946	0.7473	0.5219	0.5037
DenseNet169	TVL1_yatay	0.6185	0.5301	0.7651	0.5722	0.5441
DenseNet201		0.6176	0.5307	0.7654	0.5614	0.5407
DenseNet121		0.6829	0.6266	0.8133	0.6576	0.6396
DenseNet169	TVL1_dikey	0.6955	0.6495	0.8248	0.6753	0.6609
DenseNet201		0.6721	0.6183	0.8092	0.6429	0.6289

Çizelge 6.7’de verilen bütünleşik veri setinde TV-L1_yatay görüntülerinden elde edilen özelliklerin sınıflandırma ölçümleri en zayıf performans gösterirken, TV-L1_dikey görüntülerinden elde edilen özelliklerin sınıflandırma ölçümleri üç model içinde en yüksek performansı göstermiştir. Buna göre DenseNet169 modeli, doğruluk 0.6955, duyarlılık değeri 0.6495 özgüllük değeri 0.8248, kesinlik değeri 0.6753 ve f1 değeri 0.6609 ile en başarılı performansı göstermiştir.

Bu model için yapılan bir sonraki deneysel çalışmada yeni bileşik veri seti, veri artırma tekniği ile her sınıf en büyük örnelemeye sahip negatif sınıfa eşitlenerek, her sınıf 252 adet örnelemeye sahip olmuştur. Böylece veri seti dengeli hale getirilip, sınıflandırıcının daha iyi sınıflandırması beklenmektedir [21]. Artırılmış bileşik veri seti, RGB tepe kare ve farklı optik akış görüntülerinin özellikleri DenseNet121, DenseNet169, DenseNet201 modellerine girdi olarak verilmiştir. Buna göre iki boyutlu görüntü verileri DenseNet121’den 50176 adet özellik, DenseNet169’den 81536 adet özellik, DenseNet201’den 94080 adet özellik kullanılarak sınıflandırılmıştır. Bu üç ESA modellerinden alınan özelliklerin doğrusal DVM sınıflandırma performanslarının sonuçları Çizelge 6.8’de verilmiştir.

Çizelge 6.8. Arttırılmış bütünleşik veri setinin DenseNet modelleri ve Doğrusal DVM ile sınıflandırma sonuçları

Model	Özellik Çıkarma	Doğruluk	Duyarlılık	Özgüllük	Kesinlik	F1
DenseNet121	RGB tepe	0.6601	0.6601	0.8300	0.6582	0.6583
DenseNet169	kare	0.6685	0.6685	0.8343	0.6664	0.6669
DenseNet201	görüntüleri	0.6574	0.6574	0.8287	0.6561	0.6565

DenseNet121		0.6685	0.6685	0.8343	0.6700	0.6685
DenseNet169	Farneback	0.6563	0.6563	0.8282	0.6560	0.6555
DenseNet201		0.6780	0.6780	0.8390	0.6788	0.6778
DenseNet121		0.7108	0.7108	0.8554	0.7113	0.7105
DenseNet169	TVL1_Dual	0.7183	0.7183	0.8591	0.7185	0.7180
DenseNet201		0.6992	0.6992	0.8496	0.6996	0.6988
DenseNet121		0.6230	0.6230	0.8115	0.6222	0.6217
DenseNet169	TVL1_yatay	0.6376	0.6376	0.8188	0.6367	0.6364
DenseNet201		0.6143	0.6143	0.8071	0.6131	0.6132
DenseNet121		0.7061	0.7061	0.8530	0.7059	0.7053
DenseNet169	TVL1_dikey	0.7294	0.7294	0.8647	0.7293	0.7290
DenseNet201		0.7048	0.7048	0.8524	0.7046	0.7046

Çizelge 6.8’ daki arttırılmış bileşik veri setinde TV-L1 yatay görüntülerinden elde edilen özelliklerin sınıflandırma ölçümleri en zayıf performans gösterirken, TV-L1 dikey görüntülerinden elde edilen özelliklerin sınıflandırma ölçümleri üç model içinde en başarılı performansı göstermiştir. Buna göre DenseNet169 modeli, doğruluk 0.7294, duyarlılık değeri 0.7294, özgüllük değeri 0.8647, kesinlik değeri 0.7293 ve f1 değeri 0.7290 ile en yüksek performansı göstermiştir. Veri arttırma tekniği ve dengeli örneklemenin sonrasında bir önceki deneysel çalışmadaki ölçümlerden (Çizelge 6.7) elde edilen en başarılı performanslar karşılaştırıldığında %3 den daha fazla bir performans artışına, bazı modellerde ise %5’den fazla performans artışına etkisi olduğu gözlemlenmiştir.

Bir sonraki deneysel çalışmada ise DenseNet ve DVM yöntemlerinden farklı olarak ÖÖE özellik seçim algortimasından yararlanılarak Mİ tanımda en verimli özellikler seçilmiş, modelin performansının arttırılması hedeflenmiştir. Özellik seçimi algoritmalarından olan ÖÖE algortimasından yararlanılarak, algortimada ‘n_features_to_select’ parametre değeri ‘6200’, ‘estimator’ parametre değeri DVM sınıflandırıcısı, ‘step’ parametre değeri ‘1000’ olarak belirlenmiştir. Elde edilen bu 6200 özellik doğrusal DVM sınıflandırıcı ile sınıflandırılmıştır. Parametre değerleri sırasıyla ‘kernel=linear’ ve ‘C=1.0’ olarak atanmıştır. Çizelge 6.9’da bütünleşik veri

setinde TV-L1_Dual optik akış görüntüleri ile birlikte kullanılan DenseNet169 ESA modeli en iyi performansı sağlamaktadır.

Çizelge 6.9. Arttırılmış bütünleşik veri setinin DenseNet modelleri doğrusal DVM-ÖÖE ile sınıflandırma sonuçları

Model	Özellik Çıkarma	ÖÖE				
		Doğruluk	Duyarlılık	Özgüllük	Kesinlik	F1
DenseNet121	RGB tepe	0.7632	0.7632	0.8816	0.7632	0.7628
DenseNet169	kare	0.7955	0.7955	0.8978	0.7957	0.7950
DenseNet201	görüntüleri	0.7997	0.7997	0.8999	0.7996	0.7993
DenseNet121	Farneback	0.8402	0.8402	0.9201	0.8410	0.8401
DenseNet169		0.8611	0.8611	0.9306	0.8632	0.8609
DenseNet201		0.8669	0.8669	0.9335	0.8676	0.8669
DenseNet121	TVL1_Dual	0.8524	0.8524	0.9262	0.8529	0.8523
DenseNet169		0.8704	0.8704	0.9352	0.8708	0.8705
DenseNet201		0.8632	0.8632	0.9316	0.8638	0.8631
DenseNet121	TVL1_yatay	0.7918	0.7918	0.8959	0.7930	0.7914
DenseNet169		0.8000	0.8000	0.900	0.8007	0.7994
DenseNet201		0.8257	0.8257	0.9128	0.8255	0.8254
DenseNet121	TVL1_dikey	0.8556	0.8556	0.9278	0.8567	0.8553
DenseNet169		0.8661	0.8661	0.9331	0.8667	0.8661
DenseNet201		0.8587	0.8587	0.9294	0.8596	0.8585

DenseNet169 modelinden elde edilen en iyi performans ölçümleri doğruluk 0.8704, duyarlılık 0.8704, özgüllük 0.9352, kesinlik 0.8708 ve F1 0.8705 olarak ölçülmüştür. Böylece başlangıta yapılan deneysel çalışmadaki (Çizelge 6.7) en başarılı metrikler referans alındığında 0.1749 değerinde doğruluk, 0.2209 değerinde duyarlılık, 0.1104değerinde özgüllük, 0.1945 değerinde kesinlik ve 0.2096 değerinde F1 performans artışı gözlemlenmektedir.

RGB tepe kare görüntüleri, FarneBack görüntüler, TVL1_Dual görüntülerinden çıkarılan ve birleştirilen özelliklere ait deneysel test sonuçları Çizelge 6.10'de

verilmiştir. Bu deneysel çalışmada yapılan aynı görüntülere ait RGB tepe kare görüntülerinin, Farneback ve TV-L1_Dual optik akış görüntülerinin özellik haritaları birleştiren üç akışlı modelde daha zengin ve daha farklı bilgilere sahip olan özellik haritaları kullanılarak modelimizin performansı arttırılmaya çalışılmıştır. Bunun sebebi bir görüntüden elde edilemeyen fakat diğer görüntüden elde edilebilen farklı Mİ özellikleri içeren görüntülerin özellik haritalarının birleştirilmesiyle derin öğrenme modellerinin performansına olumlu etkilerinin olması hedeflenmiştir. Bu birleştirmeden sonra DenseNet121 için 50176x3 özellik, DenseNet169 için 81536x3 özellik, DenseNet201 için 94080x3 özellik elde edilmiştir. Bu birleşme sonucunda Mİ tanıma için gerekli özellik haritaları arttığı kadar, faydalı olmayan özellik haritalarında arttığı için, arttırılmış bütünleşik veri setine göre performansının azaldığı Çizelge 6.10’de gözlemlenmiştir.

Çizelge 6.10. Zenginleştirilmiş özellik haritalarının DenseNet modelleri ve doğrusal DVM ile sınıflandırma sonuçları

Model	Doğruluk	Duyarlılık	Özgüllük	Kesinlik	F1
DenseNet121	0.6101	0.6101	0.8050	0.6095	0.6093
DenseNet169	0.6429	0.6429	0.8214	0.6417	0.6420
DenseNet201	0.6325	0.6325	0.8163	0.6325	0.6322

RGB tepe kare görüntüleri, Farneback ve TV-L1 optik akış görüntülerinin birleşiminden oluşan zengin özellik haritaları ÖÖE algoritması ile en verimli 6200 özellik seçilerek sınıflandırıldığında deneysel çalışmalar içerisinde performansı en başarılı sonuçlar elde edilmiştir. Arttırılmış bileşik veri setinden elde edilen farklı özellik haritalarının birleştirilmesi ile DenseNet ESA modelleri ve Doğrusal DVM-ÖÖE sınıflandırma sonuçları Çizelge 6.11’de verilmiştir.

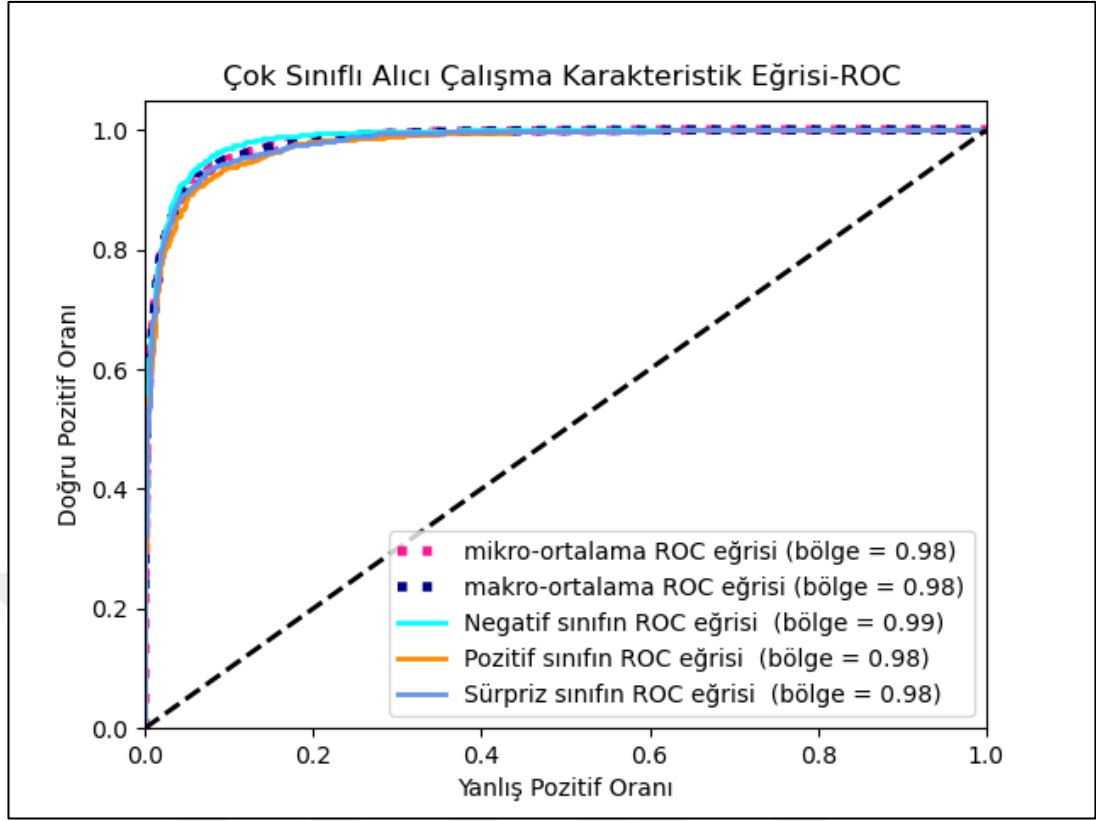
Çizelge 6.11. Zenginleştirilmiş özellik haritalarının DenseNet modelleri doğrusal DVM-ÖÖE ile sınıflandırma sonuçları

Model	Özellik çıkarma ve birleştirme	ÖÖE				
		Doğruluk	Duyarlılık	Özgüllük	Kesinlik	F1
DenseNet121	(RGB tepe kare görüntüleri	0.8849	0.8849	0.9425	0.8894	0.8851
DenseNet169	+ FarneBack	0.90	0.90	0.95	0.9007	0.8999
DenseNet201	+ TVL1_Dual)	0.8897	0.8897	0.9448	0.8944	0.8894

Buna göre en başarılı performansı DenseNet169 modeli seçilen 6200 özellik haritası ile 0.90 doğruluk ve duyarlılık, 0.95 özgüllük, 0.9007 kesinlik ve 0.8999 F1 ölçüm oranları ile göstermektedir. Bu başarı oranı, bu kapsamda yapılan deneysel çalışmalar içerisinde ulaşılan en yüksek ve en başarılı performans olmuştur. Bu modele ait karmaşıklık matrisi Şekil 6.5’de, AİK eğrisinin grafiği 6.6’de verilmiştir.

