

TC.
İSTANBUL KÜLTÜR ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

EEG VE YÜZ İFADELERİNE DAYALI ÇOK MODLU DUYGU TANIMA

YÜKSEK LİSANS TEZİ

Songül Erdem GÜLER

2100008117

Anabilim Dalı: Bilgisayar Mühendisliği

Programı: Bilgisayar Mühendisliği

Tez Danışmanı: Doç. Dr. Fatma Patlar AKBULUT

HAZİRAN 2024

TC.
İSTANBUL KÜLTÜR ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

EEG VE YÜZ İFADELERİNE DAYALI ÇOK MODLU DUYGU TANIMA

YÜKSEK LİSANS TEZİ

Songül Erdem GÜLER

2100008117

Anabilim Dalı: Bilgisayar Mühendisliği

Programı: Bilgisayar Mühendisliği

Tez Danışmanı: Doç. Dr. Fatma Patlar AKBULUT

Jüri Üyeleri: Doç. Dr. Fatma Patlar AKBULUT

Dr. Öğr Üyesi Öznur Şengel

Doç. Dr. Can Eyüpoğlu

HAZİRAN 2024

ÖNSÖZ

“EEG ve Yüz İfadelerine Dayalı Çok Modlu Duygu Tanıma” başlıklı tez çalışmamda akademik bilgi birikimi ve tecrübesiyle tezimin her aşamasında bana yol gösteren, teoriden pratiğe tüm süreçlerde araştırmalarımı derinleştirmem için gereken desteği sağlayan ve ilham veren değerli tez danışmanım Doç. Dr. Fatma Patlar AKBULUT’a çok teşekkür ederim.

Ayrıca, başta kıymetli eşim Ali Kerem Güler olmak üzere, sabır ve destekleriyle her zaman yanımda olan aileme ve tez sürecim boyunca bana esneklik sağlayarak akademik yolculuğuma destek olan şirketim ve çalışma arkadaşlarıma en içten teşekkürlerimi sunuyorum. Onların destekleri sayesinde bu zorlu süreç daha keyifli ve kolay hale geldi. Her birinin katkısı, bu çalışmanın sadece akademik bir çaba değil, aynı zamanda bir yaşam deneyimi olmasını sağladı. Bu nedenle, tezimi tamamlamamda yardımcı olan herkese yürekten teşekkür ediyorum.

Songül Erdem GÜLER

HAZİRAN 2024

İÇİNDEKİLER

ÖNSÖZ	iii
İÇİNDEKİLER.....	iv
KISALTMALAR	v
ŞEKİL LİSTESİ	vi
TABLO LİSTESİ	vii
SÖZDE KOD LİSTESİ	vii
ÖZET.....	viii
ABSTRACT	ix
1. GİRİŞ.....	1
1.1. Tezin Organizasyonu	4
2. İLGİLİ ÇALIŞMALAR.....	5
3. METODOLOJİ.....	11
3.1 Veri Seti	11
3.2 Veri Ön İşleme.....	15
3.2.1 EEG Sinyalleri ile Duygu Sınıflandırma	15
3.2.2 Çok Modlu Duygu Sınıflandırma: Yüz İfadeleri ve EEG Sinyalleri	21
3.3 Model Eğitimi	29
3.3.1 GRU	29
3.3.2 LSTM.....	31
3.3.3 Transformer	33
3.4 Müzik Öneri Sistemi	38
3.4.1 Veri İşleme	39
3.4.2 Öneri Sistemi Modelleme.....	44
4. BULGULAR VE TARTIŞMA	49
4.1 EEG Sinyalleri ile Duygu Sınıflandırma Sonuçları	49
4.2 Hem EEG Sinyalleri Hem Yüz İfadeleri ile Duygu Sınıflandırma Sonuçları	62
4.3 Genel Değerlendirme ve Genel Mimari.....	74
4.4 Öneri Sistemi Sonuçları	77
5. SONUÇ VE ÖNERİLER.....	79
6. KAYNAKÇA	81

KISALTMALAR

EEG: Electroencephalogram (Elektroensefalografi)

FER: Facial Expression Recognition (Yüz İfadesi Tanıma)

FPS: Frame Per Second (Saniyedeki Kare Sayısı)

GRU: Gated Recurrent Unit (Kapılı Yinelemeli Birim)

LSTM: Long Short-Term Memory (Uzun Kısa Süreli Hafıza)

KNN: K-Nearest Neighbors (K- En Yakın Komşu)

DEAP: A Database for Emotion Analysis using Physiological Signals (Fizyolojik Sinyaller Kullanılarak Duygu Analizi için Bir Veritabanı)

HVHA: High Valence, High Arousal (Yüksek Değerlik, Yüksek Uyarılma)

HVLA: High Valence, Low Arousal (Yüksek Değerlik, Düşük Uyarılma)

LVHA: Low Valence, High Arousal (Düşük Değerlik, Yüksek Uyarılma)

LVLA: Low Valence, Low Arousal (Düşük Değerlik, Düşük Uyarılma)

ŞEKİL LİSTESİ

Şekil 3.1 DEAP veri seti katılımcısından deney esnasında alınan görsel	12
Şekil 3.2 Elektrot yerleşimi ve kanalların gösterimi	12
Şekil 3.3 İki boyutlu uyarılma-değerlik alanı	13
Şekil 3.4 DEAP veri setinde kullanılan videoların uyarılma-değerlik ekseninde dağılımı 14	
Şekil 3.5 Katılımcı 11'e ait eeg sinyali görselleştirmesi, karşılık gelen yüz görüntüsü, duygu etiketi ve yüz ifadesi tanıma sonuçları	18
Şekil 3.6 Russel'in duygu ölçeğine göre veri setinin dağılımları.....	19
Şekil 4.1 Tek modlu LSTM modeline ait doğruluk(accuracy) grafiği	49
Şekil 4.2 Tek modlu LSTM modeline ait kayıp(loss) grafiği	50
Şekil 4.3 Tek modlu LSTM modeline ait karışıklık matrisi	51
Şekil 4.4 Tek modlu Transformer modeline ait doğruluk grafiği.....	54
Şekil 4.5 Tek modlu Transformer modeline ait kayıp grafiği	55
Şekil 4.6 Tek modlu Transformer modeline ait karışıklık matrisi	56
Şekil 4.7 Tek modlu GRU modeline ait doğruluk grafiği	58
Şekil 4.8 Tek modlu GRU modeline ait kayıp grafiği.....	59
Şekil 4.9 Tek modlu GRU modeline ait karışıklık matrisi	60
Şekil 4.10 Çok modlu LSTM modeline ait doğruluk grafiği	62
Şekil 4.11 Çok modlu LSTM modeline ait kayıp grafiği.....	63
Şekil 4.12 Çok modlu LSTM modeline ait karışıklık matrisi	64
Şekil 4.13 Çok modlu Transformer modeline ait doğruluk grafiği	66
Şekil 4.14 Çok modlu Transformer modeline ait kayıp grafiği	67
Şekil 4.15 Çok modlu Transformer modeline ait karışıklık matrisi.....	68
Şekil 4.16 Çok modlu GRU modeline ait doğruluk grafiği.....	70
Şekil 4.17 Çok modlu GRU modeline ait kayıp grafiği	71
Şekil 4.18 Çok modlu GRU modeline ait karışıklık matrisi	72
Şekil 4.19 Sistemin genel mimarisi	76
Şekil 4.20 Öneri sisteminin test kullanıcısı için yaptığı öneriler	77

TABLO LİSTESİ

Tablo 3.1 Eğitim, doğrulama ve test kümelerindeki veri sayısı	20
Tablo 3.2 Transformer modelinin hiperparametreleri ve tanımlamaları.....	36
Tablo 3.3 Duygu durumları ve etiketler ile yapılan ön işlemler sonucu oluşan dataframe	42
Tablo 3.4 Participant ratings ve video list dosyalarından gelen bilgilerin de eklenmesiyle oluşan dataframe.....	42
Tablo 3.5 Tüm ön işleme adımları sonrasında oluşan nihai dataframe.....	43
Tablo 4.1 Tek modlu LSTM modeline ait sınıflandırma raporu.....	52
Tablo 4.2 Tek modlu Transformer modeline ait sınıflandırma raporu	57
Tablo 4.3 Tek modlu GRU modeline ait sınıflandırma raporu.....	61
Tablo 4.4 Çok modlu LSTM modeline ait sınıflandırma raporu.....	65
Tablo 4.5 Çok modlu Transformer modeline ait sınıflandırma raporu	69
Tablo 4.6 Çok modlu GRU modeline ait sınıflandırma raporu	73
Tablo 4.7 Modellerin genel doğruluk oranları.....	74
Tablo 4.8 Modellerin genel kayıp performansları	75

SÖZDE KOD LİSTESİ

Sözde Kod 3.1 EEG veri işleme	16
Sözde Kod 3.2 Etiket dönüşümünü yapan fonksiyon.....	17
Sözde Kod 3.3 Haar Cascade ve MTCNN Algoritmalarının Ürettiği Veri Setlerini Karşılaştırarak Daha Geniş Veri Setini Kayıt Dizinine Kaydeden Fonksiyon	24
Sözde Kod 3.4 Sentetik veri üreten fonksiyon	26
Sözde Kod 3.5 Alt örnekleme faktörünün hesaplanması, alt örnekleme işlemi, EEG ve FER verilerinin birleştirilmesi.....	28
Sözde Kod 3.6 Duygu durumları ve Etiketlerin kullanımı ile yeni dataframe oluşturan fonksiyon	40
Sözde Kod 3.7 Veri yükleme ve ön işleme fonksiyonu	46
Sözde Kod 3.8 Kullanıcılar arası benzerlik hesaplayan fonksiyon	47
Sözde Kod 3.9 Öneri Yapma Fonksiyonu	48

Üniversite	:	İstanbul Kültür Üniversitesi
Enstitü	:	Lisansüstü Eğitim Enstitüsü
Anabilim Dalı	:	Bilgisayar Mühendisliği
Programı	:	Bilgisayar Mühendisliği
Tez Danışmanı	:	Doç. Dr. Fatma Patlar AKBULUT
Tez Türü ve Tarihi	:	Yüksek Lisans – Haziran 2024

ÖZET

EEG VE YÜZ İFADELERİNE DAYALI ÇOK MODLU DUYGU TANIMA

Songül Erdem GÜLER

Duygu tanıma alanındaki ilerlemelere rağmen, araştırma alanı hala iki temel sınırlamayla karşı karşıyadır: giderek daha karmaşık hesaplamalar yapmak için derin modellerin kullanımı ve çeşitli veri türlerinin bir karışımıyla duyguların tanımlanması. Bu çalışma, elektroensefalografi (EEG) sinyallerini yüz ifadeleriyle birleştirerek, Transformer, LSTM ve GRU gibi gelişmiş modelleri içeren multimodal duygu tanıma konusundaki bilgi birikimine katkıda bulunmaktadır. Sonuçlar, önerilen yaklaşımın etkinliğini doğrulamaktadır, çünkü GRU modelinin sırasıyla tek modlu (yalnızca EEG) ve çok modlu (EEG ve yüz ifadeleri) veri setlerinde %88 ve %97 gibi yüksek tahmin doğruluğu göstermektedir. Bulgular, geleneksel tek modlu tekniklerine kıyasla, çok modlu yaklaşımların dikkate değer ilerlemeler sağladığını göstermektedir. Çalışma, karmaşık sinirsel dinamikleri ve görünür duygusal ipuçlarını yakalayan, duygu tanıma sistemlerinin sağlamlığını ve doğruluğunu artıran bir çerçeve sunmaktadır. Bu sonuçlar, çeşitli veri kaynaklarını gelişmiş modellerle entegre ederek tek modlu sistemlerin yarattığı sınırlamanın nasıl aşılabileceğini göstermeleri açısından önemli pratik çıkarımlara sahiptir.

Anahtar Kelimeler : İnsan Bilgisayar Etkileşimi, Duygu Tanıma, Derin Öğrenme, EEG Sinyalleri, Yüz İfadeleri

University : **İstanbul Kültür University**
Institute : **Institute of Graduate Studies**
Branch : **Computer Engineering**
Program : **Computer Engineering**
Thesis Advisor : **Assoc. Prof. Dr. Fatma Patlar AKBULUT**
Thesis Type and Date : **Master Degree – June 2024**

ABSTRACT

Multimodal Emotion Recognition: Emotion Classification through the Integration of EEG and Facial Expressions

Songül Erdem GÜLER

Despite the advancements made in emotion recognition, the need for deep learning to carry out more complex calculations and identifying emotions across different modalities remain two major challenges. This study integrates facial expressions with electroencephalography (EEG) signals, adding to the body of knowledge on multimodal emotion recognition. An extensive analysis of the standalone EEG signals and EEG signals combined with facial expression data performance of emotion recognition models GRU, LSTM, and Transformer. The results show the efficacy of this method since they show the high accuracy of the GRU model with predictions of 88% and 97% on unimodal (EEG only) and multimodal (EEG and facial expressions) datasets, respectively. The results show that multimodal approaches produce significant improvements, unlike conventional unimodal methods. Using models including GRU, LSTM, and the Transformer, together with the complementary nature of EEG signals and facial expressions, the study offers a thorough framework that captures complex neural dynamics and visible emotional cues, enhancing the robustness and accuracy of emotion recognition systems. These results have important practical ramifications since they show how to get beyond the drawbacks of single-modality systems by combining different data sources with sophisticated models.