Sürpriz	Negatif	94	4	2
	Pozitif	6	89	5
	Sürpriz	6	6	87
		Negatif	Pozitif	Sürpriz

Şekil 6.5. Modelin karmaşıklık matris tablosu.



Şekil 6.6. Modelin AİK eğrisi grafiği.

Son olarak birleştirilen özellik haritalarının ikili kombinasyonları da incelenerek sonuçları Çizelge 6.12’de sunulmaktadır.

Çizelge 6.12. Zenginleştirilmiş ikili özellik haritalarının DenseNet modelleri doğrusal DVM-ÖÖE ile sınıflandırma sonuçları

Model	Özellik çıkarma ve birleştirme	ÖÖE				
		Doğruluk	Duyarlılık	Özgüllük	Kesinlik	F1
DenseNet169	RGB tepe kare görüntüleri +TVL1_Dual	0.8386	0.8386	0.9193	0.8424	0.8379
DenseNet169	FarneBack + TVL1_Dual	0.8849	0.8849	0.9425	0.8853	0.8849

	RGB tepe kare					
DenseNet169	görüntüleri +	0.8881	0.8881	0.9440	0.8890	0.8880
	FarneBack					

Çizelge 6.11'nin, Çizelge 6.12'den tek farkı daha fazla görüntüden elde edilen özellik haritası bulundurmasıdır. Bundan dolayı daha iyi sonuçlar elde etmemizi sağlamıştır. Daha zengin ve daha farklı bilgilere sahip olan özellik haritaları ile daha performanslı bir model elde etme hedefini elde eden bir deneysel ölçüm olmuştur.

Tez kapsamında önerilen DenseNet169-ÖÖE model çalışması değerlendirildiğinde, birleştirilmiş veri seti üzerinde DenseNet ESA modelleri kullanılarak yapılan sınıflandırmada doğruluk oranlarının %70 oranını geçemediği görülmüştür. Başarıyı artırmak için veri arttırma tekniği ile veri seti çoğaltılıp veri seti dengelendiğinde yapılan sınıflandırmada başarı oranlarının arttığı gözlemlenmiştir. Veri arttırma sonrası artan oranlar incelendiğinde; en başarılı oranlar (Çizelge 6.7 ve Çizelge 6.8) kıyaslandığında %3 leri geçen doğruluk oranı, bazı modellerde ise %5'den fazla artan doğruluk oranı değerleri elde edildiği görülmektedir. Buradan veri arttırma tekniğinin ve dengeli örneklemeye sahip sınıflandırmanın Mİ tanıma performansını arttırmada, birçok ESA modelinde olumlu etkiler yapabileceği sonucuna varılabilir. Özellik seçim metodu ile etkisi daha fazla olan özelliklerin filitrelenerek kullanılmasıyla yapılan Mİ sınıflandırmalarda, performans ölçümlerinde önemli başarı artışları gözlemlenmiştir. ÖÖE özellik seçimi tekniği kullanılarak yapılan ölçümlerin başlangıç performansı ile karşılaştırdığında (Çizelge 6.7 ve Çizelge 6.9) ise TV-L1_Dual görüntülerinde %18.66 oranında doğruluk, %23.52 oranında F1 performans artışı gözlemlenmiştir. Ayrıca aynı görüntüye ait farklı özelliklerin birleştirilmesi özellik seçimi yapılmadığında doğruluk oranlarını ve modelin başarı performansını düşürsede, özellik seçim algoritması kullanıldığında, tek tip görüntü olarak en iyi performansı veren TV-L1 Dual görüntüsünden alınan performansın üstüne çıktığı gözlemlenmiştir. Buda özellik farklılıklarının ve zenginliğinin Mİ tanımaya katkısını göstermektedir.

Modelimizde kullanılan DenseNet ESA modelinin avantajlarından biriside diğer son model teknolojilerle kıyaslandığında daha az parametre ve hesaplama maliyeti

gerektirmenin yanı sıra parametre artışı ile doğru orantılı bir başarı göstermesidir [134]. Ayrıca [184]'daki araştırmada farklı optik akış yöntemlerinin karşılaştırılmasında farklı blok boyutlarında birbirlerine üstünlük sağlayan en iyi iki yöntem olarak Farneback ve TV-L1 olarak verilmiştir. Çalışmamızda, TV-L1 yatay ve dikey görüntüleri kullanarak kendi aralarında karşılaştırılmış, hem de Mİ tanımada dikey görüntülerin yatay görüntülere göre daha iyi performans verdiği gözlemlenmiştir. Modelimizde ise yatay ve dikey TV-L1 görüntüleri yerine TV-L1 Dual görüntüleri işleme girmesinin sebebi hem deneylerde daha iyi sonuç vermesi hemde yatay ve dikey TV-L1 görüntüleri içerisinde barındırmasıdır. Modelimizde gözlemlenen dezavantajlardan birisi ise ÖÖE algoritması özyineleyici bir algoritma olduğu için görüntü verisi arttıkça özellik haritası seçime süresi de o kadar uzamaktadır. Farklı özellik seçim algoritmaları kullanılarak modelin maliyetinde ve başarısında artış sağlanılabilir.

6.3. ÖNERİLEN XNET-ÇDÖÖE MODELİ İÇİN PERFORMANS SONUÇLARI

Bu tez kapsamında önerilen Xnet-ÇDÖÖE modelin mimarisi, önışleme, özellik çıkarma, özellik seçme ve sınıflandırma olmak üzere dört adımdan oluşmaktadır. Kullanılan metotların performanslarına ait deneysel test sonuçları sırasıyla verilmiştir.

Kamuya açık CASME-II Mİ veri setinde ön işlem olarak optik akış özelliklerinin çıkarılması süreci gerçekleştirildikten sonra geleneksel yöntemlerden gabor modeli ve veriye dayalı yöntemlerden Xception derin öğrenme modeli ile özellik çıkarma, ÇDÖÖE algoritması ile özellik seçme ve doğrusal DVM ile sınıflandırma süreçleri uygulanarak deneysel test sonuçları elde edilmiştir.

Önerilen model Xnet-ÇDÖÖE için yapılan ilk deneysel çalışmada, seçilen veri setinden 94 negatif, 32 pozitif, 28 sürpriz video kare dizisi örneklerinden elde edilen doğal RGB görüntülerinin tepe kareleri, Farneback ve TV-L1 optik akış görüntü kareleri Xception ve Gabor modellerine girdi olarak verilmiştir. Buna göre 2B görüntü verileri Xception'dan 100352 adet özellik haritası ve Gabor'dan 50176 adet özellik

haritası çıkarılarak sınıflandırılmıştır. Bu iki modelden alınan özellik haritalarının doğrusal DVM sınıflandırma algoritması ile performanslarının sonuçları Çizelge 6.13’de verilmiştir.

Çizelge 6.13. Modellerin arttırılmamış CASME-II veri setine uygulanması ve doğrusal DVM ile sınıflandırma sonuçları

Model	Özellik Çıkarma	Doğruluk	Duyarlılık	Özgüllük	Kesinlik	F1
Xception	RGB tepe	0.6766	0.5818	0.7909	0.6138	0.5944
Gabor	karesi	0.6922	0.6016	0.8008	0.6166	0.6075
Xception	FarneBack	0.6325	0.4878	0.7439	0.5869	0.5112
Gabor		0.6013	0.5056	0.7528	0.5188	0.5112
Xception	TVL1_Dual	0.6597	0.5543	0.7772	0.6319	0.5795
Gabor		0.5870	0.4922	0.7461	0.5101	0.4992
Xception	TVL1_yatay	0.6442	0.5363	0.7682	0.5576	0.5447
Gabor		0.7870	0.7403	0.8701	0.7318	0.7354
Xception	TVL1_dikey	0.7805	0.7246	0.8623	0.7670	0.7430
Gabor		0.8260	0.7919	0.8960	0.8003	0.7952

Çizelge 6.13’ deki CASME-II veri setinde TV-L1 dikey görüntülerinden elde edilen özellik haritalarının sınıflandırma sonuçları iki model içinde en başarılı performansı göstermiştir. Buna göre Gabor modeli ile doğruluk 0.8260, duyarlılık değeri 0.7919 özgüllük değeri 0.8960, kesinlik değeri 0.8003 ve f1 değeri 0,7952 ile en yüksek performans sonuçları elde edilmiştir.

Bir sonraki deneysel çalışmada veri seti, veri arttırma tekniği ile her sınıf en fazla örnekleme sahip negatif sınıf sayısına eşitlenerek, her sınıf 94 adet görüntü örneğine sahip olmuştur. Böylece veri seti dengeli hale getirilip, DVM sınıflandırıcının daha iyi sınıflandırması beklenmektedir [21]. Arttırılmış veri setindeki doğal RGB görüntüleri ve farklı optik akış görüntüleri Xception derin öğrenme modeline ve geleneksel Gabor

modeline girdi olarak verilmiştir. Buna göre 100352 özellik ESA'dan ve Gabor modellerinden alınan 50176 özellik haritalarının doğrusal DVM sınıflandırma ile karşılaştırılmalı performanslarının sonuçları Çizelge 6.14' da verilmiştir.

Çizelge 6.14. Modellerin dengeli arttırılmış CASME-II veri setine uygulanması ve doğrusal DVM ile sınıflandırma sonuçları.

Model	Özellik Çıkarma	Doğruluk	Duyarlılık	Özgüllük	Kesinlik	F1
Xception	RGB tepe	0.7504	0.7504	0.8752	0.7501	0.7492
Gabor	karesi	0.7447	0.7447	0.8723	0.7430	0.7428
Xception	FarneBack	0.6723	0.6723	0.8362	0.6722	0.6722
Gabor		0.6191	0.6191	0.8096	0.6167	0.6173
Xception	TVL1	0.6979	0.6979	0.8489	0.6983	0.6980
Gabor		0.5837	0.5837	0.7918	0.5879	0.5822
Xception	TVL1_yatay	0.6709	0.6709	0.8355	0.6679	0.6667
Gabor		0.7099	0.7099	0.8550	0.7092	0.7061
Xception	TVL1_dikey	0.7702	0.7702	0.8851	0.7707	0.7700
Gabor		0.8277	0.8277	0.9138	0.8279	0.8272

Çizelge 6.14' daki artırma sonucu elde edilen dengeli veri setinde, TV-L1 dikey görüntülerinden elde edilen özellik haritalarının sınıflandırma sonuçları iki model içinde en yüksek performansı göstermiştir. Buna göre geleneksel yöntem Gabor modeli, doğruluk değeri 0.8277, duyarlılık değeri 0.8277 özgüllük değeri 0.9138, kesinlik değeri 0.8279 ve f1 değeri 0.8272 ile en başarılı performansı göstermiştir. Veri artırma tekniği kullanılarak veri setindeki dengeli örnekleme sonrasında ilk deneysel çalışmada (Çizelge 6.13) elde edilen en başarılı performanslar karşılaştırıldığında duyarlılık ve F1 ölçümlerinde %3 den daha fazla bir performans artışına, bazı modellerde ise %16'den fazla F1 artışı, %18'den fazla duyarlılık artışı ölçülerek modelin başarısına olumlu etkisi olduğu gözlemlenmiştir.