Keywords: Human Computer Interaction, Emotion Recognition, Deep Learning, EEG Signals, Facial Expressions

1. GİRİŞ

Duygular, yaşam üzerindeki çok yönlü etkileri nedeniyle, dünyayı nasıl algıladığımız ve başkalarıyla nasıl iletişim kurduğumuz üzerinde önemli bir etkiye sahiptir. Duyguların karmaşık etkileşimi sadece bireysel yaşamımıza anlam katmakla kalmaz, aynı zamanda çevremizi algılayışımızı da etkiler. Ayrıca kişisel etkileşimleri şekillendirdiği ve çevresel bakış açılarımızı etkilediği için karar verme süreçlerimizi, hafızamızı ve kişiler arası bağlantılarımızı etkilemede önemli bir rol oynar. Bu önemli etki, insan-bilgisayar etkileşimi gibi sürekli gelişen ve karmaşık alanlarda duyguların anlaşılması ve kullanılmasının kritik önemini artırır. Teknoloji hızla ilerledikçe, duygusal durumlarımızı algılayan ve bunlara tepki veren etkileşimli sistemlerin geliştiğine tanık oluyoruz (Houssein v.d.). Teknoloji ve duygunun bu dikkat çekici birleşimi, insan-bilgisayar etkileşimini dönüştürmek için önemli fırsatlar sunar. Bu çerçevede, duygusal durumlarımıza uyum sağlayan, içsel durumlarımıza göre kişiselleştirilmiş yanıtlar veren ve dijital dünya ile insan dünyasını sorunsuz bir şekilde entegre eden makineler, daha anlamlı etkileşimler geliştirmek için oldukça etkili araçlar olma potansiyeline sahiptir.

Duygu tanıma teknolojileri, insan-makine etkileşimi, tıbbi teşhis, eğitim ve güvenlik gibi alanlarda önemli ilerlemeler kaydetmiştir. Bu alanda sıklıkla kullanılan elektroensefalografi (EEG) sinyalleri, insan beyninin durumunu karakterize etme konusundaki güçlü kapasitesi sayesinde önemli bir konuma sahiptir. EEG sinyalleri, beyindeki elektriksel aktivite kalıplarını yüksek zamansal hassasiyetle yakalayarak, duygusal tepkilerin altında yatan sinir dinamiklerine doğrudan bir pencere sağlar (Pessoa v.d.; Akbulut v.d.; Dogan v.d.). Bu yöntem, çeşitli duygusal durumlarla ilişkili farklı elektrokortikal belirteçlerin tanımlanmasına olanak tanıyarak biliş ve duygu arasındaki karmaşık etkileşimin incelikli bir şekilde anlaşılmasını sağlar (Huang v.d.; Derdiyok v.d.).

Duygu tanıma karmaşık bir süreçtir. Tek modlu yaklaşımlar, duygu tanıma süreçlerinde yalnızca bir tür veri kaynağı kullanırlar. Bu sistemler, genellikle yüz ifadeleri, ses analizi veya yazılı metin gibi tek bir modaliteye odaklanırlar. Tek modlu yaklaşımların temel avantajı, spesifik veri türü üzerine yoğunlaşarak duyguları tanıma sürecindeki basitlik ve verimliliğidir. Ancak, bu sistemler karmaşık insan duygularını anlama konusunda sınırlı kalabilirler çünkü insan duygularının ifadesi çok boyutlu ve çeşitli yollarla gerçekleşebilir.

Çok modlu duygu tanıma sistemlerinde ise insan duygularını algılamak için birden fazla veri kaynağı entegre edilerek kullanılır. Bu sistemler, genellikle EEG sinyalleri, yüz ifadeleri, ses tonu ve vücut dili gibi farklı modalitelerden gelen verileri birleştirir. Bu çok yönlü yaklaşım, her bir veri kaynağının kendi başına sağlayabileceğinden daha zengin ve doğru duygu tanıma imkânı sunar. Çok modlu sistemler, makine öğrenmesi ve derin öğrenme algoritmalarını kullanarak, entegre edilen veri setlerinden duygusal durumların sınıflandırılmasını sağlar. Bu sistemler, sanal asistanlar, akıllı ev cihazları gibi insan-makine etkileşimini geliştiren uygulamalar için önemli bir temel oluşturur ve kullanıcı deneyimini zenginleştirir.

Bu çalışmanın odak noktası, özellikle EEG sinyalleri ve yüz ifadeleri kullanılarak geliştirilen çok modlu duygu tanıma sistemleridir. Elektroensefalografi (EEG) sinyalleri, insan beyninin elektriksel aktivitesini kaydeden bir tekniktir. EEG, beyin dalgalarının zamanla değişimlerini izleyerek, çeşitli beyin durumları ve işlevleri hakkında bilgi sağlar. EEG sinyalleri, duygu tanıma alanında, insanların duygusal durumlarını objektif bir şekilde ölçmek ve analiz etmek için kullanılmaya başlanmıştır. Duygu tanımada EEG'nin kullanımı, 2000'li yıllardan itibaren ilgi görmeye başlamıştır. Bu yaklaşımda, EEG sinyallerindeki belirli desenlerin, farklı duygusal durumlarla ilişkilendirilebileceği düşünülmektedir. Örneğin, mutluluk, üzüntü, korku gibi duyguların, EEG sinyallerinde farklı desenler oluşturduğu ve bu desenlerin makine öğrenmesi algoritmaları kullanılarak sınıflandırılabilceği gösterilmiştir. Duygu tanıma için EEG sinyallerinin kullanılması, yüksek doğruluk oranlarına ulaşabilme potansiyeli sunarken, duygusal durumların daha objektif bir şekilde değerlendirilmesine olanak tanır. Bu teknik, insan-makine etkileşimi, psikolojik araştırmalar ve tıbbi teşhis gibi alanlarda uygulanmaktadır.

Yüz ifadeleri ise içsel duygusal durumların dışsal bir tezahürü olarak hizmet eder ve duygu tanıma sistemleri için değerli ipuçları sağlar. İnsan yüzü, çok sayıda duygusal bilgi aktaran mikro ifadelerden, kas hareketlerinden ve ince değişikliklerden oluşan bir tuvaldir. Duygu tanımada EEG sinyalleri ve yüz ifadeleri arasındaki sinerji, bu yöntemlerin tamamlayıcı doğasından kaynaklanmaktadır. EEG sinyalleri sinirsel tepkilerin inceliklerini açığa çıkarırken, yüz ifadeleri duygusal ipuçlarının görünür ve yorumlanabilir bir katmanını sunar. Bu yöntemlerin birleştirilmesi, duygu sınıflandırmasının derinliğini ve doğruluğunu zenginleştiren bütünsel bir yaklaşım sağlar (Leong v.d. ; Zhang v.d.).

Yüz ifadelerinin analizi, duygu tanıma teknolojilerinin geliştirilmesinde EEG sinyalleriyle birlikte kullanılarak, çok modlu bir yaklaşımın temelini oluşturur. Bu çok modlu sistemler, duygu tanıma doğruluğunu artırarak, daha doğal ve etkili insan-makine etkileşimine katkıda bulunur. Çok modlu yaklaşımlar, duygu tanıma doğruluğunu önemli ölçüde artırarak, tek modlu sistemlerin sınırlılıklarını aşma potansiyeline sahiptir. Bu tezin temel amacı, çok modlu duygu tanıma sistemlerinin geliştirilmesi ve bunların insan davranışını anlama ve yorumlama kapasitesini artırmadaki rolünün incelenmesidir. Bu sistemler, özellikle insan-makine etkileşimi, eğitim, sağlık hizmetleri ve güvenlik gibi çeşitli alanlarda uygulama potansiyeline sahiptir. Ancak, duyguların karmaşık ve çok boyutlu doğası, bu sistemlerin geliştirilmesini zorlaştırmaktadır. Tek modlu sistemler, duyguları tanıma konusunda belirli bir başarı düzeyine ulaşmış olsalar da genellikle belirli sınırlılıklar ve doğrulukta eksiklikler göstermektedir. Çok modlu duygu tanıma sistemleri, farklı veri kaynaklarından (örneğin, EEG sinyalleri ve yüz ifadeleri) elde edilen bilgileri birleştirerek, bu sınırlılıkların üstesinden gelmeyi vaat etmektedir.

Bu süreçte, çeşitli derin öğrenme modelleri (GRU, LSTM ve Transformer) performansları karşılaştırmalı olarak incelenerek ve bu modellerin tek modlu veri setleri ve çok modlu veri setleri üzerindeki etkileri analiz edilmiştir. Araştırmanın nihai hedefi, duygu tanıma doğruluğunu maksimize eden ve gerçek dünya uygulamalarında etkili olan çok modlu bir duygu tanıma sistemi geliştirmektir. Bu amaç doğrultusunda, çalışmada öncelikle EEG sinyalleri ve yüz ifadeleri gibi farklı modlardan elde edilen verilerin nasıl entegre edilebileceğini ve işlenebileceği analiz edilmiştir. Daha sonra, entegre veri setlerinin duygu tanıma üzerindeki etkisini değerlendirmek için derin öğrenme modelleri kullanılmıştır.

Bu modellerin her birinin avantajları ve sınırlılıkları, tek modlu ve çok modlu veri setleri ile karşılaştırılarak ve çok modlu duygu tanıma sistemlerinin etkinliği üzerine derinlemesine bir analiz yapılmıştır.

Özetle, çalışmanın temel katkıları aşağıdaki gibidir:

- Bu çalışmada, yüz ifadesi verileri ile EEG verileri birleştirilirken ortaya çıkan boyut uyumsuzluğunu ele almak için özel dönüşüm fonksiyonları geliştirilmiş ve detaylı bir veri analizi yapılmıştır.
- Çalışmada, katılımcı video kayıtları MTCNN ve HaarCascade yüz tespit algoritmaları kullanılarak incelenerek duygular tespit edilmiştir. En fazla duyguyu tanımlayan

algoritmanın karşılaştırılmasıyla daha doğru duygu tespiti yapılmış, bu sayede yüz ifadesi verilerinin kullanımı en üst düzeye çıkarılmıştır.

- EEG sinyalleri ve yüz ifadesi verileri birleştirildiğinde, çok modlu duygu tanıma yöntemleri tek modlu yöntemlerden daha iyi performans göstermiştir. Çok modlu GRU modeline ait doğruluk oranının (%97) tek modlu senaryoya (%88) göre daha yüksek olması genel performansın iyileştiğini ve duygusal durumları daha iyi temsil etmek için birden fazla veri kaynağını entegre etmenin değerini göstermiştir.
- Bazı katılımcıların videolarından çıkarılan yüz ifadesi verilerinin yetersiz olduğu durumlarda, boşlukları doldurmak ve veri miktarını ilgili EEG sinyali verileriyle eşleştirmek için sentetik veriler üretilmiştir.
- Literatürde DEAP veri setini kullanan araştırmalar genellikle duyguları iki sınıfa, esas olarak değerlik ve uyarılma olarak sınıflandırır. Ancak bu çalışmada hem değerlik hem de uyarılma eksenleri ile bunların yüksek ve düşük durumlarını içeren dört farklı sınıf için çok sınıflı sınıflandırma yapılmıştır.

Bu kapsamlı araştırma, çok modlu duygu tanıma sistemlerinin anlaşılmasına katkıda bulunmayı ve bu alandaki bilgi birikimini genişletmeyi hedeflemektedir.

1.1. Tezin Organizasyonu

Bölüm 1’de tezin konusu belirtilmiş, literatüre katkısına değinilmiş ve çalışmaya giriş yapılmıştır.

Bölüm 2’de tezin çalışma konusuyla alakalı literatürdeki ilgili çalışmalar gösterilmiştir.

Bölüm 3’te tezde kullanılan metodoloji açıklanmıştır.

Bölüm 4’de bulgular sunulmuştur ve tartışma bölümüne yer verilmiştir.

Bölüm 5’te sonuç ve öneriler yer almaktadır.

Bölüm 6’da çalışmada kullanılan kaynaklar sıralanmıştır.

2. İLGİLİ ÇALIŞMALAR

"DEAP: Fizyolojik Sinyaller Kullanılarak Duygu Analizi İçin Bir Veritabanı" çalışması duygusal bilişim alanında önemli bir katkı sağlamaktadır. Koelstra ve arkadaşlarının yürüttüğü bu araştırma, EEG ve çevresel fizyolojik sinyallerin kaydedilmesiyle oluşturulan kapsamlı bir çok modlu veri seti üzerinden, insanların duygusal tepkilerini derinlemesine analiz etmeyi amaçlamaktadır. Çalışma, 32 katılımcı tarafından izlenen ve her biri farklı duygusal tepkiler uyandıran 40 müzik videosu üzerinden toplanan verileri içermektedir. Katılımcıların her bir videoya ilişkin uyarılma, değerlik (pozitif veya negatif duygusal durum), beğeni, baskınlık ve aşinalık gibi çeşitli boyutlarda değerlendirmeleri yapılmıştır. Araştırmada, EEG sinyal frekansları ile katılımcıların değerlendirmeleri arasındaki korelasyonlar incelenmiş ve EEG ile duygusal durumların sınıflandırılması için metodolojiler geliştirilmiştir. Bu kapsamlı araştırma, duygusal bilişim alanında gelecekte yapılacak çalışmalara temel oluşturacak değerli bir veri setini kamuoyu ile paylaşmıştır (Koelstra v.d.).

"EEG ve Göz İzleme Verileri Kullanılarak Çok Modlu Duygu Tanıma" çalışması, EEG sinyalleri ve göz izleme verilerini birleştirerek duygu tanıma alanında yeni bir yöntem sunmaktadır. Zheng ve arkadaşları tarafından gerçekleştirilen bu araştırma, katılımcıların izledikleri 15 duygusal film klibi sırasında toplanan EEG sinyalleri ve göz takibi verilerinden duyguyla ilgili özellikleri çıkarmış ve bu verileri kullanarak duygu tanıma performansını artırmayı amaçlayan bir füzyon modeli oluşturmuştur. Araştırmanın sonuçları, özellik düzeyinde füzyon ve karar düzeyinde füzyon stratejileri kullanılarak EEG sinyalleri ve göz takibi verilerinin birleştirilmesinin, duygu tanıma modelinin performansını önemli ölçüde artırabileceğini göstermektedir. Bu çok modlu yaklaşım, EEG tabanlı sınıflandırma ve göz takibi verilerine dayalı sınıflandırma için sırasıyla %71.77 ve %58.90 gibi ortalama doğruluk oranları elde etmiştir. Ayrıca, farklı füzyon stratejileri kullanılarak %73.59 ve %72.98 gibi ortalama doğruluk oranlarına ulaşılmıştır. Bu sonuç hem özellik düzeyinde hem de karar düzeyinde füzyonun duygu tanıma modelinin performansını iyileştirebileceğini ortaya koymaktadır (Zheng v.d.).

Liu ve diğerlerinin "Çoklu Modlu Derin Öğrenme Kullanarak Duygu Tanıma" başlıklı çalışması, duygu tanıma konusunda derin öğrenme tekniklerinin çoklu modlu yaklaşımlarla nasıl geliştirilebileceğini araştırmaktadır. Bu çalışma, EEG (elektroensefalogram) sinyalleri ve göz hareketleri gibi farklı modalitelerden elde edilen verilerin bütünleştirilmesiyle duyguları

tanıma yeteneğini artırmayı hedeflemektedir. İki modlu Derin Otoenkoder (BDAE) ağının kullanımıyla, SEED ve DEAP veri setleri üzerinde sırasıyla %91.01 ve %83.25'lik ortalama doğruluk oranları elde edilmiştir. Bu oranlar, alandaki mevcut yaklaşımlara kıyasla oldukça yüksektir. Çalışma, EEG ve göz hareketi verilerinin birbirini tamamlayıcı bilgiler içerdiğini ve BDAE ağının bu tamamlayıcı özelliklerden faydalanarak duygu tanıma doğruluğunu önemli ölçüde artırabileceğini göstermiştir. Bu bulgular, duygu tanıma modellerinin geliştirilmesinde çoklu modlu yaklaşımların önemini vurgulamaktadır (Liu v.d.).

Ma ve diğer yazarlar tarafından yazılan "Çoklu Modlu Residual LSTM Ağı Kullanarak Duygu Tanıma" başlıklı çalışmada ise, EEG ve periferik fizyolojik sinyalleri (PPS) kullanarak duygu tanıma alanında derin öğrenme tabanlı bir çoklu modlu yaklaşım ele alınmıştır. Bu araştırma, uzun kısa süreli hafıza ağları (LSTM) ötesine geçen ve residual öğrenme ile ağırlık paylaşımı mekanizmalarıyla güçlendirilmiş yeni bir Çoklu Modlu Residual LSTM (MMResLSTM) ağı önermektedir. Önerilen model, farklı modaliteler arasındaki zaman ve mekansal korelasyonları öğrenmek üzere tasarlanmıştır ve duygu tanıma performansını ölçüde artırmayı amaçlamaktadır. DEAP veri seti üzerinde gerçekleştirilen deneyler, MMResLSTM ağının uyarılma ve değerlik boyutlarında sırasıyla %92.87 ve %92.30 doğruluk oranlarına ulaştığını göstermektedir. Bu sonuçlar, duygu tanıma alanındaki araştırmalarda çoklu modlu yaklaşımların önemini vurgulamaktadır (Ma v.d.).

Huang ve arkadaşları tarafından gerçekleştirilen "Yüz İfadeleri ve EEG'nin Füzyonu ile Çok Modlu Duygu Tanıma" çalışması, EEG ve yüz ifadeleri arasındaki füzyon yöntemlerini kullanarak duygu tanımanın iki çok modlu yöntemini önermektedir. Araştırma, dört temel duygu durumunu (mutluluk, nötr, üzüntü ve korku) tanımak için sinir ağı sınıflandırıcıları ve iki destek vektör makinesi (SVM) kullanmaktadır. Çalışmanın sonuçları, çok modlu füzyon yöntemlerinin, yalnızca yüz ifadesi veya EEG algılama yöntemlerine kıyasla daha yüksek doğruluk oranlarına ulaştığını göstermektedir (%81.25 ve %82.75). Bu bulgular, duygu tanıma için yüz ifadeleri ve EEG bilgilerinin birleşiminin, tekil bilgi kaynaklarının eksikliklerini giderilebileceğini göstermektedir (Huang v.d.).

Chaparro ve arkadaşlarının "EEG ve Yüz İfadelerinden Duygu Tanıma: Çok Modlu Bir Yaklaşım" çalışması, EEG ve yüz ifadeleri verilerinin özellik seviyesinde füzyonu ile sinir ağı modeli ve rastgele orman sınıflandırıcısının kullanıldığı bir yöntem sunar. MAHNOB-HCI dengeli çok modlu veritabanı kullanılarak gerçekleştirilen deneysel kurulum, diğer çalışmalarla karşılaştırıldığında %97 doğruluk oranıyla başarılı sonuçlar sunmaktadır. Bu çalışma, tek

modlu tekniklerine göre duygu başına %12'ye kadar iyileştirme sağlayan özellik seviyesinde füzyon yaklaşımının etkinliğini göstermektedir (Chaparro v.d.).

Keshari ve Palaniswamy'nin "Yüz İfadeleri ve Vücut Jestleri Kullanılarak Duygu Tanıma" başlıklı çalışması, yüz ifadeleri ve üst vücut jestlerinin özellik seviyesinde füzyonu temel alarak iki modlu bir duygu tanıma modeli önermektedir. Bu yaklaşım, yüz ifadeleri ve vücut dili gibi iki farklı veri kaynağından elde edilen bilgilerin bütünleştirilmesine odaklanmıştır. Özellikle, yüz ifadelerinden ve üst vücut hareketlerinden duyguları bağımsız olarak tanıma ve daha sonra bu özellikleri birleştirerek duyguyu tanıma stratejileri geliştirilmiştir. Bu çalışma, farklı modalitelerden gelen verilerin birleştirilmesinin, duygu tanıma doğruluğunu önemli ölçüde artırabileceğini göstermiş ve bu alandaki araştırmalar için bir perspektif sunmuştur (Keshari v.d.).

Huang ve diğerlerinin "Yüz İfadeleri ve Elektroensefalografiyi Kullanarak Duygu Tanıma Performansının Artırılması" başlıklı çalışması, yüz ifadeleri ve EEG verilerinin füzyonuna dayanan çok modlu bir duygu tanıma çerçevesi sunmaktadır. Çalışma, yüz ifadeleri için çok görevli CNN (Evrişimli Sinir Ağı) mimarilerini ve EEG tanıma için SVM (Destek Vektör Makineleri) sınıflandırıcılarını kullanarak değerlik ve uyarılma duygusal modeline dayanır. Sonuçlar, her iki veri kaynağının karar düzeyinde birleştirilmesinin, tekil yöntemlere kıyasla duygu tanıma doğruluğunu artırabileceğini göstermektedir (Huang v.d.).

Wu ve diğerlerinin "Derin Öğrenme Kullanılarak İfade-EEG Etkileşimi ve İş birliği Yoluyla Çok Modlu Duygu Tanıma" başlıklı çalışması, video ve EEG sinyallerinden duygu tanıma için hiyerarşik LSTM tabanlı bir model sunar. Bu model, yüz ifadesi ve EEG sinyallerini bütünleştirerek duygu tanımada önemli bir iyileşme sağlamayı hedefler. Önerilen yöntem, karmaşık duygu durumlarını daha doğru bir şekilde tanıyabilmek için hem mekânsal hem de zamansal dikkat mekanizmalarını kullanır. Çalışma, özellikle EEG sinyallerinden elde edilen bilgileri ve duygu verilerinin zamansal bilgilerini etkin bir şekilde kullanmayı amaçlar. Deney sonuçları, modelin diğer yöntemlere göre daha yüksek sınıflandırma oranı ve tanıma performansı elde ettiğini göstermektedir (Wu v.d.).

Rayatdoost ve arkadaşlarının "EEG Sinyalleri ve Yüz Davranışlarından Duygu Tanıma için Çok Modlu Kapılı Bilgi Füzyonu" çalışması, EEG sinyalleri ve yüz davranışları arasında derin öğrenme ile çok modlu bir yaklaşım sunmaktadır. Araştırma, EEG sinyalleri ve yüz videolarından elde edilen bilgiler arasındaki karşılıklı ve tamamlayıcı bilgileri öğrenme yeteneğine sahip bir derin temsil öğrenme modeli geliştirmiştir (Rayatdoost v.d.).

Muhammad ve diğerleri tarafından yapılan "Elektroensefalografi ve Yüz Dizilerinin Füzyonu ile İki Modlu Duygu Tanıma Yaklaşımı" başlıklı çalışma, EEG ve yüz video kliplerini derin kanonik korelasyon analizi (DCCA) kullanarak birleştiren bir duygu tanıma yöntemi sunmaktadır. Bu yaklaşım, her iki modaliteden yüksek derecede korelasyona sahip özellikleri çıkararak temel insan duygularını (mutlu, nötr, üzgün) yüksek doğrulukla sınıflandırmaya odaklanmaktadır. MAHNOB-HCI ve DEAP veri setleri kullanılarak yapılan araştırma, multimodal füzyonun insan duygularını anlamada etkinliğini gösteren yüksek doğruluk oranları sergilemiştir. Bu yöntem, EEG ve yüz verilerinin tamamlayıcı güçlerinden faydalanma yeteneğiyle öne çıkmakta, HCI (İnsan-Bilgisayar Etkileşimi) sistemlerindeki duygu tanıma zorluklarına bir çözüm sunmaktadır (Muhammad v.d.).

Houssein ve diğerleri tarafından yazılan "EEG Tabanlı Beyin-Bilgisayar Arayüzleri Kullanılarak Duygu Tanıma: Yöntemler ve İlerlemeler Üzerine Bir Derleme" başlıklı makale, EEG tabanlı beyin-bilgisayar arayüzleri (BCI) kullanılarak duygu tanıma alanındaki yöntemler ve ilerlemeler üzerine kapsamlı bir bakış sunar. Makale, duygu tanıma için EEG sinyallerine uygulanan çeşitli makine ve derin öğrenme algoritmalarını inceler, bunların etkinliklerini ve karşılaşılan zorlukları tartışır. EEG sinyal ön işleme, özellik çıkarma, seçim ve sınıflandırma yöntemlerinin duygu tanıma sistemlerinin doğruluğunu artırmadaki önemini vurgular. Ayrıca, EEG sinyallerini diğer fizyolojik ölçümlerle bütünleştiren çok modlu yaklaşımların, duygu tanıma sistemlerinin sağlamlığını ve güvenilirliğini artırmadaki önemine dikkat çeker (Houssein v.d.).

Huang ve diğer yazarlar tarafından gerçekleştirilen "Elektroensefalogram Tabanlı Bir Beyin-Bilgisayar Arayüzü ile Nöro-geribildirim Eğitiminin Duygu Düzenlemesini Geliştirmesi" başlıklı çalışma, duygu düzenleme yeteneğini artırmak için EEG tabanlı bir beyin-bilgisayar arayüzü (BCI) ve nöro-geribildirim eğitimi sunar. Araştırma, katılımcıların duygusal durumlarını gerçek zamanlı olarak tahmin edebilen ve bu tahminlere dayanarak kullanıcılara geribildirim sağlayan yenilikçi bir BCI sistemi geliştirmiştir. Bu sistem, katılımcıların belirli bir duygusal duruma yönelik duygularını etkili bir şekilde düzenlemelerine imkân tanır. Çalışmanın sonuçları, katılımcıların BCI tabanlı nöro-geribildirim eğitimiyle duygularını düzenleme yeteneklerinde önemli bir gelişme gösterdiğini ve uzun vadeli eğitimin, EEG tabanlı nöro-geribildirim eğitiminin duygusal veya mental bozuklukları olan bireylere yardımcı olabilecek bir araç olabileceğini öne sürmektedir (Huang v.d.).

Leong ve diğerkleri tarafından yürütölen "Yüz İfadeleri ve Beden Jestleri Duygu Tanıma: Görsel Verilerin Duygusal Bilişimde Kullanımı Üzerine Sistematik Bir İnceleme" başlıklı çalışma, görsel veriler kullanılarak duygu tanıma üzerine kapsamlı bir inceleme sunar. Bu derleme, yüz ifadeleri ve beden jestlerinin duygu tanıma alanında kullanımını, uygulanan teknikleri, karşılaşılan zorlukları ve gelecekteki araştırma yönlerini detaylı bir şekilde ele alır. Yapılan incelemeler, görsel verilerin duygu tanıma sistemlerindeki etkinliğini ve bu sistemlerin geliştirilmesinde kullanılan çeşitli yaklaşımları ortaya koyar. Bu çalışma, görsel verilerin duygu tanıma konusunda nasıl etkili bir şekilde kullanılabileceğine dair kapsamlı bilgiler sunarak, bu alandaki araştırmacılara yol gösterir (Leong v.d.).