Önerilen model Xnet-ÇDÖÖE için uygulanan son deneysel çalışmada, özellik seçim algoritması ile Mİ tanımada en verimli özellikler seçilerek, modelin başarısının artırılması hedeflenmiştir. Özellik seçimi algoritmalarından olan ÇDÖÖE algoritması kullanılıp, algortmada ‘estimator’ parametresi DVM sınıflandırıcısı, ‘step’ parametre değeri ‘1000’, ‘scoring’ parametre değeri ‘accuracy’, ‘verbose’ parametre değeri ise ‘1’ olarak ayarlanmıştır. Elde edilen bu özellik haritaları doğrusal DVM sınıflandırıcı ile sınıflandırılmıştır. ÇDÖÖE algoritması ile özellik seçimi yapılmasından sonra ESA’dan 10256 özellik ve Gabor modellerinden alınan 27176 özellik haritalarının doğrusal DVM ile sınıflandırma performanslarının karşılaştırmalı sonuçları Çizelge 6.15’ da verilmiştir.

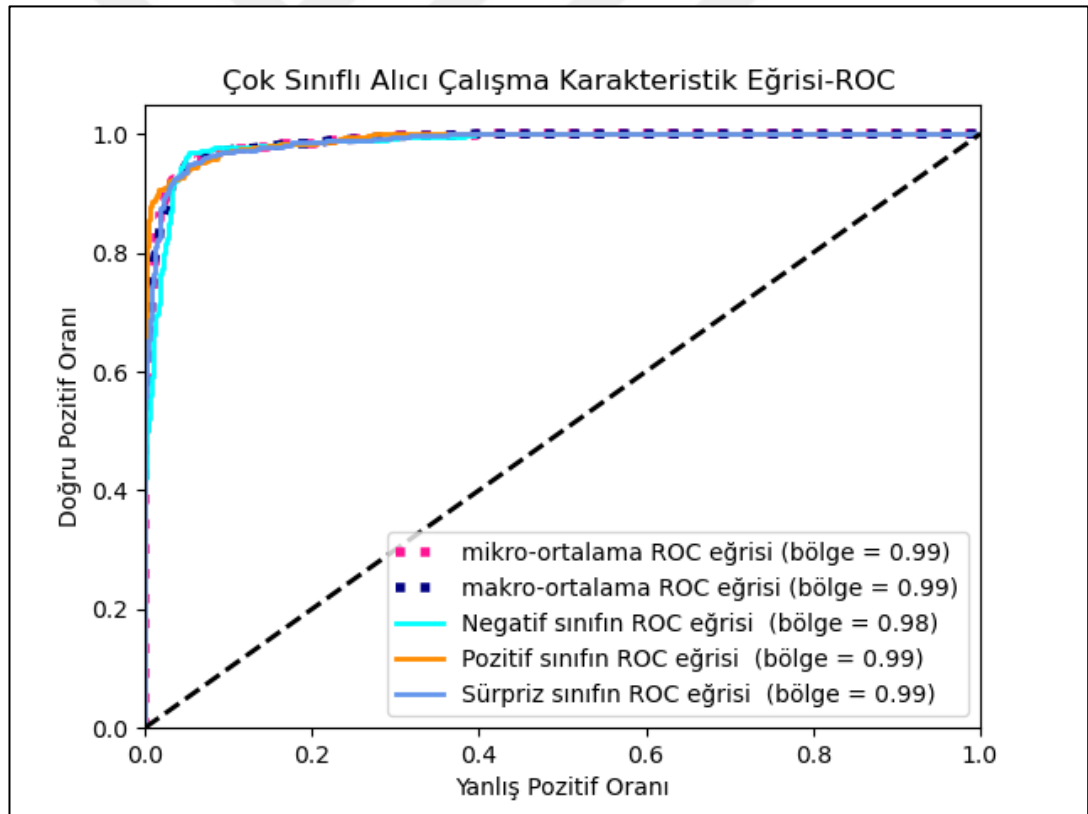
Çizelge 6.15. Filtrelenmiş özellik haritalarının Xception ve Gabor modelleri ve doğrusal DVM ile sınıflandırma sonuçları.

Model	Özellik Çıkarma	Özellik Seçimi	Doğruluk	Duyarlılık	Özgüllük	Kesinlik	F1
Xception	TVL1_	ÇDÖÖE	0.9248	0.9248	0.9624	0.9251	0.9249
Gabor	dikey		0.8773	0.8773	0.9386	0.8777	0.8770

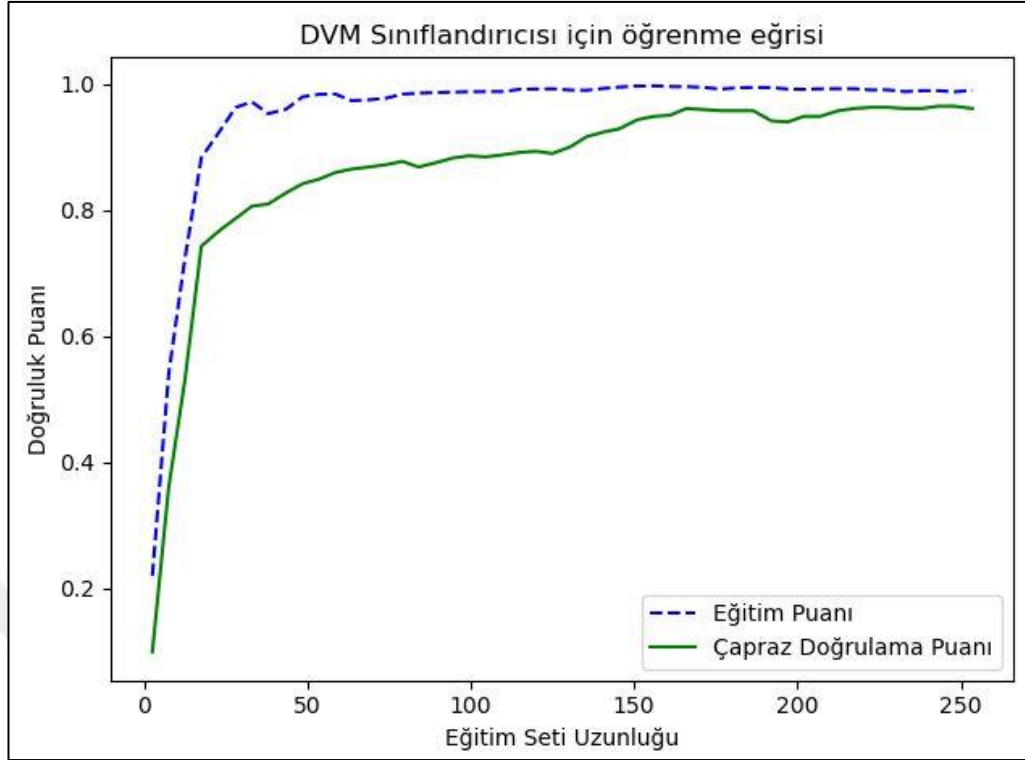
Çizelge 6.15’de arttırılmış dengeli CAME-II veri setinde TV-L1_dikey optik akış görüntü kareleri ile birlikte kullanılan Xception derin öğrenme modeli en iyi performansı sağlamaktadır. Xception modelinden elde edilen en iyi performans ölçüm sonuçları doğruluk 0.9248, duyarlılık 0.9248, özgüllük 0.9624, kesinlik 0.9251 ve f1 0.9249 olarak alınmıştır. Böylece üçüncü model için yapılan ilk deneysel çalışmadaki en başarılı ölçüm sonuçları Çizelge 6.13 kaynak alındığında, yeni deneysel çalışma ile birlikte önerilen modelin performansında %9.88 doğruluk, %13.29 duyarlılık, %6.64 özgüllük, %12.48 kesinlik, %12.97 F1 oranında olumlu bir artış gözlemlenmektedir. Önerilen üçüncü modele ait karmaşıklık matrisi Şekil 6.7’de, AİK eğrisinin grafiği Şekil 6.8’de ve öğrenme eğrisine ait grafik Şekil 6.9’da verilmiştir.

	Negatif	Pozitif	Sürpriz
Negatif	93	3	4
Pozitif	5	92	3
Sürpriz	5	3	93

Şekil 6.7. Xception modeline ait karmaşıklık matrisi.



Şekil 6.8. Xception modeline ait AİK eğrisi grafiği.



Şekil 6.9. Xception modeline ait öğrenme eğrisi grafiği.

Tez kapsamında önerilen Xnet-ÇDÖÖE model çalışması değerlendirildiğinde, yapılan ilk deneysel çalışmada CASME-II veri seti üzerinde kullanılan modellerin ölçüm sonuçlarında F1 ve duyarlılık değerleri ile doğruluk değerleri arasında büyük farklar olduğu görülmektedir. Bunun sebebi negatif sınıfa ait örneklerin diğer iki sınıfa ait örneklerin toplamından daha fazla olması ve sınıflar arasında örneklem dengesizliğidir. Dominant örneklemeye sahip sınıfların bulunduğu veri setlerinin kullanıldığı modellerin performansı değerlendirildiğinde, doğruluk ölçütünün tek başına değerlendirilmesi önyargı oluşturabilmektedir. Dolayısıyla F1 ve duyarlılık ölçümleri, dengesiz örneklemeli sınıfları bulunan veri setinin sınıflandırılmasında tüm sınıfların eşit derecede iyi tahmin edilip edilmediği hakkında kararlı bir değerlendirme sağlamaktadır [15,19]. Her ne kadar doğruluk ölçümleri kısmen iyi gibi görünse de F1 ve duyarlılık değerleri bazı ölçümlerde %50'nin altında olduğu gözlemlenmiştir. Diğer performans ölçütlerinin de başarısını artırıp daha dengeli ölçümler elde etmek için veri arttırma tekniği kullanılmış, veri seti çoğaltılıp dengelenmiştir. Veri arttırma sonrası

Çizelge 6.13 ve Çizelge 6.14'daki Farneback görüntü kareleri kullanılarak yapılan Xception modeli deneyleri karşılaştırıldığında %16 yı geçen F1 ve %18'i geçen duyarlılık ölçümleri gözlenmektedir. Buradan veri artırma tekniğinin kullanılarak dengeli örneklemeye sahip veri setinin Xception ve Gabor modelleri ile sınıflandırılması, Mİ tanıma performansını arttırdığı sonucuna ulaşılabilir [14].

Özellik seçim yönteminden faydalandığında ise Mİ tanımada diğer özelliklerden daha etkili olan özelliklerin seçilmesiyle gerçekleştirilen sınıflandırmaya ait deneysel sonuçlar modelin başarısında olumlu sonuçlar vermiştir. ÇDÖÖE özellik seçim algoritmasından faydalanılarak yapılan ölçümlerde başlangıçtaki (Çizelge 6.13) deneysel çalışmadaki en başarılı ölçüm sonuçları ile karşılaştırıldığında %9.88 doğruluk ve %12.97 F1 oranında modelin performansını geliştirdiği gözlemlenmiştir. Özellik seçim algoritması kullanılarak elde edilen özelliklerin, Mİ tanıma performansını geliştirdiği deneysel ölçümler sonunda gözlemlenmiştir.

Xception teknolojisinin üstünlüklerinden biriside Inception modelinin yapısında olduğu gibi derinlemesine ayrıştırılabilir evrişimden ve noktasal evrişimden yararlanılarak parametre sayısını büyük oranda azaltıp hesaplama maliyetini azaltmaktadır. Xception modelinin kendisinden önceki versiyonu olan Inception modelinden farkı, önce noktasal evrişim daha sonra derinlemesine ayrıştırılabilir evrişim uygulamaktadır. Bu fark sayesinde Inception modelinden hem daha az parametre kullanmakta hem de doğruluk ve hata oranı performansında daha fazla performans göstermektedir [130].

Çalışma [184] 'de farklı optik akış yöntemleri analiz edildiğinde farklı blok boyutlarında en sağlam iki yöntem olarak Farneback ve TV-L1 sonucuna varılmıştır. Bundan dolayı Mİ tanıma görevi için önerilen modelimizde bu iki yöntemin dışında TV-L1 yatay ve dikey görüntü kareleri de kullanılmıştır. Optik akış tekniğinin kullanılmasının başka bir üstünlüğü ise video kare dizilerinin tamamının kullanmak yerine, referans olarak hareketin başladığı başlangıç karelerinin ve hareketin zirve yaptığı tepe karelerinin kullanılması hesaplama yükü azaltılmaktadır. En başarılı sonuçların TV-L1 dikey bileşenlere sahip görüntü karelerinden elde edildiği deneysel

sonuçlar ile gözlemlenmiştir. Önerilen model yapısının dezavantajlarından birisi ise ÇDÖÖE algoritmasının özyinelemeli yapısından dolayı özellik haritası sayısı arttıkça, filtreleme süresi de doğrusal orantıda artmaktadır. Farklı özellik seçim algoritmalarından faydalanılarak model maliyetinde ve performansında daha başarılı sonuçlar gözlemlenebilir.