Zhang ve arkadaşlarının "Çok Modlu Veri ve Makine Öğrenimi Teknikleri Kullanılarak Duygu Tanıma: Bir Eğitim ve İnceleme" başlıklı çalışması, duygu tanımanın son yıllarda makine öğrenimi ve bilgi füzyonu alanlarındaki hızlı gelişmeler sayesinde önemli bir ilgi alanı haline geldiğini ele alır. Araştırma, insan duygularının yüz ifadeleri, konuşma, davranışlar (jest/postür) veya fizyolojik sinyallerden tanınabileceğini, ancak ilk üç yöntemin sosyal maskelenme nedeniyle etkisiz olabileceğini belirtir. Fizyolojik sinyaller, özellikle EEG sinyalleri, duygusal durumları daha objektif ve güvenilir bir şekilde tanıma potansiyeline sahip olduğunu vurgular. Çalışma, çok kanallı EEG sinyalleri ve çok modlu fizyolojik sinyaller temelinde duygu tanıma yöntemlerini, özellik çıkarma, özellik indirgeme ve makine öğrenimi sınıflandırıcı tasarım yöntemlerini kapsamlı bir şekilde incelemektedir (Zhang v.d.).

Çimtay ve diğerk yazarların çalışması, doğal olarak aldatıcı yüz ifadelerinin olduğu durumlarda doğru duygusal durumu belirlemede etkili olan çok modlu bir duygu tanıma sistemi sunmaktadır. Çalışmada, yüz ifadeleri, galvanik deri tepkisi (GSR) ve elektroensefalogram (EEG) olmak üzere birden fazla modalite kullanılmıştır. Bu modalitelerden elde edilen veriler, hibrit bir füzyon stratejisi kullanılarak işlenmiş ve maksimum tek-kullanıcı dışı doğruluk oranı %81.2 olarak belirlenmiştir. Aynı zamanda, özellikle doğal aldatıcı yüz ifadeleri durumunda duygusal durumu doğru bir şekilde belirlemede kullanışlı olan model, literatürdeki benzer bağımsız çok modlu duygu tanıma çalışmalarıyla karşılaştırıldığında iyi bir performans sergilemektedir. Özellikle, DEAP veri seti üzerinde gerçekleştirilen farklı duygu sınıfları için, ortalama doğruluk oranı %53.8 olarak ölçülmüştür. Bu çalışma, çok modlu duygu tanıma sistemlerinin doğruluk ve dayanıklılık açısından nasıl iyileştirilebileceğini göstermektedir. (Cimtay v.d.)

Tan ve diğer yazarların araştırmasında, yüz ifadeleri ve elektroensefalogram (EEG) verileri kullanılarak çok modlu bir duygu tanıma yöntemi önerilmiştir. Bu yöntem, açık kaynak veri setlerinden elde edilen yüz ifadesi sınıflandırma ve EEG özellik çıkarma yöntemlerinin incelenmesiyle başlamış, ardından bu yöntemler kendi toplanan veriler üzerinde uygulanmıştır. Ayrıca, küçük veri setlerinden kaynaklanan sorunları çözmek için Monte Carlo yöntemi kullanılmıştır. Bu çok modlu duygu tanıma yöntemi, insan-bilgisayar etkileşimi (HRI) sistemi ile bütünleştirilmiş ve %83.33 tanıma oranı elde edilmiştir. Kullanıcı deneyimlerine dayalı olarak sistem, katılımcılar tarafından 0 ile 10 arasında ortalama 7 puan almıştır. Bu sonuçlar, yöntemin HRI sistemlerindeki uyumsuzluk hissini azaltma potansiyelini ve etkinliğini göstermektedir. (Tan v.d.)

Chen ve diğer yazarların çalışması, EEG ve çevresel fizyolojik sinyalleri kullanarak duygu tanımayı geliştirmek için çok modlu bir yöntem önermektedir. Araştırma, tek modlu özelliklerin yetersiz kaldığı durumları ele almakta ve EEG ile çevresel fizyolojik sinyaller arasındaki ilişkileri keşfetmektedir. Çalışmada, tek boyutlu dizin özelliklerini iki boyutlu harita özelliklerine dönüştüren, dikkat mekanizmaları ve yinelenen graf konvolüsyonel ağları içeren bir model kullanılmıştır. Bu model, zamansal, mekânsal ve frekans özelliklerini çıkararak derin öğrenme teknikleriyle duygusal durumları sınıflandırmaktadır. Çapraz birey doğrulama kullanılarak yapılan deneylerde, DEAP veri seti üzerinde %93.19 ve %91.82'lik sınıflandırma başarıları elde edilmiştir ki bu, EEG tek modalitesine kıyasla önemli bir iyileşmedir. (Chen v.d.)

Wang ve diğer yazarlar, EEG sinyalleri ve yüz ifadelerinin füzyonuna dayalı çok modlu duygu tanıma yöntemini sunmaktadır. Öncelikle, yüz ifadelerinden yüz özelliklerini çıkarmak için önceden eğitilmiş bir CNN kullanılmıştır. Daha sonra, dikkat mekanizması kullanılarak daha kritik yüz çerçeve özellikleri elde edilmiştir. EEG sinyallerinden uzaysal özellikler çıkarmak için de CNN'ler uygulanmış, bu sırada yerel ve global konvolüsyon çekirdekleri kullanılarak sol ve sağ beyin kanalları ile tüm EEG kanallarının özellikleri öğrenilmiştir. Özellik seviyesinde füzyon yapıldıktan sonra, elde edilen çok modlu özellikler duygu tanıma için sınıflandırıcıya beslenmiştir. Çalışma, DEAP ve MAHNOB-HCI veri setleri üzerinde gerçekleştirilen deneylerle modelin performansını değerlendirmiş, valans boyutu sınıflandırması %96.63, uyarılma boyutu sınıflandırması %97.15 doğruluk oranına ulaşmıştır. MAHNOB-HCI veri setinde ise sırasıyla %96.69 ve %96.26 doğruluk elde edilmiştir (Wang v.d.).

Chakravarthy ve diğer yazarlar, EEG sinyalleri ve video verilerini birleştirerek duyguları tanımak için optimizasyon destekli çok modlu bir model önerdiler. Çalışmada, Weiner filtreleme yöntemi ile ön işlenen EEG sinyalleri ve video kareleri kullanılmıştır. Özellik çıkarımı ve geliştirilmiş özellik seviyesi füzyonu ile elde edilen özellikler, CNN ve Deep Maxout sınıflandırıcılarından oluşan hibrit bir model ile analiz edilmiştir. Combined Arithmetic and Pelican with hybrid exploration (CAPHE) algoritması ile optimize edilen model, %90.84 gibi yüksek bir doğruluk oranı elde etmiştir. (Chakravarthy v.d.)

3. METODOLOJİ

3.1 Veri Seti

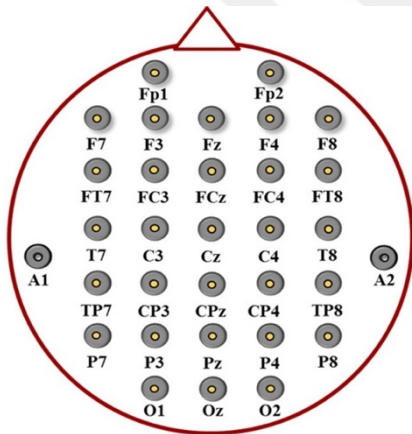
DEAP veri seti, elektroensefalografi (EEG) sinyallerini yüz ifadeleri ile bütünleştiren, duygu tanıma alanında yaygın olarak kullanılan bir veri setidir. Özellikle multimedya uyarıcıları tarafından tetiklenen duygusal durumları keşfetmek için tasarlanmıştır. Katılımcıların EEG kayıtlarını, fizyolojik sinyallerini ve yüz ifadelerinin video kayıtlarını içeren DEAP, duygu sınıflandırma algoritmalarının geliştirilmesi ve test edilmesi, multimedya içeriğine karşı duygusal tepkilerin incelenmesi ve duygu tanımda farklı modaliteler arasındaki etkileşimin araştırılması gibi çeşitli uygulamalarda kullanılmaktadır.

Veri setinde bulunan 32 kanallı EEG sinyallerinin alınması sürecinde, ilk olarak katılımcı işlem hakkında bilgilendirilmiş ve rahat bir pozisyonda oturması sağlanmıştır. Daha sonra uzman yardımıyla kafa derisine 32 adet elektrot yerleştirmek üzere bir işaretleyici ve ölçüm aleti kullanılmıştır. Bu elektrotlar, uluslararası kabul görmüş 10-20 yerleştirme sistemine göre konumlandırılmıştır. Her elektrotun kafa derisine iyi bir temas sağlaması için, özel bir jel kullanılarak, elektrotlar aracılığıyla beyin dalgalarının daha net bir şekilde kaydedilmesi sağlanmıştır. Elektrotlar yerleştirildikten sonra, elektrotların doğru pozisyonunu korumak ve hareket ederek gürültü oluşturmalarını engellemek amacıyla elektrot kabı kullanılmıştır. Bir katılımcının kayıttan kısa süre önce alınan görseliyle (Şekil 3.1) elektrot kabı, elektrotların yerleşimi, bilgisayara aktarılma düzeni gibi detaylı bilgiler sunulmuştur. Şekil3.2’ de ise

elektrotların detaylı yerleşimi ve bunların hangi kanallara karşılık geldiği yer almaktadır. Burada görülen kanallar veri setindeki özniteliklere karşılık gelmektedir.

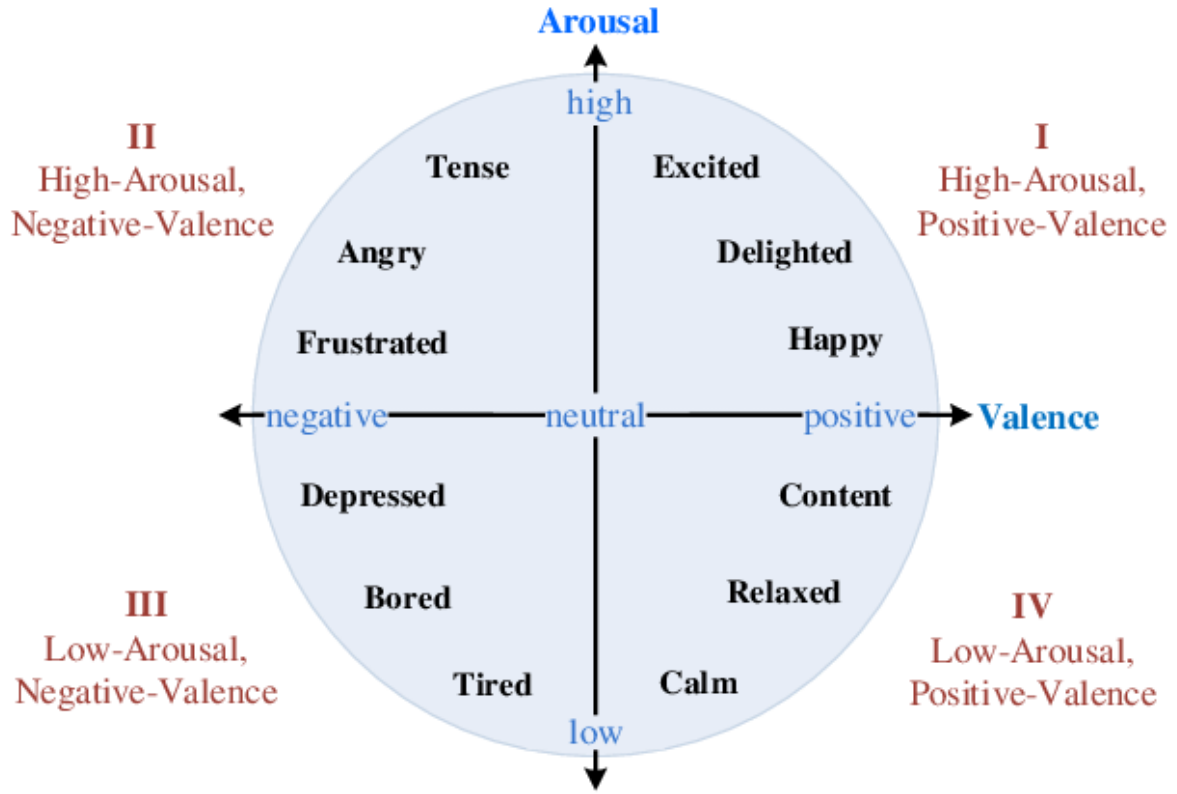


Şekil 3.1 DEAP veri seti katılımcısından deney esnasında alınan görsel



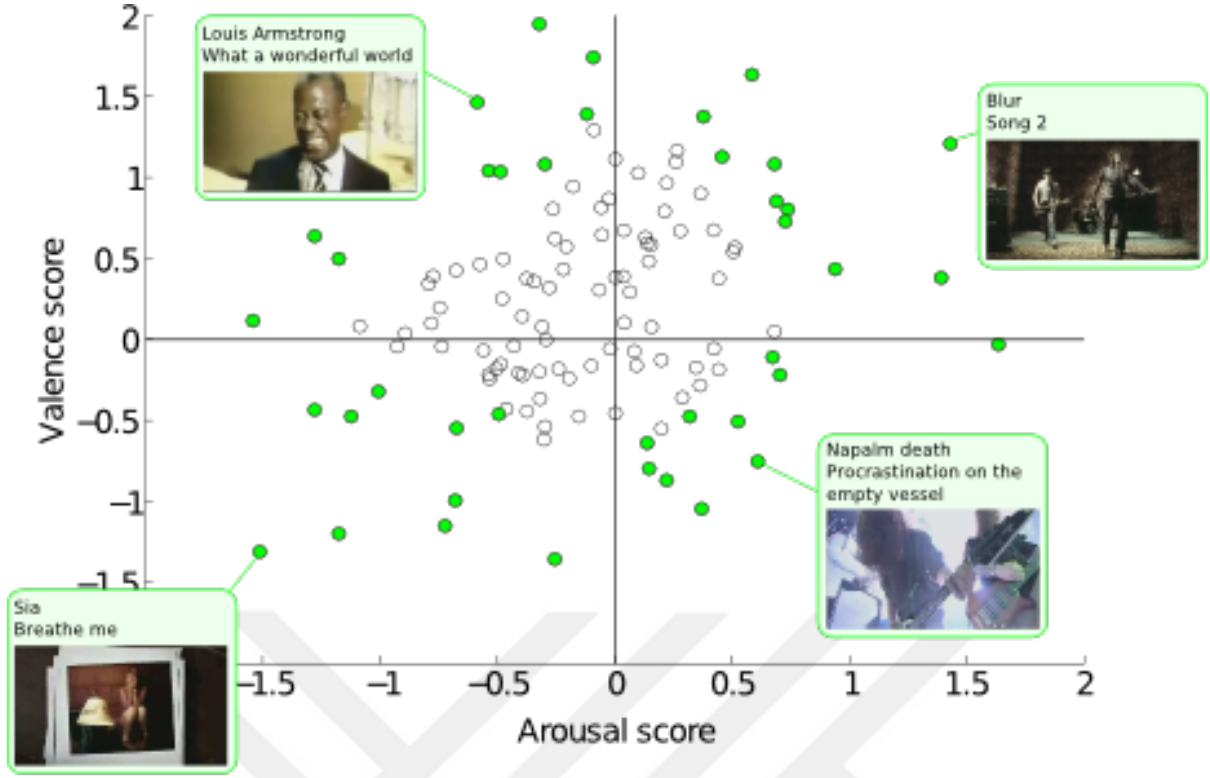
Şekil 3.2 Elektrot yerleşimi ve kanalların gösterimi

DEAP'ın önemi, duygu araştırmalarında kritik boyutlar olan uyarılma, değerlik ve baskınlık ölçümlerini içeren kapsamlı bir çerçeve sunmasında yatmaktadır. Bu sayede her katılımcının video uyarıcılarına verdiği duygusal yanıtın standartlaştırılmış bir şekilde kaydedilmesi ve analiz edilmesi mümkün olmaktadır. Uyarılma ve değerlik eksenlerinde duyguların ifade edilmesi yaklaşımı, duygusal durumları çoklu boyutlarda sürekli sayısal değerler olarak temsil eder (Russell). Değerlik, hoş ve nahoş (yani, pozitif ve negatif) duyguların derecesini temsil ederken, uyarılma heyecan ve sakinlik derecesini temsil eder. Bu temsile dayanarak, herhangi bir duygusal durum, uyarılma-değerlik koordinat düzleminde bir nokta olarak gösterilebilir. (Şekil 3.3)



Şekil 3.3 İki boyutlu uyarılma-değerlik alanı

DEAP veri seti hazırlanırken deneyde kullanılan videoların uyarılma-değerlik alanına göre dağılımı Şekil 3.4'te sunulmuştur. Kullanılan videolar yeşil renkle vurgulanmıştır. Her çeyrek için bir video, şarkı adı ve videodan bir ekran görüntüsü ile detaylandırılmıştır.



Şekil 3.4 DEAP veri setinde kullanılan videoların uyarılma-değerlik ekseninde dağılımı

EEG ve yüz ifadesi verilerinin birlikte değerlendirilmesi, DEAP'ın en önemli özelliklerinden biridir. EEG, beyin aktivitesini doğrudan ölçerken, yüz ifadeleri duygusal durumun dışavurumudur. Bu iki veri türünün birleştirilmesi, duyguların nöral alt yapısı hakkında daha derin bir anlayış edinmemizi sağlar. DEAP'ın sağladığı kapsamlı veri seti, ileri düzey sinyal işleme ve makine öğrenimi tekniklerinin kullanımıyla daha nüanslı duygu tanıma sistemlerinin geliştirilmesine katkıda bulunmaktadır. Araştırmacılar, DEAP'ı kullanarak EEG ve yüz ifadesi verilerinden anlamlı desenler çıkarabilir ve duyguların nasıl ortaya çıktığına dair yeni bilgiler edinebilirler.

DEAP veri setinin sunduğu çok boyutlu yaklaşım, insan duygusal deneyiminin karmaşıklığını yakalamak için önemlidir. Fizyolojik ve davranışsal verileri birleştirerek, duyguların sadece beyinsel değil, aynı zamanda bedensel ve sosyal etkileşimlerle de bağlantılı olduğunu görebiliriz. DEAP veri setinin kullanımı, duygusal bilişim alanında önemli ilerlemelere yol açabilir. Daha empatik ve etkileşimli insan-bilgisayar arayüzleri geliştirmek, bu ilerlemelerin en önemli sonuçlarından biri olacaktır. DEAP, bu alandaki araştırmaların temelini oluşturarak, gelecekte duyguların daha iyi anlaşılmasına ve insan-makine etkileşiminin geliştirilmesine katkıda bulunacaktır.

3.2 Veri Ön İşleme

3.2.1 EEG Sinyalleri ile Duygu Sınıflandırma

Bu bölümde, DEAP veri seti üzerinde EEG verilerinin işlenmesi ve analizi süreci detaylandırılacaktır. Araştırmanın odak noktası, EEG sinyalleri kullanılarak duygusal durumların yüksek doğrulukla sınıflandırılması için yeni bir yöntem geliştirmektir. Bu kapsamda, EEG verileri üzerinde kapsamlı bir ön işleme süreci uygulanmış ve çeşitli makine öğrenimi modelleri değerlendirilmiştir.

Ön işleme süreci, veri setinin doğruluğunu ve analiz sırasında elde edilecek sonuçların güvenilirliğini maksimize etmek için kritik öneme sahiptir. İlk adım olarak, EEG sinyalleri özel bir fonksiyon aracılığıyla işlenmiştir. Bu fonksiyon, katılımcı numarası, EEG kanalları, veri bölümlerinin süresi ve bölümler arası mesafeyi parametre olarak almaktadır. Kullanılan 256'lık pencere ve adım boyutu, her adımın 1 saniyeyi temsil etmesini sağlamıştır. Böylece, EEG sinyalleri her biri 256 veri noktası içeren 1 saniyelik bölümlere ayrılmıştır. Fonksiyon, 40 katılımcı üzerinde EEG kanallarından alınan verileri bölümlere ayırmaktadır. Pencereleme işlemi *Denklem 1*'de gösterildiği gibi, başlangıç değişkeninden itibaren pencere süresi kadar devam etmekte ve başlangıç pozisyonu her adımda arttırılarak veri üzerinde kaydırılmaktadır.

$$\text{başlangıç pozisyonu} = \text{başlangıç pozisyonu} + \text{pencere boyutu} \quad (\text{Denklem 1})$$

Duygusal durum değişikliklerini etkili bir şekilde yakalamak için tasarlanmış olan bu yöntem, EEG verilerinin farklı zaman dilimlerinde karakteristik özelliklerinin analiz edilmesine olanak tanımıştır.

```
Function Epoching(subject_number, channel_list, window_size, step_size)
    Initialize an empty list named signal_list
    Open the data file for the given subject_number in 'rb' mode
    Load the file using pickle with 'latin1' encoding and assign it to the variable subject
    For each trial from 0 to 39
        Retrieve data and labels for the current trial from subject
        Convert labels into numerical classes using the map_label function and convert to an array
        Initialize the variable start to 0
        While start + window_size is less than or equal to the number of columns in data
            Create an empty list named array
            For each channel in channel_list
                Slice the data from the current channel, starting from start up to window_size, and
assign to X
                Add X to the list array
                Add labels to the list array
            Add the list array to signal_list
            Increment start by step_size
        Convert signal_list into a numpy array
        Save the processed data to save_path + subject_number file
```

Sözde Kod 3.1 EEG veri işleme

Ardından, her veri bölümü, katılımcının o anki duygusal durumunu yansıtan bir etiketle eşleştirilmiştir. Değerlik (valence) ve uyarılma (arousal) değerlerinden oluşan iki boyutlu etiketler, bir dönüşüm işlemi aracılığıyla tek boyutlu ve dört farklı sayısal sınıfa dönüştürülmüştür. Bu sınıflar Yüksek Değerlik Yüksek Uyarılma (HVHA), Yüksek Değerlik Düşük Uyarılma (HVLA), Düşük Değerlik Yüksek Uyarılma (LVHA) ve Düşük Değerlik Düşük Uyarılma (LVLA) olarak temsil edilmektedir. Bu dönüşüm, duygusal durumların daha sade ve sayısal bir biçimde ifade edilmesini sağlar, böylece veri analizi ve işleme süreçleri kolaylaştırılır.

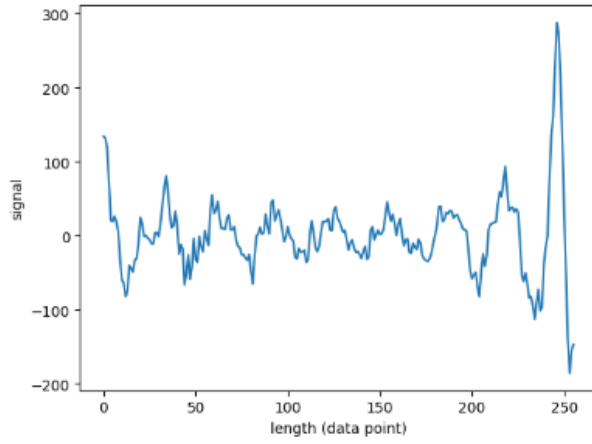
```
Define a function named label_mapping with two parameters: valence and arousal.  
  
Start of the function:  
  
Check if valence is greater than 5 AND arousal is greater than 5.  
  
    If true, set the label as 0 (representing 'HVHA').  
  
Otherwise, check if valence is greater than 5 AND arousal is less than or equal to 5.  
  
    If true, set the label as 1 (representing 'HVLA').  
  
Otherwise, check if valence is less than or equal to 5 AND arousal is greater than 5.  
  
    If true, set the label as 2 (representing 'LVHA').  
  
Otherwise, assume valence is less than or equal to 5 AND arousal is less than or equal to 5.  
  
    Set the label as 3 (representing 'LVLA').  
  
Return the label as the output of the function.  
  
End of the function.
```

Sözde Kod 3.2 Etiket dönüşümünü yapan fonksiyon

Etiket dönüşümü işlemi, Sözde kod 3.2’de tanıtılan, **label_mapping** fonksiyonu ile gerçekleştirilmiştir. Bu fonksiyon, değerlik ve uyarılmanın sayısal değerlerini alır ve bu değerlere göre bir etiket (label) döndürür. Fonksiyonun izlediği yol şu şekildedir:

- HVHA (High Valence, High Arousal), hem değerlik hem de uyarılma değerleri 5’ten büyük olan durumlardır.
- HVLA (High Valence, Low Arousal), değerlik değeri 5’ten büyük, uyarılma değeri 5 veya daha düşük olan durumları kapsamaktadır.
- LVHA (Low Valence, High Arousal), değerlik değeri 5 veya daha küçük, uyarılma değeri 5’ten büyük olan durumları içermektedir.
- LVHA (Low Valence, Low Arousal) ise hem değerlik hem de uyarılma değerleri 5 veya küçük olan durumlardır.

Bu yaklaşımla, veri setinin hem EEG sinyalleri hem de çok modlu sınıflandırmaya uygun tek boyutlu etiketler içermesi sağlanmıştır.