6.4. TARTIŞMA

Bu tez kapsamında önerilen Mİ modelleri ve literatürde bulunan diğer son teknoloji Mİ modellerinin karşılaştırılması Çizelge 6.16’de sunulmuştur. Buna göre modelimizin rekabetçi ve sonuçların tatmin edici olduğu görülmektedir.

Çizelge 6.16. Önerilen çalışmanın en son modellerle karşılaştırılması

Metot	Doğruluk	F1	Veri Seti
STSTNet [185]	0.6667	0.6758	SAMM+CASME II+SMIC/MEVIEW (eğitim / test)
OFF-ApexNet [29]	0.746	0.710	SMIC+CASME II +SAMM
From Macro to Micro[49]	0.747	0.64	CASME II +SAMM
Residual Network with Micro-Attention [43]	0.763	0.668	CASME II +SAMM
STSTNet [20]	0.769	0.739	SMIC+CASME II +SAMM
STSTNet+GA [24]	0.859	0.837	SMIC+CASME II +SAMM
Deep3DCANN [186]	0.86	0.84	CASME II
Önerilen DTÖ-PSO[14]	0.8784	0.8765	SMIC+CASME II +SAMM
Önerilen DensNet169- ÖÖE	0.9019	0.9017	SMIC+CASME II+SAMM
DFN [187]	0.9028	0.8923	CASME II
Önerilen XNet-ÇDÖÖE	0.9248	0.9249	CASME II

Önerilen modeller ile literatürde bulunan son teknoloji modeller değerlendirildiğinde, Micro-Attention [43], From Macro to Micro[49] adlı modellerinde artık ağ (residual network) kullanan çalışmalar da, önerilen modellerimizden farklı olarak ImageNet üzerinde eğitilmiş ağ kullanmak yerine Makro ifadeler üzerinde eğitip burdaki

ağırlıkları Mİ tanıma görevinde kullanılmıştır. Ayrıca bu iki model çalışmalarında hareket özelliği içeren optik akış görüntüleri yerine statik RGB tepe karelerini kullanmayı tercih etmişlerdir. Bu tez kapsamında önerilen modellerin hepsinde ImageNet ağırlıkları ve optik akış görüntüleri kullanıldığı için modellerin karşılaştırılması çok doğru ve kesin bilgiler vermeyebilir. DTÖ-PSO adlı modelimizde tek bir ağın eğitiminden elde edilen özellik yerine beş farklı ağdan özellik elde edilerek birleştirilmiş, her ağın kendine ait belirli üstünlüklerinden faydalanılarak daha değerli ve farklı özellikler elde edilmesi hedeflenmiştir. Hem beş farklı modelden farklı özelliklerin elde edilmesi, statik RGB tepe kareleri yerine hareket özelliklerini içeren daha dinamik Farneback optik akış görüntülerinin modellerin girişine verilmesi ve ayrıca PSO algortimasının Mİ tanıma görevinde ilgili özellikleri filtrelemesinden ötürü DTÖ-PSO modelimizin bu iki modelden daha üstün performans vermesinin sebebi olarak görülebilir. Ayrıca ifade tanıma görevinin diğer tanıma görevlerinden farklı olduğu [43] düşünüldüğünde modellerimizde ImageNet ağırlıkları yerine ifadeler üzerinde eğitilmiş ağ ağırlıkları kullanılabilirdi.

Modellerinde sıkı bir ESA modellerini kullanan STSTNet [185], DFN [187], STSTNet [20], STSTNet+GA [24] çalışmaları bu tez kapsamında önerilen DensNet169-ÖÖE modelinden farklı olarak TV-L1 optik akış görüntülerinin yatay, dikey ve optik gerinim özelliklerini ayrıştırarak iki ve üç akışlı ESA modellerine girdi olarak verilmiştir. Bu modellerde sınıflandırma sürecinden önce her bir akışa ait vektörleri birleştirerek sınıflandırma işlemi yapılmıştır. Önerilen DensNet169-ÖÖE modellerinde ise, TV-L1 optik akış görüntüleri içerisinde yatay, dikey ve optik gerinim özellikleri barındırdığından dolayı TV-L1 görüntülerini olduğu gibi modelimizde kullanılmış ayrıca farklı özellikleri barındıran Farneback optik akış özellikleri ve statik olarak Mİ içeren RGB tepe kareleri kullanılarak üç akışlı modelimize girdi olarak verilmiştir. Böylece daha zengin ve daha farklı bilgilere sahip olan özellik haritaları elde edilmek istenmiştir. Bu modeller arasında sadece DFN modelinin DensNet169-ÖÖE ve DTÖ-PSO modelimizden üstün performans vermesinin nedeni ise bünyesinde çok katmanlı ve her katmanda birden çok rastgele orman (Random Forest - RF) içeren Derin Orman (Deep Forest - DF) yaklaşımını benimsemeleri ile özellik temsil

yeteneğinin daha da gelişmiş olduğu fikri düşünülebilir. Bundan sonraki çalışmalarda bu ve buna benzer farklı yaklaşımların da göz önüne alınmasının faydalı olacağı düşünülmektedir.

Modellerinde sığ bir ESA modeli kullanan OFF-ApexNet [29] çalışması ise sadece yatay ve dikey optik akış görüntülerini iki akışlı sığ bir modellerinin girdisi olarak kullanmışlardır. Model sınıflandırma işleminden önce her akışa ait özellik vektörünü birleştirerek sınıflandırma işlemi gerçekleştirilmektedir. Bu tez kapsamında önerilen tek akışlı Xnet-ÇDÖÖE modelinin geliştirme aşamasında TV-L1 optik akış görüntüleri, TV-L1 optik akış görüntülerinin hem yatay ve dikey bileşenlerine ait görüntüler hem de Farneback optik akış görüntüleri ve statik RGB tepe kareleri karşılaştırılmış, en iyi performansı veren TV-L1_dikey görüntüleri olduğu deneyler sonucunda belirlenmiştir. Mİ tanıma görevi için önerilen Xnet-ÇDÖÖE modeli DensNet169-ÖÖE modelinden farklı olarak tek akışlı olmasında dolayı en iyi performansı veren, daha özet ve yalın özelliğe sahip TV-L1_dikey görüntüleri, TV-L1 optik akış görüntülerinden elde edilerek modelin performansının artırılması hedeflenmiştir. Xnet-ÇDÖÖE modelinin önerilen ve karşılaştırılan tüm modellerden daha üstün sonuçlar vermesinin nedeni ise, modelimizde daha öz, değerli bilgileri içeren TV-L1_dikey görüntülerinin ve ÇDÖÖE özellik seçim algoritmasının Mİ tanıma görevi için daha uygun olduğu fikri düşünülebilir.

Mİ veri setleri ile ilgili hala bazı sınırlamalar bulunmaktadır. Modellerde öğretilen Mİ verileri, veri setlerinde bulunan dugu kümesinin elemanlarını içerecek şekilde etiketlenmiştir. Fakat veri setlerinde bulunan duyguların dışında farklı duygulara ait mikro ifadelerde bulunmaktadır. Bu tez çalışmasının sınırlılıklarından biri de etiket kümesi haricindeki bir Mİ için önceden eğitilmemiş olması, etiket kümeleri dışındaki Mİ' ler için çözüm üretemeyeceğidir. Bir diğer sınırlama ise bu veri setlerindeki Mİ belirli bir laboratuvar ortamında elde edilmesidir. Katılımcılardan yüz ifadelerini kontrol etmeleri ve bastırmaları istenmektedir. Gerçek hayat üzerinde koşullanmadan elde edilecek Mİ örnekleri daha farklı olabileceği göz önüne alınmalıdır.

BÖLÜM 7

SONUÇ VE ÖNERİLER

Mikro ifade, kişinin kendini riskli gördüğü ortamlarda, duygularını gizlemek istemesine rağmen, kişinin yüzünün bir bölümünde, düşük yoğunluklu ve kısa zaman diliminde oluşan küçük ifadelerdir. Bu ifadeler tamamlamamış bir hareket olduğundan dolayı insan gözüyle farkedilmesi oldukça zordur. Bu ifadelerin oluştukları riskli ortamları yapay olarak meydana getirmek ve o ifadeleri kayıt altına almak, doğal mikro ifadedeki duyguları yakalamanın zorluğundan dolayı mikro ifade veri seti oluşturmak oldukça zordur. Bu yüzden günümüzde var olan halka açık doğal ortamlardaki doğal yüz ifadesi görüntülerinden oluşturulan veri setlerindeki örneklemeler yetersiz olup bunun üstesinden gelebilmek için veri setlerinin farklı kombinasyonlarının birleşimi ile yapılmış ve kullanılmıştır.

Bu tez çalışması kapsamında, halka açık en popüler doğal mikro ifade veri setleri CASME II, SMIC, SAMM birleştirilerek yeni bileşik bir veri seti oluşturulmuştur. Farklı duygu sınıflarına sahip bu veri setleri MEGC protokolüne en düşük duygu sınıfı içeren SMIC veri setinin üç sınıfına (negatif, pozitif, sürpriz) indirgenmiştir. Buna göre CASME II ve SAMM veri setlerinde bulunan baskı, öfke, küçümseme, iğrenme, korku ve üzüntü sınıfları negatif sınıf altında toplanırken, mutluluk sınıfı pozitif sınıf olarak sürpriz sınıfı ise değiştirilmeden sürpriz sınıfı adıyla deneylerde kullanılmıştır. CASME II setinde önışleme aşamasında gerçekleştirilen hem yüz bölgesinin çıkartılması hem de anahtar kare olarak adlandırılan başlangıç, tepe ve bitiş karelerinin konum indeksleri kullanıcıların hizmetine sunulmaktadır. SAMM veri setinde sadece anahtar kareler sağlanırken SMIC veri setinde ise sadece yüz bölgesinin çıkarıldığı görüntüler kullanıcılara hazır olarak sağlanmaktadır. SAMM ve SMIC veri setlerinin bulunduğu bileşik veri seti ile yapılan deneysel çalışmalarda SAMM için yüz bölgesinin çıkarılması, SMIC için ise anahtar karelerin tespit edilmesi önışlem süreci gerçekleştirilmiştir.

Bu tez çalışması kapsamında önerilen DTÖ-PSO modelinde, bileşik veri setindeki mikro ifade görüntülerini sınıflandırmak için beş farklı ESA modellerden elde edilen özelliklerin bileşimine dayanan bir tanıma modeli önerilmiştir. Farneback optik akış görüntülerini girdi olarak alan AlexNet, VGG16, MobilenetV2, EfficientNet, Squeezenet ağlarının tam bağlı katmanlarından 1000'er adet toplamda 5000 adet özellik elde edilmiştir. Bu 5000 görüntü özelliği birleştirilmiş, içlerinde mikro ifade tanıma görevinde en ayırt edici özelliğe sahip olanlar, PSO algoritması ile seçilmiştir. Deneysel sonuçlar veri setlerindeki öreneklem yetersizliğinden kaynaklı olarak bu beş ESA modelinin sağladığı özelliklerin kullanılmasının, sağlam ve tutarlı bir sitem sağlamak için yeterli olmadığını göstermiştir. Bundan dolayı veri artırma tekniğinin, özellik seçimi için PSO algoritmasının ve çapraz doğrulama tekniğinin önerilen modelin yapısında yer alması modelin verimliliğini arttırmıştır. Sonuç olarak, en iyi sonuçlar ortalama Doğruluk % 87.84 Duyarlılık % 87.84 Özgüllük % 93.92 Kesinlik % 88.39 ve F1 % 87.65 ile sağlanmıştır.