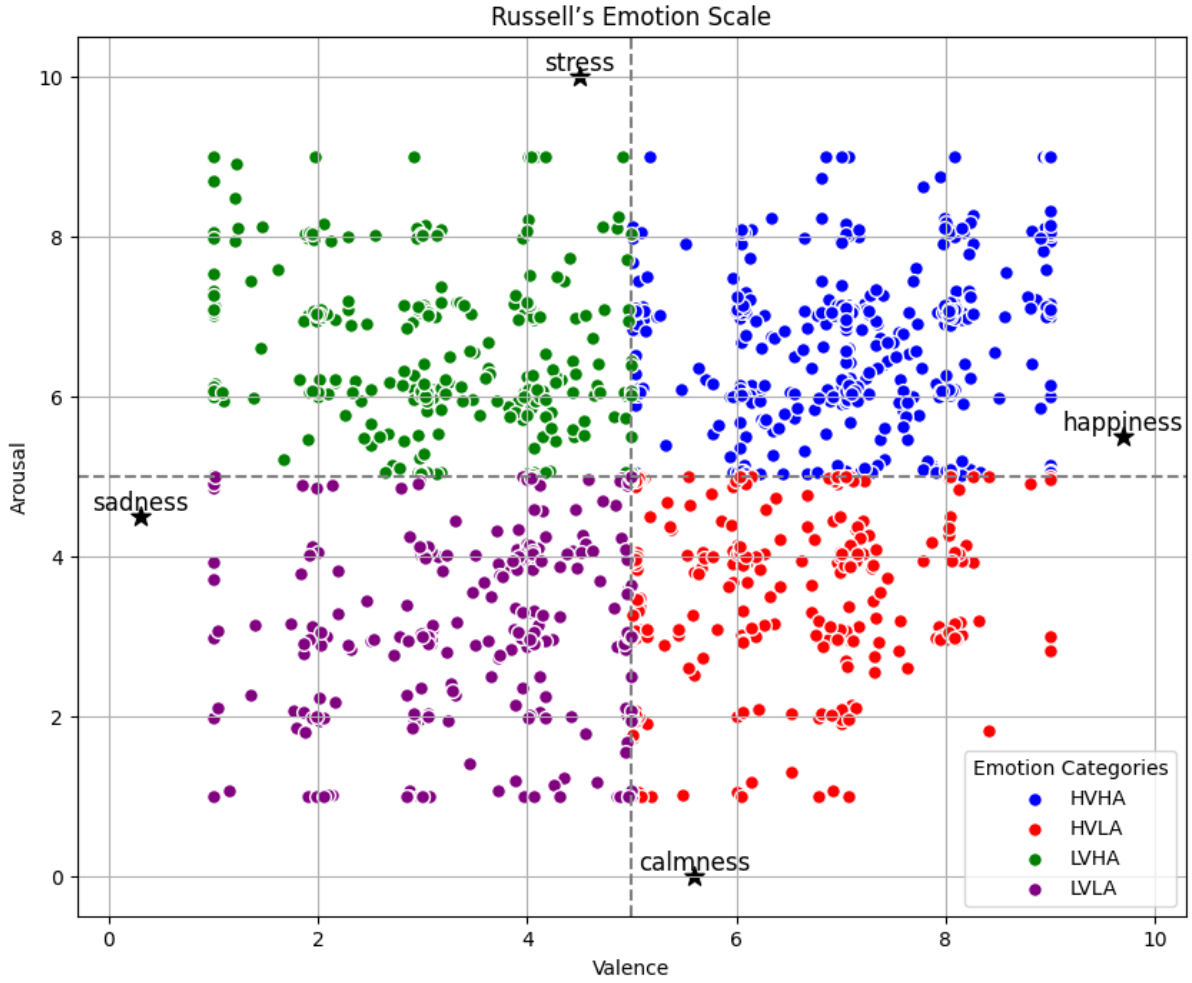
Label	Corresponding Emotion
High Valence Low Arousal (HVLA)	Calmness



Face Expression Recognition Results	
Emotion	Rate (%)
neutral	41
sad	28
others (angry, fear etc.)	31

Şekil 3.5 Katılımcı 11'e ait eeg sinyali görselleştirmesi, karşılık gelen yüz görüntüsü, duygu etiketi ve yüz ifadesi tanıma sonuçları

Katılımcı ID'si 11 olan bir katılımcının izlediği videodan alınan 1 saniyelik kesit üzerinden, bu zaman dilimine ait EEG verisi görselleştirilmiştir. Bu EEG verisine karşılık gelen duygu etiketinin HVLA olduğu tespit edilmiştir.



Şekil 3.6 Russel'in duygu ölçeğine göre veri setinin dağılımları

EEG sinyallerine karşılık gelen dört farklı etiketin (HVHA, HVLA, LVHA, LVLA) hangi duyguları içerdiği, Russell'ın duygu ölçeği (Şekil 3.6) kullanılarak tespit edilmiştir. Bu bağlamda, HVLA etiketiyle eşleşen duygunun 'calmness' olduğu belirlenmiştir. Aynı katılımcının 1 saniyelik EEG sinyaline denk gelen zaman dilimindeki yüz ifadesi videosundan alınan görüntü üzerinde yürütülen yüz ifadesi tanıma analizleri, Şekil 3.5'te "Yüz İfadesi Tanıma Sonuçları" altında sunulmuştur. Bu analizler sonucunda, görüntü karesinde %41 ile tespit edilen en yüksek yoğunluklu duygunun nötr olduğu saptanmıştır. EEG sinyalleri ve yüz ifadesi tanıma sonuçları birlikte değerlendirilerek, bütüncül bir duygu analizi yapılmıştır.

Kullanılan veri seti eğitim için %70, validasyon için %15 ve test için %15 olacak şekilde bölünmüştür.

Tablo 3.1 Eğitim, doğrulama ve test kümelerindeki veri sayısı

Veri seti tipi	Tek modlu	Çok Modlu
Eğitim	388864	267344
Validasyon	83328	57288
Test	83328	57288

Tek modlu ve çok modlu çalışmalarda veriler aynı oranda bölünmüş olsa da her iki yöntemde kullanılan veri miktarları farklıdır. Çok modlu yöntemde, DEAP veri setindeki yüz ifadeleri verileri ile EEG verilerinin farklı boyutta olmaları, birleştirildikleri zaman veri boyutunun değişmesine sebep olmuştur. Bu durum daha detaylı bir şekilde Bölüm 3.2.2'de açıklanmıştır. Her iki yöntem için kullanılan veri miktarları sayısal olarak Tablo 3.1'de sunulmuştur.

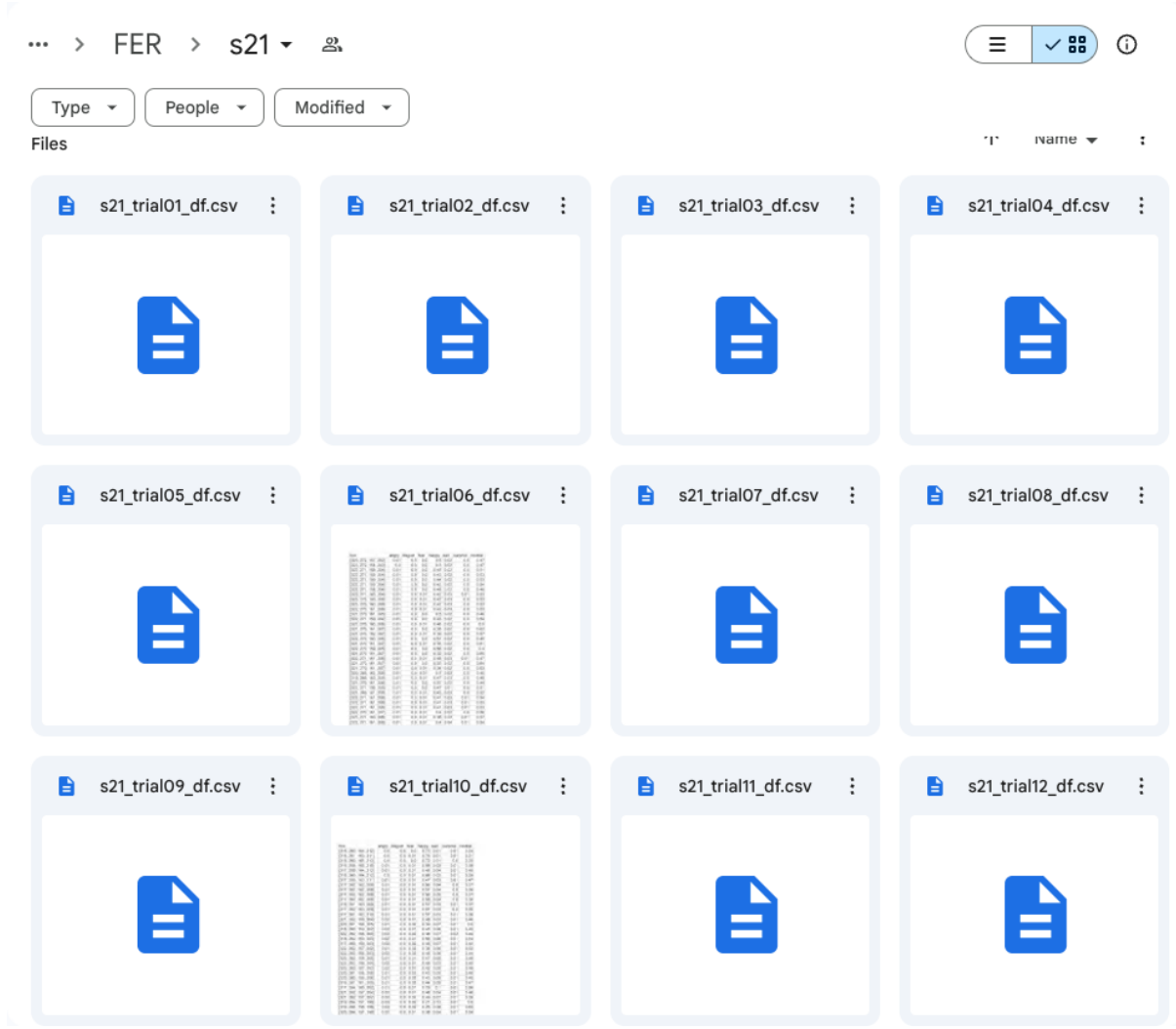
Verilerin eğitim, doğrulama ve test olarak bölünmesi işlemi, modelin tutarlılığını sağlamak amacıyla tek seferlik yapılmış ve eğitim sürecinde tüm modeller için aynı veri setleri kullanılmıştır. Veri setlerinin görselleştirilmesi yapılmış ve bu görselleştirmelere dayalı analiz süreçleri yürütülmüştür. Eğitim, doğrulama ve test setlerindeki sınıf etiketleri, ikili sınıf matrislerine (0 ve 1) dönüştürülerek, sınıf etiketleri arasındaki hiyerarşik ilişki önlenmiştir. Bu sayede, modelin her sınıfı ayrı bir kategori olarak değerlendirmesi ve doğru şekilde sınıflandırması sağlanmıştır.

Bu dönüşümün ardından, veriler normalizasyon, ölçeklendirme ve yeniden şekillendirme işlemlerinden geçirilmiştir. Farklı EEG kanalları ve katılımcılar arasında tutarlılık sağlamak için verilerin normalizasyonu ve ölçeklendirilmesi gerçekleştirilmiştir. Normalizasyon, verileri ortalama değeri 0 ve standart sapması 1 olacak şekilde dönüştürürken; min-max ölçeklendirme, verileri 0 ile 1 arasında ölçeklendirmektedir. Bu işlemler, modelin eğitim sürecini optimize ederken aşırı uyuma karşı koruma sağlamaktadır.

Son olarak, veriler modelin işleyebileceği formata getirilmiş ve gereken durumlarda veri artırma teknikleri uygulanmıştır. Bu yeniden boyutlandırma ve veri artırma işlemleri, modelin verimliliğini ve sınıflandırma performansını artırmak için oldukça önemlidir. Bu aşamaların her biri, EEG verilerinden maksimum bilgiyi çıkararak duygusal durumların doğru bir şekilde sınıflandırılmasını sağlamak için dikkatlice planlanmıştır. Ön işleme sürecinin titizlikle uygulanması, makine öğrenimi modellerinin performansını artırarak duygu tanıma çalışmalarında daha başarılı sonuçların elde edilmesine olanak tanımaktadır.

3.2.2 Çok Modlu Duygu Sınıflandırma: Yüz İfadeleri ve EEG Sinyalleri

Bu bölümde, DEAP veri setinden alınan ön işlenmiş EEG verileriyle katılımcıların yüz videolarından elde edilen duygu analizi sonuçlarının entegrasyonuna odaklanılmıştır. DEAP veri seti kapsamında, katılımcıların izlediği 40 farklı videoya verdikleri yüz tepkileri FER (Facial Expression Recognition) algoritması (Luz v.d.) ile analiz edilmiştir. FER, gelişmiş görüntü işleme teknikleriyle, insan yüzündeki mikro ifadeleri algılayarak, bunları belirlenen duygu kategorilerine atamaktadır. Tüm katılımcıların yüzlerinin videoya alınmasına onay vermemesi nedeniyle yalnızca 22 katılımcının yüz videoları analiz edilebilmiştir. Ayrıca, her katılımcının toplamda 40 tane yüz videosu olması gerekirken, bazı katılımcıların son videolarında meydana gelen teknik aksaklıklardan dolayı birkaçında bu sayı 40'tan az olmuştur.



Şekil 3.7 Yüz ifadesi veri setleri

Her bir video, 1 dakika uzunluğunda ve 50 FPS (Frame Per Second) hızındadır. Toplamda 3000 (50 FPS x 60 sn) tane görüntüye sahip videolar, FER aracıyla işlenmiş ve analiz sonuçları, yüz koordinatları (x,y,w,h) ve 7 farklı duygu durumu ('öfke', 'iğrenme', 'korku', 'mutluluk', 'üzüntü', 'şaşkınlık', 'nötr') için tablo olarak dosyalar hâlinde (Şekil 3.7) karşılık gelen katılımcı id klasörlerine kaydedilmiştir.

Bu çalışmada, MTCNN (Zhang v.d.) yüz tanıma algoritması, yüz ifadelerinin duygu analizi için FER aracı ile kullanılmıştır. MTCNN, yüzleri algılayıp koordinatlarını duygu analizi için FER aracına aktarmaktadır, FER ise bu yüz verilerinden duygu analizi yapmaktadır. Bu süreç, tüm katılımcıların yüz videoları için tekrarlanmıştır.