DenseNet169-ÖÖE modelinde ise mikro ifade görüntülerini sınıflandırmak için veriye dayalı yöntemlerden DenseNet121, DenseNet169, DenseNet201 ESA karşılaştırıldığı ve en iyi sonu veren DenseNet169' un da yapısında bulunduğu bir model önerilmiştir. Bileşik veri setinin kullanıldığı üç akışlı modelde, Farneback, TV-L1 ve optik akış görüntüleri ve mikro ifade video kare dizindeki tepe kare RGB görüntüsü girdi olarak alınmaktadır. Her bir görüntüden 81536 toplamda 244608 adet özellik, modelin sınıflandırma katmanında bulunan küresel ortalama havuzlamadan hemen önceki Relu Activation işleminden sonra elde edilmiştir. Mikro ifade tanıma için en ayırt edici 6200 özellikler özyinelemeli özellik eleme algoritması ile seçilmiş, seçilen bu özellikler doğrusal DVM kullanılarak sınıflandırılmıştır. Deneysel sonuçlar veri setindeki öreneklem yetersizliğinden kaynaklı olarak DenseNet169 modelinin sağladığı özelliklerin kullanılmasının, sağlam ve tutarlı bir sitem sağlamak için yeterli olmadığını göstermiştir. Veri artırma tekniği ve özellik seçim algoritması modelin performansında çok verimli artışlara neden olmuştur. Sonuç olarak, modele ait en iyi deneysel ölçümler ortalama doğruluk %90, duyarlılık %90, özgüllük % 95, kesinlik % 90.07 ve F1 % 89.99 ile elde edilmiştir.

Önerilen Xnet-ÇDÖÖE modelinde, mikro ifade görüntülerini sınıflandırmak için veriye dayalı yöntemlerden Xception ESA'nın yer aldığı uygulamada ayrıca geleneksel yöntemlerden Gabor modeli ile de karşılaştırılmıştır. Farneback, TV-L1 ve TV-L1 yatay

ve dikey optik akış görüntülerini ve mikro ifade video kare dizindeki tepe kare RGB görüntüsünü girdi olarak alan bu iki modelde önceki deneysel çalışmaların aksine sadece CASME II verisetinden yararlanılmıştır. Bu görüntülerinden Xception'dan 100352 adet özellik ve Gabor'dan 50176 adet özellik çıkarılarak negatif, pozitif ve sürpriz olmak üzere üç kategoriye sınıflandırılmıştır. Xception açısından elde edilen özellik haritaları, küresel ortalama havuzlama katmanından hemen önceki hemen önceki 'block14_sepconv2_act' isimli aktivasyon katmanının çıktısından elde edilmiştir. Mikro ifade tanıma görevinde en ayırt edici özelliğe sahip olanlar çapraz doğrulama ile özyinelemeli özellik eleme algoritması ile seçilmiştir. Deneysel sonuçlar veri setindeki örnekleme yetersizliğinden kaynaklı olarak bu iki modelinin sağladığı özelliklerin kullanılmasının, sağlam ve tutarlı bir sistem sağlamak için yeterli olmadığını göstermiştir. Bundan dolayı veri artırma tekniğinin ve özellik seçimi için çapraz doğrulama ile özyinelemeli özellik eleme algoritmasının önerilen modelin yapısında yer alması modelin verimliliğini arttırmıştır. Sonuç olarak, modele ait en iyi ölçüm sonuçları doğruluk % 92.48, duyarlılık % 92.48, özgüllük % 96.24, kesinlik % 92.51 ve F1 % 92.49 ile elde edilmiştir

Sonuç olarak mikro ifade tepe görüntülerinin, optik akış görüntülerinin ve optik akış görüntü bileşenlerinin girdi olarak kullanıldığı önerilen modeller önışleme, özellik çıkarma, özellik seçme ve sınıflandırma bölümlerinden oluşmaktadır. Önerilen modeller değerlendirildiğinde % 92.48 doğruluk, % 92.49 F1 skor ile en yüksek başarımlar elde edilmiştir. Elde edilen bu sonuç literatürde benzer veri setleri üzerinde yapılan çalışmalar ile kıyaslaması yapılarak modelin başarısı gösterilmiştir.

Gelecekte yapılması planlanan çalışmalardan ilki mikro ifade video klip görüntülerinden uzamsal ve zamansal 3B özelliklerin kullanılarak sınıflandırıcıların performanslarının artırılması hedeflenmektedir. Bir diğer hedeflenen çalışma ise gerçek zamanlı mikro ifade tanıyan bilgisayar destekli uygulama geliştirmek olarak belirlenmiştir. Amaçlanan çalışmalar üzerinde optimizasyon algoritmaları ve yeni evrimsel modeller kullanılarak test edilebilir.

KAYNAKLAR

- [1] Y. Li, J. Wei, Y. Liu, J. Kauttonen, ve G. Zhao, “Deep Learning for Micro-Expression Recognition: A Survey”, *IEEE Trans. Affect. Comput.*, c. 13, sayı 4, ss. 2028–2046, 2022, doi: 10.1109/TAFFC.2022.3205170.
- [2] M. Peng, C. Wang, T. Chen, G. Liu, ve X. Fu, “Dual temporal scale convolutional neural network for micro-expression recognition”, *Front. Psychol.*, c. 8, s. 1745, 2017.
- [3] D. Patel, X. Hong, ve G. Zhao, “Selective deep features for micro-expression recognition”, içinde *2016 23rd international conference on pattern recognition (ICPR)*, 2016, ss. 2258–2263.
- [4] X. Huang, S.-J. Wang, G. Zhao, ve M. Piteikainen, “Facial micro-expression recognition using spatiotemporal local binary pattern with integral projection”, içinde *Proceedings of the IEEE international conference on computer vision workshops*, 2015, ss. 1–9.
- [5] L. Fan, J. He, Y. Zheng, Y. Nie, T. Chen, ve H. Zhang, “Facial micro-expression recognition impairment and its relationship with social anxiety in internet gaming disorder”, *Curr. Psychol.*, 2022, doi: 10.1007/s12144-022-02958-7.
- [6] W.-J. Yan, Q. Wu, J. Liang, Y.-H. Chen, ve X. Fu, “How Fast are the Leaked Facial Expressions: The Duration of Micro-Expressions”, *J. Nonverbal Behav.*, c. 37, sayı 4, ss. 217–230, 2013, doi: 10.1007/s10919-013-0159-8.
- [7] Y.-H. Oh, A. C. Le Ngo, J. See, S.-T. Liong, R. C.-W. Phan, ve H.-C. Ling, “Monogenic Riesz wavelet representation for micro-expression recognition”, içinde *2015 IEEE International Conference on Digital Signal Processing (DSP)*, 2015, ss. 1237–1241.
- [8] V. Esmaeili, M. Mohassel Feghhi, ve S. O. Shahdi, “A comprehensive survey on facial micro-expression: approaches and databases”, *Multimed. Tools Appl.*, c. 81, sayı 28, ss. 40089–40134, 2022, doi: 10.1007/s11042-022-13133-2.
- [9] Z. Xia, X. Feng, J. Peng, X. Peng, ve G. Zhao, “Spontaneous micro-expression spotting via geometric deformation modeling”, *Comput. Vis. Image Underst.*, c. 147, ss. 87–94, 2016.
- [10] X. Li vd., “Towards Reading Hidden Emotions: A Comparative Study of Spontaneous Micro-Expression Spotting and Recognition Methods”, *IEEE Trans. Affect. Comput.*, c. 9, sayı 4, ss. 563–577, 2018, doi: 10.1109/TAFFC.2017.2667642.
- [11] Y. Zhao ve J. Xu, “Compound Micro-Expression Recognition System”, içinde *2020 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, 2020, ss. 728–733, doi: 10.1109/ICITBS49701.2020.00161.
- [12] S.-T. Liong, J. See, K. Wong, ve R. C.-W. Phan, “Less is more: Micro-expression recognition from video using apex frame”, *Signal Process. Image Commun.*, c. 62, ss. 82–92, 2018.

- [13] S. Zhao vd., “A two-stage 3D CNN based learning method for spontaneous micro-expression recognition”, *Neurocomputing*, c. 448, ss. 276–289, 2021.
- [14] M. Z. Uzun, Y. Celik, ve E. Basaran, “Micro-Expression Recognition by Using CNN Features with PSO Algorithm and SVM Methods”, *learning*, c. 2, sayı 3, ss. 5–8.
- [15] Y.-H. Oh, J. See, A. C. Le Ngo, R. C.-W. Phan, ve V. M. Baskaran, “A survey of automatic facial micro-expression analysis: databases, methods, and challenges”, *Front. Psychol.*, c. 9, s. 1128, 2018.
- [16] G. Zhao ve X. Li, “Automatic micro-expression analysis: open challenges”, *Front. Psychol.*, c. 10, s. 1833, 2019.
- [17] K. M. Goh, C. H. Ng, L. L. Lim, ve U. U. Sheikh, “Micro-expression recognition: an updated review of current trends, challenges and solutions”, *Vis. Comput.*, c. 36, sayı 3, ss. 445–468, 2020, doi: 10.1007/s00371-018-1607-6.
- [18] X. Jia, X. Ben, H. Yuan, K. Kpalma, ve W. Meng, “Macro-to-micro transformation model for micro-expression recognition”, *J. Comput. Sci.*, c. 25, ss. 289–297, 2018, doi: <https://doi.org/10.1016/j.jocs.2017.03.016>.
- [19] J. See, M. H. Yap, J. Li, X. Hong, ve S. Wang, “MEGC 2019 – The Second Facial Micro-Expressions Grand Challenge”, içinde *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)*, 2019, ss. 1–5, doi: 10.1109/FG.2019.8756611.
- [20] S.-T. Liong, Y. S. Gan, J. See, H.-Q. Khor, ve Y.-C. Huang, “Shallow triple stream three-dimensional cnn (ststnet) for micro-expression recognition”, içinde *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)*, 2019, ss. 1–5.
- [21] Y. Liu, H. Du, L. Zheng, ve T. Gedeon, “A neural micro-expression recognizer”, içinde *2019 14th IEEE international conference on automatic face & gesture recognition (FG 2019)*, 2019, ss. 1–4.
- [22] M. Takalkar, M. Xu, Q. Wu, ve Z. Chaczko, “A survey: facial micro-expression recognition”, *Multimed. Tools Appl.*, c. 77, sayı 15, ss. 19301–19325, 2018.
- [23] M. H. Yap, J. See, X. Hong, ve S.-J. Wang, “Facial Micro-Expressions Grand Challenge 2018 Summary”, içinde *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, 2018, ss. 675–678, doi: 10.1109/FG.2018.00106.
- [24] K.-H. Liu, Q.-S. Jin, H.-C. Xu, Y.-S. Gan, ve S.-T. Liong, “Micro-expression recognition using advanced genetic algorithm”, *Signal Process. Image Commun.*, c. 93, s. 116153, 2021, doi: <https://doi.org/10.1016/j.image.2021.116153>.
- [25] S.-J. Wang vd., “Micro-expression recognition with small sample size by transferring long-term convolutional neural network”, *Neurocomputing*, c. 312, ss. 251–262, 2018, doi: <https://doi.org/10.1016/j.neucom.2018.05.107>.
- [26] L. Yao, X. Xiao, R. Cao, F. Chen, ve T. Chen, “Three Stream 3D CNN with SE

- Block for Micro-Expression Recognition”, içinde *2020 International Conference on Computer Engineering and Application (ICCEA)*, 2020, ss. 439–443.
- [27] X. Zeng, X. Zhao, X. Zhong, ve G. Liu, “A Survey of Micro-expression Recognition Methods Based on LBP, Optical Flow and Deep Learning”, *Neural Process. Lett.*, 2023, doi: 10.1007/s11063-022-11123-x.
 - [28] Z. Xia, X. Hong, X. Gao, X. Feng, ve G. Zhao, “Spatiotemporal recurrent convolutional networks for recognizing spontaneous micro-expressions”, *IEEE Trans. Multimed.*, c. 22, sayı 3, ss. 626–640, 2019.
 - [29] Y. S. Gan, S.-T. Liong, W.-C. Yau, Y.-C. Huang, ve L.-K. Tan, “OFF-ApexNet on micro-expression recognition system”, *Signal Process. Image Commun.*, c. 74, ss. 129–139, 2019.
 - [30] M. Shreve, S. Godavarthy, D. Goldgof, ve S. Sarkar, “Macro- and micro-expression spotting in long videos using spatio-temporal strain”, içinde *2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG)*, 2011, ss. 51–56, doi: 10.1109/FG.2011.5771451.
 - [31] Y. He, S.-J. Wang, J. Li, ve M. H. Yap, “Spotting Macro-and Micro-expression Intervals in Long Video Sequences”, içinde *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*, 2020, ss. 742–748, doi: 10.1109/FG47880.2020.00036.
 - [32] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, ve W. Freeman, “Eulerian video magnification for revealing subtle changes in the world”, *ACM Trans. Graph.*, c. 31, sayı 4, ss. 1–8, 2012.
 - [33] L. Cai, H. Li, W. Dong, ve H. Fang, “Micro-expression recognition using 3D DenseNet fused Squeeze-and-Excitation Networks”, *Appl. Soft Comput.*, c. 119, s. 108594, 2022, doi: <https://doi.org/10.1016/j.asoc.2022.108594>.
 - [34] J. Li, Y. Wang, J. See, ve W. Liu, “Micro-expression recognition based on 3D flow convolutional neural network”, *Pattern Anal. Appl.*, c. 22, sayı 4, ss. 1331–1339, 2019, doi: 10.1007/s10044-018-0757-5.
 - [35] J. Wei, G. Lu, ve J. Yan, “A comparative study on movement feature in different directions for micro-expression recognition”, *Neurocomputing*, c. 449, ss. 159–171, 2021, doi: <https://doi.org/10.1016/j.neucom.2021.03.063>.
 - [36] N. Liu, X. Liu, Z. Zhang, X. Xu, ve T. Chen, “Offset or Onset Frame: A Multi-Stream Convolutional Neural Network with CapsuleNet Module for Micro-expression Recognition”, içinde *2020 5th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)*, 2020, ss. 236–240.
 - [37] L. Zhou, Q. Mao, ve L. Xue, “Dual-inception network for cross-database micro-expression recognition”, içinde *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)*, 2019, ss. 1–5.
 - [38] S.-J. Wang vd., “Micro-expression recognition using color spaces”, *IEEE Trans. Image Process.*, c. 24, sayı 12, ss. 6034–6047, 2015.

- [39] Y. Wang, J. See, R. C.-W. Phan, ve Y.-H. Oh, “Efficient spatio-temporal local binary patterns for spontaneous facial micro-expression recognition”, *PLoS One*, c. 10, sayı 5, s. e0124674, 2015.
- [40] C. Lin, F. Long, J. Huang, ve J. Li, “Micro-Expression Recognition Based on Spatiotemporal Gabor Filters”, içinde *2018 Eighth International Conference on Information Science and Technology (ICIST)*, 2018, ss. 487–491, doi: 10.1109/ICIST.2018.8426088.
- [41] P. Zhang, X. Ben, R. Yan, C. Wu, ve C. Guo, “Micro-expression recognition system”, *Optik (Stuttg.)*, c. 127, sayı 3, ss. 1395–1400, 2016, doi: <https://doi.org/10.1016/j.ijleo.2015.10.217>.
- [42] R. Zhi, H. Xu, M. Wan, ve T. Li, “Combining 3D convolutional neural networks with transfer learning by supervised pre-training for facial micro-expression recognition”, *IEICE Trans. Inf. Syst.*, c. 102, sayı 5, ss. 1054–1064, 2019.
- [43] C. Wang, M. Peng, T. Bi, ve T. Chen, “Micro-attention for micro-expression recognition”, *Neurocomputing*, c. 410, ss. 354–362, 2020.
- [44] X. Ben, X. Jia, R. Yan, X. Zhang, ve W. Meng, “Learning effective binary descriptors for micro-expression recognition transferred by macro-information”, *Pattern Recognit. Lett.*, c. 107, ss. 50–58, 2018, doi: <https://doi.org/10.1016/j.patrec.2017.07.010>.
- [45] N. Van Quang, J. Chun, ve T. Tokuyama, “CapsuleNet for micro-expression recognition”, içinde *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)*, 2019, ss. 1–7.
- [46] B. Xia, W. Wang, S. Wang, ve E. Chen, “Learning from Macro-expression: a Micro-expression Recognition Framework”, içinde *Proceedings of the 28th ACM International Conference on Multimedia*, 2020, ss. 2936–2944.
- [47] M. A. Takalkar ve M. Xu, “Image Based Facial Micro-Expression Recognition Using Deep Learning on Small Datasets”, içinde *2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, 2017, ss. 1–7, doi: 10.1109/DICTA.2017.8227443.
- [48] H.-X. Xie, L. Lo, H.-H. Shuai, ve W.-H. Cheng, “Au-assisted graph attention convolutional network for micro-expression recognition”, içinde *Proceedings of the 28th ACM International Conference on Multimedia*, 2020, ss. 2871–2880.
- [49] M. Peng, Z. Wu, Z. Zhang, ve T. Chen, “From macro to micro expression recognition: Deep learning on small datasets using transfer learning”, içinde *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, 2018, ss. 657–661.
- [50] L. Zhou, Q. Mao, ve L. Xue, “Cross-database micro-expression recognition: a style aggregated and attention transfer approach”, içinde *2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, 2019, ss. 102–107.
- [51] Y.-J. Liu, J.-K. Zhang, W.-J. Yan, S.-J. Wang, G. Zhao, ve X. Fu, “A main directional mean optical flow feature for spontaneous micro-expression

- recognition”, *IEEE Trans. Affect. Comput.*, c. 7, sayı 4, ss. 299–310, 2015.
- [52] X. Huang, G. Zhao, X. Hong, W. Zheng, ve M. Pietikäinen, “Spontaneous facial micro-expression analysis using spatiotemporal completed local quantized patterns”, *Neurocomputing*, c. 175, ss. 564–578, 2016.
- [53] B. Sun, S. Cao, D. Li, J. He, ve L. Yu, “Dynamic Micro-Expression Recognition Using Knowledge Distillation”, *IEEE Trans. Affect. Comput.*, c. 13, sayı 2, ss. 1037–1043, 2022, doi: 10.1109/TAFFC.2020.2986962.
- [54] V. Mayya, R. M. Pai, ve M. M. M. Pai, “Combining temporal interpolation and DCNN for faster recognition of micro-expressions in video sequences”, içinde *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2016, ss. 699–703, doi: 10.1109/ICACCI.2016.7732128.
- [55] M. A. Takalkar, M. Xu, ve Z. Chaczko, “Manifold feature integration for micro-expression recognition”, *Multimed. Syst.*, c. 26, sayı 5, ss. 535–551, 2020, doi: 10.1007/s00530-020-00663-8.
- [56] B. Martinez, M. F. Valstar, B. Jiang, ve M. Pantic, “Automatic analysis of facial actions: A survey”, *IEEE Trans. Affect. Comput.*, c. 10, sayı 3, ss. 325–347, 2017.
- [57] J. F. Cohn, Z. Ambadar, ve P. Ekman, “Observer-based measurement of facial expression with the Facial Action Coding System”, *Handb. Emot. elicitation Assess.*, c. 1, sayı 3, ss. 203–221, 2007.
- [58] P. Ekman ve W. V Friesen, “Facial action coding system”, *Environ. Psychol. Nonverbal Behav.*, 1978.
- [59] R. Zhi, M. Liu, ve D. Zhang, “A comprehensive survey on automatic facial action unit analysis”, *Vis. Comput.*, c. 36, ss. 1067–1093, 2020.
- [60] J. K. Thompson, “Right Brain, Left Brain; Left Face, Right Face: Hemisphericity and the Expression of Facial Emotion”, *Cortex*, c. 21, sayı 2, ss. 281–299, 1985, doi: [https://doi.org/10.1016/S0010-9452\(85\)80033-2](https://doi.org/10.1016/S0010-9452(85)80033-2).
- [61] F. Trémeau vd., “Facial expressiveness in patients with schizophrenia compared to depressed patients and nonpatient comparison subjects”, *Am. J. Psychiatry*, c. 162, sayı 1, ss. 92–101, 2005.
- [62] P. Ekman, W. V Freisen, ve S. Ancoli, “Facial signs of emotional experience.”, *J. Pers. Soc. Psychol.*, c. 39, sayı 6, s. 1125, 1980.
- [63] Q. Li, S. Zhan, L. Xu, ve C. Wu, “Facial micro-expression recognition based on the fusion of deep learning and enhanced optical flow”, *Multimed. Tools Appl.*, c. 78, sayı 20, ss. 29307–29322, 2019, doi: 10.1007/s11042-018-6857-9.
- [64] G. Warren, E. Schertler, ve P. Bull, “Detecting deception from emotional and unemotional cues”, *J. Nonverbal Behav.*, c. 33, sayı 1, ss. 59–69, 2009.
- [65] A. Metin, “Yüz ifadelerindeki duygular: Derleme çalışması”, *OPUS Int. J. Soc. Res.*, c. 10, sayı 17, ss. 2027–2055, 2019.

- [66] C. M. Hurley, “Do you see what I see? Learning to detect micro expressions of emotion”, *Motiv. Emot.*, c. 36, sayı 3, ss. 371–381, 2012.
- [67] R. W. Levenson, “The autonomic nervous system and emotion”, *Emot. Rev.*, c. 6, sayı 2, ss. 100–112, 2014.
- [68] R. Zhang, N. He, Y. Wu, Y. He, ve K. Yan, “To balance: balanced micro-expression recognition”, *Multimed. Syst.*, c. 28, sayı 1, ss. 335–345, 2022, doi: 10.1007/s00530-021-00842-1.
- [69] C. M. Hurley, A. E. Anker, M. G. Frank, D. Matsumoto, ve H. C. Hwang, “Background factors predicting accuracy and improvement in micro expression recognition”, *Motiv. Emot.*, c. 38, sayı 5, ss. 700–714, 2014.
- [70] S. Porter, L. Ten Brinke, ve B. Wallace, “Secrets and lies: Involuntary leakage in deceptive facial expressions as a function of emotional intensity”, *J. Nonverbal Behav.*, c. 36, sayı 1, ss. 23–37, 2012.
- [71] K. McDonald, I. R. Newby-Clark, J. Walker, ve K. Henselwood, “It is written all over your face: Socially rejected people display microexpressions that are detectable after training in the Micro Expression Training Tool (METT)”, *Eur. J. Soc. Psychol.*, c. 48, sayı 2, ss. O175–O181, 2018.
- [72] W.-J. Yan vd., “CASME II: An improved spontaneous micro-expression database and the baseline evaluation”, *PLoS One*, c. 9, sayı 1, s. e86041, 2014.
- [73] A. K. Davison, C. Lansley, N. Costen, K. Tan, ve M. H. Yap, “SAMM: A Spontaneous Micro-Facial Movement Dataset”, *IEEE Trans. Affect. Comput.*, c. 9, sayı 1, ss. 116–129, 2018, doi: 10.1109/TAFFC.2016.2573832.
- [74] X. Li, T. Pfister, X. Huang, G. Zhao, ve M. Pietikäinen, “A spontaneous micro-expression database: Inducement, collection and baseline”, içinde *2013 10th IEEE International Conference and Workshops on Automatic face and gesture recognition (fg)*, 2013, ss. 1–6.
- [75] Y. Li, X. Huang, ve G. Zhao, “Can Micro-Expression be Recognized Based on Single Apex Frame?”, içinde *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018, ss. 3094–3098, doi: 10.1109/ICIP.2018.8451376.
- [76] S.-T. Liong, J. See, K. Wong, A. C. Le Ngo, Y.-H. Oh, ve R. Phan, “Automatic apex frame spotting in micro-expression database”, içinde *2015 3rd IAPR Asian conference on pattern recognition (ACPR)*, 2015, ss. 665–669.
- [77] Y. S. Gan ve S.-T. Liong, “Bi-directional vectors from apex in cnn for micro-expression recognition”, içinde *2018 IEEE 3rd international conference on image, vision and computing (ICIVC)*, 2018, ss. 168–172.
- [78] G. Farnebäck, “Two-frame motion estimation based on polynomial expansion”, içinde *Scandinavian conference on Image analysis*, 2003, ss. 363–370.
- [79] C. Zach, T. Pock, ve H. Bischof, “A duality based approach for realtime tv-l 1 optical flow”, içinde *Joint pattern recognition symposium*, 2007, ss. 214–223.
- [80] B. Allaert, I. R. Ward, I.-M. Bilasco, C. Djeraba, ve M. Bennamoun, “Optical

flow techniques for facial expression analysis: Performance evaluation and improvements”, 2019.