Ancak, bazen MTCNN'nin yüz algılamada başarısız olması veya FER'in duygusal ifadeleri doğru çıkaramaması durumlarından ötürü, bazı videolardaki duygu analizi veri setlerinin beklenen 3000 satır yerine daha az satırla oluşmasına neden olmuştur. DEAP veri setinde, her bir katılımcının izlediği 40 farklı videoya karşılık gelen 17360 adet işlenmiş EEG sinyali örneği bulunmaktadır. Bu örnekler tüm katılımcılar için aynı sayıdadır. Katılımcı başına elde edilen bütün EEG sinyali örneklerinin video sayısına bölünmesi neticesinde, her bir videodan 434 adet EEG sinyali örneğinin çıkarıldığı sonucuna varılmıştır. Bu hesaplama, EEG sinyallerinin yüz ifadesi verileriyle birleştirilmesi sürecinde, her bir video analizinden en az 434 satırlık duygu analizi veri setinin oluşması gerektiğini ifade etmektedir. Ancak, bazı videolardan alınan yüz analizi sonuçları beklenen örnek sayısına ulaşamamıştır. Bu sorunu çözmek adına yüz tanıma için FER aracında Haar Cascade (Viola v.d.) algoritması kullanılmıştır. Bu yaklaşım, eksik veri olan videoları yeniden işleyerek daha fazla veri içeren bazı sonuçlar üretmiştir.

```

If carry_to_new_path is set to 'True':

    # Define directory paths

    first_FER_directory_path = MTCNN_algorithm_path

    second_FER_directory_path = Haar_Cascade_algorithm_path

    save_FER_directory_path =
    Combined_data_path_for_both_algorithm_to_more_detailed_data

    # If the save directory exists, delete it and create a new directory

    If save_FER_directory exists:

        Delete save_FER_directory

    Create save_FER_directory

    # Copy every item from the first directory to the new directory, skipping '.ipynb' files

    For each item in the first_FER_directory:

        If the item ends with '.ipynb', skip it

        If the item is a directory:

            Copy the item from the first directory to the save directory

    # Check and compare every folder and file in the second directory

    For each folder in the second_FER_directory:

        If the folder exists in the save directory:

            For each file in the folder:

                If the file already exists in the save directory and the second file has more lines:

                    Copy the file from the second directory to the save directory

    # Copy missing folders and files from the second directory to the new directory

    For each folder in the second_FER_directory:

        If the folder does not exist in the save directory:

            Copy the folder from the second directory to the save directory

    Else:

        For each file in the folder:

            If the file does not exist in the save directory:

                Copy the file from the second directory to the save directory

    Print "File transfer process is completed."

```

Sözde Kod 3.3 Haar Cascade ve MTCNN Algoritmalarının Ürettiği Veri Setlerini Karşılaştırarak Daha Geniş Veri Setini Kayıt Dizinine Kaydeden Fonksiyon

Bu aşamada, yeni üretilen sonuçlar eğer daha fazla veri içeriyor ve boyutları daha büyükse, daha az veri içeren önceki sonuçlarla değiştirilmiştir. Eğer MTCNN algoritması ile elde edilen sonuçlar, Haar Cascade ile elde edilenlerden daha fazla veya eşit miktarda bilgi sağlıyorsa, bu sonuçlar tercih edilmiştir. Sonuç olarak, her durumda, mümkün olan en kapsamlı analiz sonuçları elde edilmiştir.

Yapılan bu işlemler sonucunda, yüz videolarından çıkarılan duygu analizi sonuçlarında minimum 434 örneğin elde edilemediği durumlar için, eksik örnek sayısı kadar sentetik veri üretilmesini sağlayan bir fonksiyon yazılmıştır. Bu fonksiyon, orijinal veri setinden yola çıkarak, belirlenen hedef uzunluğa ulaşana dek sentetik veri üretme işlemini gerçekleştirmektedir. Sentetik veri üretimi sırasında, mevcut veri setinin uzunluğu ile hedeflenen uzunluk arasındaki fark kadar ekstra veri örnekleri üretilir. Her bir sentetik veri örneği, orijinal veri setinden rastgele seçilen bir satırın sayısal değerlerine belirli bir değişiklik uygulanarak oluşturulmaktadır. Bu değişiklik, ortalama değeri sıfır ve standart sapması 0.01 olan normal bir dağılım tarafından üretilen rastgele bir "gürültü" miktarının eklenmesi şeklinde gerçekleştirilmektedir. Bu yöntemle, veri setine küçük varyasyonlar eklenerek daha gerçekçi ve çeşitliliği yüksek sentetik verilerin oluşturulması hedeflenmiştir. Orijinal veri setindeki değerler 0 ile 1 arasında olduğundan, yeni elde edilen veri satırında negatif değerlerin 0'a ve 1'den büyük değerlerin ise 1'e yuvarlanması sağlanmıştır. Bu yaklaşım, veri değerlerinin belirlenen aralıkta kalmasını sağlayarak veri bütünlüğünün korunmasına katkıda bulunmuştur. Sonrasında, tüm değerler iki ondalık basamağa yuvarlanır, bu da önemsiz farkların ortadan kaldırılmasına ve veri setinde tutarlılığının sağlanmasına yardımcı olmaktadır. Son olarak, bu işlem tüm eksik veriler üzerinde tekrar edilir ve elde edilen tüm sentetik veri satırları bir veri çerçevesine dönüştürülür. İşlem sonucunda, oluşturulan sentetik veri seti kayıt dizinine boyutu daha küçük olan orijinal veriyle yer değiştirilerek kaydedilir. Bu yöntem, orijinal veri setinin yapısını ve özelliklerini korumuş ve aynı zamanda, veri çeşitliliğini ve hacmini artırmıştır.

```

Function GenerateSyntheticData(original_df, target_length):
    Begin
        1. Identify non-numeric columns in the original DataFrame and store their names.
            non_numeric_columns = Identify and store names of columns in original_df that are not
            numeric

        2. Calculate the current length of the DataFrame and the additional rows needed to reach
        the target length.
            current_length = Get the number of rows in original_df
            additional_length = target_length - current_length

        3. Initialize an empty list to store synthetic data rows.
            synthetic_data = Empty list

        4. Loop to generate each synthetic row until the additional_length is met.
            For each iteration from 1 to additional_length:
                a. Randomly select a row from the original DataFrame excluding non-numeric columns.
                    random_row = Randomly select one row from original_df excluding
                    non_numeric_columns

                b. Generate noise with mean=0 and standard deviation=0.01, matching the shape of the
                random_row.
                    noise = Generate normal distribution noise with mean=0 and standard deviation=0.01,
                    shaped like random_row

                c. Add the noise to the random_row to create a new synthetic row.
                    new_row = Add noise to random_row

                d. Clip the values in new_row to ensure they are between 0 and 1.
                    new_row = Clip values of new_row between 0 and 1

                e. Round the values in new_row to two decimal places.
                    new_row = Round values in new_row to two decimal places

                f. For each non-numeric column, copy the original value from the corresponding row in
                the original DataFrame to the new_row.
                    For each column in non_numeric_columns:
                        Assign the value from the original row in original_df to new_row for this column

                g. Append the new_row to the synthetic_data list.
                    Add new_row to synthetic_data

        5. Convert the synthetic_data list into a DataFrame.
            synthetic_df = Convert synthetic_data list to DataFrame

```

Sözde Kod 3.4 Sentetik veri üreten fonksiyon

Ayrıca, yüz ifadesi analiz sonuçlarının sayısının EEG sinyallerinin sayısını aştığı durumlarda, fazla olan her bir sonuçtan alt örnekleme(downsampling) ile 434 örnek seçilerek örnek sayıları eşitlenmiştir. Veri setlerinin eşitlenmesi için uygulanan alt örnekleme süreci üç temel aşamadan oluşur. Bunlar, alt örnekleme faktörünün hesaplanması, alt örnekleme işlemi ve alt örnekleme işleminden geçirilen FER verileri ile EEG verilerinin doğru şekilde entegre edilmesi aşamalarından meydana gelmektedir. Veri azaltma sürecinde, FER verilerine, EEG verileriyle zaman ekseninde uyum sağlaması amacıyla alt örnekleme uygulanmıştır. Alt örnekleme işlemlerini gerçekleştiren (downsample_fer) fonksiyon, FER verisinden alt örnekleme faktörüne karşılık gelen indisteki öğeyi seçerek veri setinin boyutunu azaltmaktadır. Bu işlem, FER veri setinin frekansını düşürerek EEG veri setinin frekansıyla uyumlu hale gelmesini sağlamıştır. Alt örnekleme faktörü 'calculate_downsample_factor' fonksiyonu kullanılarak hesaplanmıştır. Bu fonksiyon, toplam FER veri uzunluğunu EEG veri uzunluğuna bölerek bir oran elde etmektedir. Elde ettiği bu oranı tam sayıya yuvarlayarak alt örnekleme faktörünü belirlemektedir. Bu faktör, FER verisini EEG verisiyle aynı uzunlukta olacak şekilde azaltmak için kullanılmaktadır. EEG ve FER verilerini birleştirme sürecinde, her bir FER dosyası için, dosyadaki verileri okuyan ve gereksiz sütunları (örneğin, 'box') çıkartan bir fonksiyon kullanılmıştır. Bu fonksiyon ile elde edilen FER verisi, hesaplanan alt örnekleme faktörü ile azaltılmış ve EEG verisiyle zaman ekseninde uyumlu hale getirilmiştir. Bu işlem sırasında, her bir FER örneği ilgili EEG örneğiyle birleştirilerek bütünleşmiş bir veri seti oluşturulmuştur. Böylelikle, her bir videodan elde edilen EEG sinyalleri, ilgili videoların düzeltilmiş yüz ifadesi analiz sonuçları ile birleştirilmiştir. Veri işleme sürecinin geri kalanında, 'Sadece EEG sinyalleri ile duygu sınıflandırma' bölümünde açıklanan sürecin aynısı izlenmiştir. Ancak, yüz analizi sonuçlarının zaten 0 ile 1 arasında olması sebebiyle, veri setindeki son 7 özelliğe normalizasyon uygulanmamıştır.

```

Function DownsampleFER(fer_data, factor=7):
    """
    Reduces the FER data size by the specified factor through downsampling.
    """
    Begin
        1. Return every nth row of fer_data, where n equals the factor.
    End
Function CalculateDownsampleFactor(fer_total_length, eeg_length=17360):
    """
    Calculates the downsampling factor needed based on the total length of FER data and the
    target EEG data length.
    """
    Begin
        1. Calculate the factor as the integer division of fer_total_length by eeg_length.
        2. Return the calculated factor.
    End
Function CombineEEGAndFER(eeg_data_path, fer_data_dir):
    """
    Combines EEG and FER data by first downsampling the FER data to match the EEG data length,
    then concatenating them.
    """
    Begin
        1. Load the EEG data from the provided path.
        2. Collect paths of all FER data files in the specified directory and sort them.

        3. Calculate the total number of rows in all FER data files.
        4. Calculate the downsampling factor needed for FER data to match EEG data length.

        5. Initialize an empty list for storing combined data samples.
        6. Initialize an index for tracking the current position in EEG data.

        7. For each FER data file path:
            a. Load the FER data as a DataFrame and drop any non-essential columns.
            b. Downsample the FER data using the calculated factor.
            c. For each row in the downsampled FER data:
                i. Check if the end of EEG data has been reached; if so, exit the loop.
                ii. Combine the current EEG data sample, the FER data row, and the label from the EEG
                data into a new sample.
                iii. Append the new sample to the combined data list.
                iv. Increment the EEG data index.

        8. Convert the combined data list to a numpy array and return it.
    End

```

Sözde Kod 3.5 Alt örnekleme faktörünün hesaplanması, alt örnekleme işlemi, EEG ve FER verilerinin birleştirilmesi

3.3 Model Eğitimi

Veriler üzerinde duygu sınıflandırması yapmak için GRU (Gated Recurrent Unit), LSTM (Long Short-Term Memory) ve Transformer modelleri olmak üzere üç farklı derin öğrenme yaklaşımı kullanılmıştır. (Dey v.d. ; Hochreiter v.d. ; Vaswani v.d.)

3.3.1 GRU

GRU modellerinin güncelleme ve sıfırlama olmak üzere iki kapısı (gate) vardır. GRU'lar bilgi akışını yönetebilmek için gating birimlerinden faydalanırlar. Bu da gerekli bilgilerin daha uzun süreler boyunca tutulmasına olanak sağlamaktadır. (Dey v.d.)

Update kapısı z_t , modelin hangi bilgilerin atılacağına ve hangi yeni bilgilerin hücre durumuna ekleneceğine karar vermesini sağlar. Bu kapı, modelin geçmiş bilgilerin ne kadarının korunması gerektiğine dair kararları desteklemektedir. Denklem II'de, W_z ile update gate için weight matrisi, σ ile sigmoid aktivasyon fonksiyonu, h_{t-1} ile previous hidden state ve x_t ile t zaman adımındaki girdi temsil edilmektedir.

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t])$$

(Denklem II)

Reset kapısı r_t ne kadar geçmiş bilginin unutulacağına karar vermek için GRU modeli tarafından kullanılır. Modelin gerektiği zaman hafızasını sıfırlamasını sağlar. Denklem III'de, W_r ile reset gate için weight matrisi ifade edilmektedir.