- [81] B. Allaert, I. M. Bilasco, ve C. Djeraba, “Advanced local motion patterns for macro and micro facial expression recognition”, *arXiv Prepr. arXiv1805.01951*, 2018.
- [82] H. Lee, S. Lee, ve O. Choi, “Improved method on image stitching based on optical flow algorithm”, *Int. J. Eng. Bus. Manag.*, c. 12, s. 1847979020980928, 2020.
- [83] L. Zhou, Q. Mao, ve L. Xue, “Dual-Inception Network for Cross-Database Micro-Expression Recognition”, içinde *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)*, 2019, ss. 1–5, doi: 10.1109/FG.2019.8756579.
- [84] Y. Li, X. Huang, ve G. Zhao, “Joint Local and Global Information Learning With Single Apex Frame Detection for Micro-Expression Recognition”, *IEEE Trans. Image Process.*, c. 30, ss. 249–263, 2020.
- [85] N. Rose, “Facial Expression Classification using Gabor and Log-Gabor Filters”, içinde *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, 2006, ss. 346–350, doi: 10.1109/FGR.2006.49.
- [86] Y.-C. Lee ve C.-H. Chen, “Feature Extraction for Face Recognition Based on Gabor Filters and Two-Dimensional Locality Preserving Projections”, içinde *2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 2009, ss. 106–109, doi: 10.1109/IIH-MSP.2009.210.
- [87] J. Ou, X.-B. Bai, Y. Pei, L. Ma, ve W. Liu, “Automatic Facial Expression Recognition Using Gabor Filter and Expression Analysis”, içinde *2010 Second International Conference on Computer Modeling and Simulation*, 2010, c. 2, ss. 215–218, doi: 10.1109/ICCMS.2010.45.
- [88] T. Gao, X. M. Zhao, T. Chen, Z. W. Liu, ve C. Ni, “Face description based on adaptive local weighted Gabor comprehensive histogram feature”, *Multimed. Tools Appl.*, c. 76, sayı 10, ss. 12893–12916, 2017, doi: 10.1007/s11042-016-3701-y.
- [89] R. Mohammadian Fini, M. Mahlouji, ve A. Shahidinejad, “Performance improvement in face recognition system using optimized Gabor filters”, *Multimed. Tools Appl.*, c. 81, sayı 27, ss. 38375–38408, 2022, doi: 10.1007/s11042-022-13167-6.
- [90] J. Gu vd., “Recent advances in convolutional neural networks”, *Pattern Recognit.*, c. 77, ss. 354–377, 2018, doi: <https://doi.org/10.1016/j.patcog.2017.10.013>.
- [91] A. S. Shamsaldin, P. Fattah, T. A. Rashid, ve N. K. Al-Salihi, “A study of the applications of convolutional neural networks”, *J. Sci. Eng.*, c. 3, ss. 31–39, 2019.
- [92] K. Fukushima ve S. Miyake, “Neocognitron: A self-organizing neural network

- model for a mechanism of visual pattern recognition”, içinde *Competition and cooperation in neural nets*, Springer, 1982, ss. 267–285.
- [93] F. Sultana, A. Sufian, ve P. Dutta, “Advancements in Image Classification using Convolutional Neural Network”, içinde *2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*, 2018, ss. 122–129, doi: 10.1109/ICRCICN.2018.8718718.
- [94] Y. LeCun, L. Bottou, Y. Bengio, ve P. Haffner, “Gradient-based learning applied to document recognition”, *Proc. IEEE*, c. 86, sayı 11, ss. 2278–2324, 1998.
- [95] A. Krizhevsky, I. Sutskever, ve G. E. Hinton, “Imagenet classification with deep convolutional neural networks”, *Adv. Neural Inf. Process. Syst.*, c. 25, ss. 1097–1105, 2012.
- [96] K. Simonyan ve A. Zisserman, “Very deep convolutional networks for large-scale image recognition”, *arXiv Prepr. arXiv1409.1556*, 2014.
- [97] C. Szegedy vd., “Going deeper with convolutions”, içinde *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, ss. 1–9.
- [98] K. He, X. Zhang, S. Ren, ve J. Sun, “Deep residual learning for image recognition”, içinde *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, ss. 770–778.
- [99] K. O’Shea ve R. Nash, “An introduction to convolutional neural networks”, *arXiv Prepr. arXiv1511.08458*, 2015.
- [100] F. DOĞAN ve İ. TÜRKOĞLU, “Derin öğrenme modelleri ve uygulama alanlarına ilişkin bir derleme”, *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Derg.*, c. 10, sayı 2, ss. 409–445, 2019.
- [101] X. Lei, H. Pan, ve X. Huang, “A dilated CNN model for image classification”, *IEEE Access*, c. 7, ss. 124087–124095, 2019.
- [102] E. BAŞARAN, “TİMPANİK MEMBRAN GÖRÜNTÜ ANALİZİ VE YAPAY ZEKA KULLANILARAK SANAL OTİTİS MEDİA TANI SİSTEMİNİN GELİŞTİRİLMESİ”. 2020.
- [103] A.-D. Nguyen, S. Choi, W. Kim, S. Ahn, J. Kim, ve S. Lee, “Distribution Padding in Convolutional Neural Networks”, içinde *2019 IEEE International Conference on Image Processing (ICIP)*, 2019, ss. 4275–4279, doi: 10.1109/ICIP.2019.8803537.
- [104] S. Sharma, S. Sharma, ve A. Athaiya, “Activation functions in neural networks”, *Towar. data Sci.*, c. 6, sayı 12, ss. 310–316, 2017.
- [105] X. Li, Z. Hu, ve X. Huang, “Combine Relu with Tanh”, içinde *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, 2020, c. 1, ss. 51–55, doi: 10.1109/ITNEC48623.2020.9084659.
- [106] M. Saeed, “A gentle introduction to sigmoid function”, 2021.

- [107] N. Aloysius ve M. Geetha, “A review on deep convolutional neural networks”, içinde *2017 International Conference on Communication and Signal Processing (ICCSP)*, 2017, ss. 588–592, doi: 10.1109/ICCSP.2017.8286426.
- [108] O. Sharma, “A New Activation Function for Deep Neural Network”, içinde *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, 2019, ss. 84–86, doi: 10.1109/COMITCon.2019.8862253.
- [109] A. L. Maas, A. Y. Hannun, ve A. Y. Ng, “Rectifier nonlinearities improve neural network acoustic models”, içinde *Proc. icml*, 2013, c. 30, sayı 1, s. 3.
- [110] K. He, X. Zhang, S. Ren, ve J. Sun, “Delving deep into rectifiers: Surpassing human-level performance on imagenet classification”, içinde *Proceedings of the IEEE international conference on computer vision*, 2015, ss. 1026–1034.
- [111] S. Gokulan, S. Narmadha, M. Pavithra, R. Rajmohan, ve T. Ananthkumar, “Determination of Various Deep Learning Parameter for Sleep Disorder”, içinde *2020 International Conference on System, Computation, Automation and Networking (ICSCAN)*, 2020, ss. 1–6, doi: 10.1109/ICSCAN49426.2020.9262331.
- [112] X. Glorot, A. Bordes, ve Y. Bengio, “Deep sparse rectifier neural networks”, içinde *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, 2011, ss. 315–323.
- [113] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, ve R. R. Salakhutdinov, “Improving neural networks by preventing co-adaptation of feature detectors”, *arXiv Prepr. arXiv1207.0580*, 2012.
- [114] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, ve R. Salakhutdinov, “Dropout: a simple way to prevent neural networks from overfitting”, *J. Mach. Learn. Res.*, c. 15, sayı 1, ss. 1929–1958, 2014.
- [115] J. Brownlee, “What is the Difference Between a Batch and an Epoch in a Neural Network”, *Mach. Learn. Mastery*, c. 20, 2018.
- [116] D. R. Wilson ve T. R. Martinez, “The need for small learning rates on large problems”, içinde *IJCNN’01. International Joint Conference on Neural Networks. Proceedings (Cat. No.01CH37222)*, 2001, c. 1, ss. 115–119 c.1, doi: 10.1109/IJCNN.2001.939002.
- [117] R. Zaheer ve H. Shaziya, “A Study of the Optimization Algorithms in Deep Learning”, içinde *2019 Third International Conference on Inventive Systems and Control (ICISC)*, 2019, ss. 536–539, doi: 10.1109/ICISC44355.2019.9036442.
- [118] S. H. Haji ve A. M. Abdulazeez, “Comparison of optimization techniques based on gradient descent algorithm: A review”, *PalArch’s J. Archaeol. Egypt/Egyptology*, c. 18, sayı 4, ss. 2715–2743, 2021.
- [119] S. Ruder, “An overview of gradient descent optimization algorithms”, *arXiv Prepr. arXiv1609.04747*, 2016.

- [120] J. Zhang, “Gradient descent based optimization algorithms for deep learning models training”, *arXiv Prepr. arXiv1903.03614*, 2019.
- [121] R. Sun, “Optimization for deep learning: theory and algorithms”, *arXiv Prepr. arXiv1912.08957*, 2019.
- [122] M. Sarigül ve M. Avci, “Performance comparison of different momentum techniques on deep reinforcement learning”, içinde *2017 IEEE International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*, 2017, ss. 302–306, doi: 10.1109/INISTA.2017.8001175.
- [123] J. Lu, “Gradient Descent, Stochastic Optimization, and Other Tales”, *arXiv Prepr. arXiv2205.00832*, 2022.
- [124] M. D. Zeiler, “Adadelta: an adaptive learning rate method”, *arXiv Prepr. arXiv1212.5701*, 2012.
- [125] D. P. Kingma ve J. Ba, “Adam: A method for stochastic optimization”, *arXiv Prepr. arXiv1412.6980*, 2014.
- [126] J. Yosinski, J. Clune, Y. Bengio, ve H. Lipson, “How transferable are features in deep neural networks?”, *Adv. Neural Inf. Process. Syst.*, c. 27, 2014.
- [127] K. N. Akpınar, S. Genc, ve S. Karagöl, “Chest X-Ray Abnormality Detection Based on SqueezeNet”, içinde *2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, 2020, ss. 1–5, doi: 10.1109/ICECCE49384.2020.9179404.
- [128] K. He ve J. Sun, “Convolutional neural networks at constrained time cost”, içinde *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, ss. 5353–5360.
- [129] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, ve K. Keutzer, “SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size”, *arXiv Prepr. arXiv1602.07360*, 2016.
- [130] F. Chollet, “Xception: Deep learning with depthwise separable convolutions”, içinde *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, ss. 1251–1258.
- [131] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, ve L.-C. Chen, “Mobilenetv2: Inverted residuals and linear bottlenecks”, içinde *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, ss. 4510–4520.
- [132] A. G. Howard vd., “Mobilenets: Efficient convolutional neural networks for mobile vision applications”, *arXiv Prepr. arXiv1704.04861*, 2017.
- [133] M. Tan ve Q. Le, “Efficientnet: Rethinking model scaling for convolutional neural networks”, içinde *International Conference on Machine Learning*, 2019, ss. 6105–6114.
- [134] G. Huang, Z. Liu, L. Van Der Maaten, ve K. Q. Weinberger, “Densely Connected Convolutional Networks”, içinde *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, ss. 2261–2269, doi: 10.1109/CVPR.2017.243.

- [135] K. DONUK, A. R. I. Ali, M. F. ÖZDEMİR, ve D. HANBAY, “Deep Feature Selection for Facial Emotion Recognition Based on BPSO and SVM”, *Politek. Derg.*, s. 1.
- [136] J. Kennedy ve R. Eberhart, “Particle swarm optimization”, içinde *Proceedings of ICNN’95-international conference on neural networks*, 1995, c. 4, ss. 1942–1948.
- [137] A. Robson ve L. Zhang, “Intelligent Facial Expression Recognition Using Particle Swarm Optimization Based Feature Selection”, içinde *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2018, ss. 305–311.
- [138] B. S. G. de Almeida ve V. C. Leite, “Particle swarm optimization: A powerful technique for solving engineering problems”, *Swarm Intell. Adv. New Perspect. Appl.*, 2019.
- [139] J. Kennedy ve R. C. Eberhart, “A discrete binary version of the particle swarm algorithm”, içinde *1997 IEEE International conference on systems, man, and cybernetics. Computational cybernetics and simulation*, 1997, c. 5, ss. 4104–4108.
- [140] M. A. Khanesar, M. Teshnehlab, ve M. A. Shoorehdeli, “A novel binary particle swarm optimization”, içinde *2007 Mediterranean Conference on Control & Automation*, 2007, ss. 1–6, doi: 10.1109/MED.2007.4433821.
- [141] M. Mafarja, R. Jarrar, S. Ahmad, ve A. A. Abusnaina, “Feature selection using binary particle swarm optimization with time varying inertia weight strategies”, içinde *Proceedings of the 2nd International Conference on Future Networks and Distributed Systems*, 2018, ss. 1–9.
- [142] E. Basaran, Z. Cömert, Y. Çelik, Ü. Budak, ve A. Sengür, “Otitis media diagnosis model for tympanic membrane images processed in two-stage processing blocks”, *IOP Sci*, c. 14, ss. 1–27, 2020.
- [143] M. S. Zulfiker, N. Kabir, A. A. Biswas, T. Nazneen, ve M. S. Uddin, “An in-depth analysis of machine learning approaches to predict depression”, *Curr. Res. Behav. Sci.*, c. 2, s. 100044, 2021, doi: <https://doi.org/10.1016/j.crbha.2021.100044>.
- [144] P. M. Granitto, C. Furlanello, F. Biasioli, ve F. Gasperi, “Recursive feature elimination with random forest for PTR-MS analysis of agroindustrial products”, *Chemom. Intell. Lab. Syst.*, c. 83, sayı 2, ss. 83–90, 2006, doi: <https://doi.org/10.1016/j.chemolab.2006.01.007>.
- [145] B. Richhariya, M. Tanveer, ve A. H. Rashid, “Diagnosis of Alzheimer’s disease using universum support vector machine based recursive feature elimination (USVM-RFE)”, *Biomed. Signal Process. Control*, c. 59, s. 101903, 2020, doi: <https://doi.org/10.1016/j.bspc.2020.101903>.
- [146] X. Chen ve J. C. Jeong, “Enhanced recursive feature elimination”, içinde *Sixth International Conference on Machine Learning and Applications (ICMLA 2007)*, 2007, ss. 429–435, doi: 10.1109/ICMLA.2007.35.
- [147] A. Z. Mustaqim, S. Adi, Y. Pristyanto, ve Y. Astuti, “The Effect of Recursive

- Feature Elimination with Cross-Validation (RFECV) Feature Selection Algorithm toward Classifier Performance on Credit Card Fraud Detection”, içinde *2021 International Conference on Artificial Intelligence and Computer Science Technology (ICAICST)*, 2021, ss. 270–275, doi: 10.1109/ICAICST53116.2021.9497842.
- [148] C. Wang, Z. Xiao, ve J. Wu, “Functional connectivity-based classification of autism and control using SVM-RFECV on rs-fMRI data”, *Phys. Medica*, c. 65, ss. 99–105, 2019, doi: <https://doi.org/10.1016/j.ejmp.2019.08.010>.
- [149] V. Vapnik, “The support vector method of function estimation”, içinde *Nonlinear modeling*, Springer, 1998, ss. 55–85.
- [150] Ö. Tonkal, H. Polat, E. Başaran, Z. Cömert, ve R. Kocaoğlu, “Machine Learning Approach Equipped with Neighbourhood Component Analysis for DDoS Attack Detection in Software-Defined Networking”, *Electronics*, c. 10, sayı 11, s. 1227, 2021.
- [151] G. Bassma ve S. Tayeb, “Support vector machines for improving vehicle localization in urban canyons”, içinde *MATEC Web of Conferences*, 2018, c. 200, s. 4.
- [152] G. Mercier ve M. Lennon, “Support vector machines for hyperspectral image classification with spectral-based kernels”, içinde *IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No.03CH37477)*, 2003, c. 1, ss. 288–290 c.1, doi: 10.1109/IGARSS.2003.1293752.
- [153] B. Schölkopf, “The kernel trick for distances”, *Adv. Neural Inf. Process. Syst.*, c. 13, 2000.
- [154] J. D. Rodriguez, A. Perez, ve J. A. Lozano, “Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation”, *IEEE Trans. Pattern Anal. Mach. Intell.*, c. 32, sayı 3, ss. 569–575, 2010, doi: 10.1109/TPAMI.2009.187.
- [155] T.-T. Wong ve P.-Y. Yeh, “Reliable Accuracy Estimates from k-Fold Cross Validation”, *IEEE Trans. Knowl. Data Eng.*, c. 32, sayı 8, ss. 1586–1594, 2020, doi: 10.1109/TKDE.2019.2912815.
- [156] T. Fushiki, “Estimation of prediction error by using K-fold cross-validation”, *Stat. Comput.*, c. 21, sayı 2, ss. 137–146, 2011, doi: 10.1007/s11222-009-9153-8.
- [157] D. Anguita, L. Ghelardoni, A. Ghio, L. Oneto, ve S. Ridella, “The ‘K’ in K-fold Cross Validation.”, içinde *ESANN*, 2012, ss. 441–446.
- [158] L. Ma ve K. Khorasani, “Facial expression recognition using constructive feedforward neural networks”, *IEEE Trans. Syst. Man, Cybern. Part B*, c. 34, sayı 3, ss. 1588–1595, 2004, doi: 10.1109/TSMCB.2004.825930.
- [159] P. Tarnowski, M. Kołodziej, A. Majkowski, ve R. J. Rak, “Emotion recognition using facial expressions”, *Procedia Comput. Sci.*, c. 108, ss. 1175–1184, 2017, doi: <https://doi.org/10.1016/j.procs.2017.05.025>.

- [160] W. Merghani, A. K. Davison, ve M. H. Yap, “A review on facial micro-expressions analysis: datasets, features and metrics”, *arXiv Prepr. arXiv1805.02397*, 2018.
- [161] R. Susmaga, “Confusion matrix visualization”, içinde *Intelligent Information Processing and Web Mining: Proceedings of the International IIS: IIPWM '04 Conference held in Zakopane, Poland, May 17–20, 2004*, 2004, ss. 107–116.
- [162] V. Mayya, R. M. Pai, ve M. M. Manohara Pai, “Automatic Facial Expression Recognition Using DCNN”, *Procedia Comput. Sci.*, c. 93, ss. 453–461, 2016, doi: <https://doi.org/10.1016/j.procs.2016.07.233>.
- [163] E. Başaran, Z. Cömert, ve Y. Çelik, “Convolutional neural network approach for automatic tympanic membrane detection and classification”, *Biomed. Signal Process. Control*, c. 56, s. 101734, 2020.
- [164] X. Zhang, X. Li, Y. Feng, ve Z. Liu, “The use of ROC and AUC in the validation of objective image fusion evaluation metrics”, *Signal Processing*, c. 115, ss. 38–48, 2015, doi: <https://doi.org/10.1016/j.sigpro.2015.03.007>.
- [165] Z. H. Hoo, J. Candlish, ve D. Teare, “What is an ROC curve?”, *Emergency Medicine Journal*, c. 34, sayı 6. BMJ Publishing Group Ltd and the British Association for Accident ..., ss. 357–359, 2017.
- [166] L. Gonçalves, A. Subtil, M. R. Oliveira, ve P. de Zea Bermudez, “ROC curve estimation: An overview”, *REVSTAT-Statistical J.*, c. 12, sayı 1, ss. 1–20, 2014.
- [167] Z. Li, L. Xie, ve G. Wang, “Deep learning features in facial identification and the likelihood ratio bound”, *Forensic Sci. Int.*, c. 344, s. 111576, 2023, doi: <https://doi.org/10.1016/j.forsciint.2023.111576>.
- [168] J. N. Mandrekar, “Receiver operating characteristic curve in diagnostic test assessment”, *J. Thorac. Oncol.*, c. 5, sayı 9, ss. 1315–1316, 2010.
- [169] J. Y. Verbakel vd., “ROC curves for clinical prediction models part 1. ROC plots showed no added value above the AUC when evaluating the performance of clinical prediction models”, *J. Clin. Epidemiol.*, c. 126, ss. 207–216, 2020, doi: <https://doi.org/10.1016/j.jclinepi.2020.01.028>.
- [170] S. H. Park, J. M. Goo, ve C.-H. Jo, “Receiver operating characteristic (ROC) curve: practical review for radiologists”, *Korean J. Radiol.*, c. 5, sayı 1, ss. 11–18, 2004.
- [171] T. Lu, B. Han, ve F. Yu, “Detection and classification of marine mammal sounds using AlexNet with transfer learning”, *Ecol. Inform.*, c. 62, s. 101277, 2021, doi: <https://doi.org/10.1016/j.ecoinf.2021.101277>.
- [172] T. Lu, F. Yu, C. Xue, ve B. Han, “Identification, classification, and quantification of three physical mechanisms in oil-in-water emulsions using AlexNet with transfer learning”, *J. Food Eng.*, c. 288, s. 110220, 2021, doi: <https://doi.org/10.1016/j.jfoodeng.2020.110220>.
- [173] I. Singh, G. Goyal, ve A. Chandel, “AlexNet architecture based convolutional neural network for toxic comments classification”, *J. King Saud Univ.* -

- Comput. Inf. Sci.*, c. 34, sayı 9, ss. 7547–7558, 2022, doi: <https://doi.org/10.1016/j.jksuci.2022.06.007>.
- [174] S. Lu, Z. Lu, ve Y.-D. Zhang, “Pathological brain detection based on AlexNet and transfer learning”, *J. Comput. Sci.*, c. 30, ss. 41–47, 2019, doi: <https://doi.org/10.1016/j.jocs.2018.11.008>.
- [175] M. Juneja ve S. K. Nagar, “Particle swarm optimization algorithm and its parameters: A review”, içinde *2016 International Conference on Control, Computing, Communication and Materials (ICCCCM)*, 2016, ss. 1–5, doi: 10.1109/ICCCCM.2016.7918233.
- [176] İ. YükselÇelik Alper TalhaKaradeniz, “Son Üç Yılda Geliştirilen Metasezgisel Algoritmalar Hakkında Kısa Bir İnceleme”, *Journal*, ss. 463–477, 2019.
- [177] I. P. Adegun ve H. B. Vadapalli, “Facial micro-expression recognition: A machine learning approach”, *Sci. African*, c. 8, s. e00465, 2020, doi: <https://doi.org/10.1016/j.sciaf.2020.e00465>.
- [178] N. Peng, Y. Zhang, ve Y. Zhao, “A SVM-kNN method for quasar-star classification”, *Sci. China Physics, Mech. Astron.*, c. 56, sayı 6, ss. 1227–1234, 2013, doi: 10.1007/s11433-013-5083-8.
- [179] A. R. Tenpe ve A. Patel, “Utilization of Support Vector Models and Gene Expression Programming for Soil Strength Modeling”, *Arab. J. Sci. Eng.*, c. 45, sayı 5, ss. 4301–4319, 2020, doi: 10.1007/s13369-020-04441-6.
- [180] E. BAŞARAN ve Y. ÇELİK, “Skin Cancer Diagnosis Using CNN Features with Genetic Algorithm and Particle Swarm Optimization Methods”, 2022.
- [181] L. Jiang vd., “Machine Learning Based on Diffusion Kurtosis Imaging Histogram Parameters for Glioma Grading”, *Journal of Clinical Medicine*, c. 11, sayı 9, 2022, doi: 10.3390/jcm11092310.
- [182] J. P. Appadurai ve R. Bhargavi, “Eye movement feature set and predictive model for dyslexia: feature set and predictive model for dyslexia”, *Int. J. Cogn. Informatics Nat. Intell.*, c. 15, sayı 4, ss. 1–22, 2021.
- [183] S.-T. Liong vd., “Evaluation of the Spatio-Temporal Features and GAN for Micro-Expression Recognition System”, *J. Signal Process. Syst.*, c. 92, sayı 7, ss. 705–725, 2020, doi: 10.1007/s11265-020-01523-4.
- [184] B. Allaert, I. R. Ward, I. M. Bilasco, C. Djeraba, ve M. Bennamoun, “A comparative study on optical flow for facial expression analysis”, *Neurocomputing*, c. 500, ss. 434–448, 2022, doi: <https://doi.org/10.1016/j.neucom.2022.05.077>.
- [185] Y. S. Gan, J. See, H.-Q. Khor, K.-H. Liu, ve S.-T. Liong, “Needle in a Haystack: Spotting and recognising micro-expressions ‘in the wild’”, *Neurocomputing*, c. 503, ss. 283–298, 2022, doi: <https://doi.org/10.1016/j.neucom.2022.06.101>.
- [186] S. Thuseethan, S. Rajasegarar, ve J. Yearwood, “Deep3DCANN: A Deep 3DCNN-ANN Framework for Spontaneous Micro-expression Recognition”, *Inf. Sci. (Ny)*, 2022, doi: <https://doi.org/10.1016/j.ins.2022.11.113>.

- [187] M.-X. Sun, S.-T. Liong, K.-H. Liu, ve Q.-Q. Wu, “The heterogeneous ensemble of deep forest and deep neural networks for micro-expressions recognition”, *Appl. Intell.*, c. 52, sayı 14, ss. 16621–16639, 2022, doi: 10.1007/s10489-022-03284-y.



ÖZGEÇMİŞ

Mehmet Zahit UZUN 2004 yılında Süleyman Demirel Üniversitesi Teknik Eğitim Fakültesi Bilgisayar ve Kontrol Öğretmenliği Bölümü'nde öğrenime başlayıp 2008 yılında mezun oldu. 2010 yılında Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Elektronik Bilgisayar Eğitimi Anabilim Dalında yüksek lisans eğitimine başladı ve 2013 yılında mezun oldu. 2016 yılında Karabük Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalında doktora eğitimine başladı. 2008-2010 eğitim öğretim yılında MEB'de ücretli olarak, 2009-2011 yılları arasında ise Konya Meslek Edindirme Kurslarında (KOMEK) bilgisayar öğretmeliği yaptı. 2011 yılında Karamanoğlu Mehmetbey Üniversitesi Ermenek Meslek Yüksek Okulunda öğretim görevlisi olarak göreve başladı ve halen aynı kurumda görevine devam etmektedir.