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t])$$

(Denklem III)

Model: "sequential"

Layer (type)	Output Shape	Param #
gru (GRU)	(None, 256, 128)	50304
dropout (Dropout)	(None, 256, 128)	0
batch_normalization (Batch Normalization)	(None, 256, 128)	512
gru_1 (GRU)	(None, 256, 64)	37248
dropout_1 (Dropout)	(None, 256, 64)	0
batch_normalization_1 (Batch Normalization)	(None, 256, 64)	256
gru_2 (GRU)	(None, 32)	9408
dropout_2 (Dropout)	(None, 32)	0
dense (Dense)	(None, 4)	132
Total params: 97860 (382.27 KB)		
Trainable params: 97476 (380.77 KB)		
Non-trainable params: 384 (1.50 KB)		

Şekil 3.8 GRU modelinin özeti

Çalışmada kullanılan GRU modeli, üç katmandan (128, 64, 32 birim) oluşmaktadır. İlk iki katmanda “return_sequences=True” parametresi ile GRU katmanları arasındaki geçişte sıralı veri yapısı korunmaktadır. Bu sayede, modelin daha derin katmanlarda zamanla ilişkili özellikleri öğrenmesi sağlanmaktadır. Üçüncü ve son GRU katmanında ise “return_sequences” parametresi “False” olarak ayarlanmıştır. Bu ayar, son katmanda tüm zaman adımlarındaki bilgileri tek bir vektörde toplayıp, çıkış katmanına iletmektedir.

Şekil 3.8'de modelin özellikleri ve parametreleri gösterilmektedir. Bu modelde aşırı uyumu önlemek ve modelin genelleştirme yeteneğini arttırmak amacıyla L1 ve L2 regülerizasyonu (Ng v.d.) uygulanmıştır. L1 regülerizasyonu, modelin ağırlık matrisindeki bazı değerleri sıfıra yaklaştırarak daha basit bir model oluşturmayı hedeflerken, L2 regülerizasyonu ağırlıkların karelerinin toplamını minimize ederek ağırlık değerlerinin büyük ölçüde artmasını engellemektedir. Her GRU katmanının ardından eklenen Dropout katmanları (Srivastava v.d.), eğitim sırasında rastgele seçilen nöronları devre dışı bırakarak modelin farklı nöron kombinasyonlarını kullanmasını ve böylece daha kararlı bir yapıya ulaşmasını sağlamaktadır.

Dropout yaklaşımı, derin öğrenme modellerinde aşırı uyumu önlemek için yaygın olarak kullanılan bir tekniktir. Bunlara ek olarak, modelin daha hızlı ve kararlı bir şekilde eğitilmesine olanak tanıyan BatchNormalization (Ioffe v.d.) katmanları eklenmiştir. Bu katmanlar, her bir mini-batch içindeki aktivasyonların dağılımını normalleştirerek, ağ katmanları arasındaki aktivasyonların ölçeklenmesini optimize eder. Çıkış katmanında ise tahmin edilen sınıf olasılıklarını üretecek şekilde softmax aktivasyon fonksiyonuna (Sharma v.d.) sahip bir Dense katmanı yer almaktadır.

3.3.2 LSTM

LSTM'ler, ileriye doğru bilgi akışı sağlamak üzere verileri işleyebilen bir kontrol akışına sahiptir. LSTM hücrelerinde, hangi bilgilerin hatırlanacağı ve hangilerinin unutulacağını belirleyen kapılar (gate) ve verileri diğer hücrelere taşıyan hücre durumu (cell state) bulunmaktadır. LSTM'lerde üç ana kapı bulunmaktadır: Unutma Kapısı (Forget Gate), Girdi Kapısı (Input Gate) ve Çıktı Kapısı (Output Gate). Bu kapılar, hücre durumuna giren ve çıkan bilgi akışını düzenleyerek, LSTM'lerin önemli bilgileri uzun süre hatırlamasını sağlar (Hochreiter v.d.).

Forget gate f_t önceki hücre durumundan hangi bilgilerin atılması veya saklanması gerektiğine karar verir. Önceki hücre durumundaki her bir sayı için 0 ve 1 arasında bir değer üretir. Bu değer 0'a yaklaştıkça unutmak, 1'e yaklaştıkça tutmak anlamına gelir. Denklem IV'te, W_f sembolü forget gate için weight matrisini, b_f sembolü forget gate için bias'ı temsil etmektedir.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

(Denklem IV)

Giriş kapısı (Denklem V), hücre durumunu güncellemek için girişten gelen hangi değerlerin kullanılacağına karar verir. Aday hafızası (Denklem VI) ile cell state'e yeni potansiyel değerlerin eklenmesi sağlanmaktadır.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

(Denklem V)

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$$

(Denklem VI)

Çıkış kapısı (Denklem VII), hücre durumuna göre bir sonraki state'in ne olması gerektiğini belirler.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

(Denklem VII)

Model: "sequential"

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 256, 128)	66560
dropout (Dropout)	(None, 256, 128)	0
batch_normalization (Batch Normalization)	(None, 256, 128)	512
lstm_1 (LSTM)	(None, 256, 64)	49408
dropout_1 (Dropout)	(None, 256, 64)	0
batch_normalization_1 (Batch Normalization)	(None, 256, 64)	256
lstm_2 (LSTM)	(None, 32)	12416
dropout_2 (Dropout)	(None, 32)	0
dense (Dense)	(None, 4)	132
Total params: 129284 (505.02 KB)		
Trainable params: 128900 (503.52 KB)		
Non-trainable params: 384 (1.50 KB)		

Şekil 3.9 LSTM modelinin özeti

LSTM modelinin eğitim süreci için benzer şekilde üç farklı LSTM katmanı kullanılmıştır. İlk iki katmanda “return_sequences” parametresinin ayarı True olarak tanımlanmıştır. Bu, katman çıkışlarının sonraki katman için sıralı veri olarak kalmasını sağlamaktadır. Böylece, zamanla ilişkili bilgilerin korunmasını mümkün kılmaktadır. Üçüncü LSTM katmanında ise “return_sequences” parametresi “False” ayarlanmıştır. Bu sayede, zaman boyutu son katmanda sıkıştırılarak sonraki Dense katmanına tek bir vektör olarak iletilmiştir. Her LSTM katmanında, Şekil 3.9’da belirtildiği üzere, GRU modelinde olduğu gibi aşırı uyumu önlemek amacıyla L1 ve L2 regularizasyonları uygulanmıştır. Her bir LSTM katmanından sonra modelin daha robust olması için Dropout katmanları eklenmiştir. Eğitim sürecini hızlandırmak ve daha stabil bir hale getirmek amacıyla katmanlar arası aktivasyonları normalize eden Batch Normalization katmanları da dâhil edilmiştir. Modelin çıkış katmanı, sınıf olasılıklarını temsil etmek için kullanılan softmax aktivasyon fonksiyonu ile oluşturulmuş bir Dense katmandır. Son olarak model, kategorik çapraz entropi kaybı (categorical cross entropy) ve Adam optimizasyonu kullanılarak derlenmiştir.

3.3.3 Transformer

Transformer, tekrarlanan bağlantılara dayanmadan, sequence-to-sequence görevleri verimli bir şekilde gerçekleştiren özellikle doğal dil işleme (NLP) görevlerinde oldukça başarılı olan bir derin öğrenme modelidir. Transformer, tüm sequence’i aynı anda işlemek için self attention mekanizmalarını kullanır. Bu mekanizmalar, daha iyi paralelleştirmeye ve uzun vadeli bağımlılıkların daha başarılı bir şekilde yakalanmasına imkân tanırlar. Son zamanlarda, doğal dil işlemenin yanı sıra zaman serisi tahminleme, sınıflandırma gibi farklı uygulama alanlarında da kullanılmaktadır. Transformer modeli, Şekil 3.10, Şekil 3.11 ve Şekil 3.12’te temsil edilen üç farklı görseldeki model özetine göre, MultiHeadAttention, katman normalizasyonu, dropout ve konvolüsyonel katmanları içeren transformer kodlayıcı (encoder) bloklarından oluşmaktadır. Her bir kodlayıcı blok, dikkat mekanizmasıyla başlar ve ardından ileri beslemeli (feed-forward) bir ağ içerir.

Dikkat mekanizmalarının içerdiği MultiHead Attention katmanları (Denklem VIII), modelin veri集中的 özellikler arasındaki ilişkileri birden fazla perspektiften değerlendirmesine olanak tanımaktadır (Vaswani v.d.).

$$MultiHead(Q, K, V) = Concat(head_1, \dots, head_h)W^o$$

(Denklem VIII)

Model: "model"

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	[(None, 256, 1)]	0	[]
multi_head_attention (MultiHeadAttention)	(None, 256, 1)	7169	['input_1[0][0]', 'input_1[0][0]']
dropout (Dropout)	(None, 256, 1)	0	['multi_head_attention[0][0]']
layer_normalization (LayerNormalization)	(None, 256, 1)	2	['dropout[0][0]']
tf.__operators__.add (TFOpLambda)	(None, 256, 1)	0	['layer_normalization[0][0]', 'input_1[0][0]']
conv1d (Conv1D)	(None, 256, 8)	16	['tf.__operators__.add[0][0]']
dropout_1 (Dropout)	(None, 256, 8)	0	['conv1d[0][0]']
conv1d_1 (Conv1D)	(None, 256, 1)	9	['dropout_1[0][0]']
layer_normalization_1 (LayerNormalization)	(None, 256, 1)	2	['conv1d_1[0][0]']
tf.__operators__.add_1 (TFOpLambda)	(None, 256, 1)	0	['layer_normalization_1[0][0]', 'tf.__operators__.add[0][0]']
multi_head_attention_1 (MultiHeadAttention)	(None, 256, 1)	7169	['tf.__operators__.add_1[0][0]', 'tf.__operators__.add_1[0][0]']
dropout_2 (Dropout)	(None, 256, 1)	0	['multi_head_attention_1[0][0]']
layer_normalization_2 (LayerNormalization)	(None, 256, 1)	2	['dropout_2[0][0]']
tf.__operators__.add_2 (TFOpLambda)	(None, 256, 1)	0	['layer_normalization_2[0][0]', 'tf.__operators__.add_1[0][0]']
conv1d_2 (Conv1D)	(None, 256, 8)	16	['tf.__operators__.add_2[0][0]']
dropout_3 (Dropout)	(None, 256, 8)	0	['conv1d_2[0][0]']
conv1d_3 (Conv1D)	(None, 256, 1)	9	['dropout_3[0][0]']

Şekil 3.10 Transformer model mimarisi birinci bölüm

layer_normalization_3 (LayerNormalization)	(None, 256, 1)	2	['conv1d_3[0][0]']
tf.__operators__.add_3 (TF OpLambda)	(None, 256, 1)	0	['layer_normalization_3[0][0]', 'tf.__operators__.add_2[0][0]']
multi_head_attention_2 (MultiHeadAttention)	(None, 256, 1)	7169	['tf.__operators__.add_3[0][0]', 'tf.__operators__.add_3[0][0]']
dropout_4 (Dropout)	(None, 256, 1)	0	['multi_head_attention_2[0][0]']
layer_normalization_4 (LayerNormalization)	(None, 256, 1)	2	['dropout_4[0][0]']
tf.__operators__.add_4 (TF OpLambda)	(None, 256, 1)	0	['layer_normalization_4[0][0]', 'tf.__operators__.add_3[0][0]']
conv1d_4 (Conv1D)	(None, 256, 8)	16	['tf.__operators__.add_4[0][0]']
dropout_5 (Dropout)	(None, 256, 8)	0	['conv1d_4[0][0]']
conv1d_5 (Conv1D)	(None, 256, 1)	9	['dropout_5[0][0]']
layer_normalization_5 (LayerNormalization)	(None, 256, 1)	2	['conv1d_5[0][0]']
tf.__operators__.add_5 (TF OpLambda)	(None, 256, 1)	0	['layer_normalization_5[0][0]', 'tf.__operators__.add_4[0][0]']
multi_head_attention_3 (MultiHeadAttention)	(None, 256, 1)	7169	['tf.__operators__.add_5[0][0]', 'tf.__operators__.add_5[0][0]']
dropout_6 (Dropout)	(None, 256, 1)	0	['multi_head_attention_3[0][0]']
layer_normalization_6 (LayerNormalization)	(None, 256, 1)	2	['dropout_6[0][0]']
tf.__operators__.add_6 (TF OpLambda)	(None, 256, 1)	0	['layer_normalization_6[0][0]', 'tf.__operators__.add_5[0][0]']

Şekil 3.11 Transformer model mimarisi ikinci bölüm

conv1d_6 (Conv1D)	(None, 256, 8)	16	['tf.__operators__.add_6[0][0]']
dropout_7 (Dropout)	(None, 256, 8)	0	['conv1d_6[0][0]']
conv1d_7 (Conv1D)	(None, 256, 1)	9	['dropout_7[0][0]']
layer_normalization_7 (LayerNormalization)	(None, 256, 1)	2	['conv1d_7[0][0]']
tf.__operators__.add_7 (TF OpLambda)	(None, 256, 1)	0	['layer_normalization_7[0][0]', 'tf.__operators__.add_6[0][0]']
global_average_pooling1d (GlobalAveragePooling1D)	(None, 256)	0	['tf.__operators__.add_7[0][0]']
dense (Dense)	(None, 128)	32896	['global_average_pooling1d[0][0]']
dropout_8 (Dropout)	(None, 128)	0	['dense[0][0]']
dense_1 (Dense)	(None, 4)	516	['dropout_8[0][0]']

=====
Total params: 62204 (242.98 KB)
Trainable params: 62204 (242.98 KB)
Non-trainable params: 0 (0.00 Byte)

Şekil 3.12 Transformer model mimarisi üçüncü bölüm

İleri-beslemeli ağ, her MultiHead Attention bloğundan sonra gelen iki konvolüsyonel katmandan meydana gelmektedir. Bu katmanlar, modelin öğrenme kapasitesini artırmak ve daha karmaşık özellikler çıkarabilmek için kullanılmıştır. Katman normalizasyonu ve dropout katmanı, modelde aşırı uyumu önlemek ve daha kararlı bir öğrenme süreci sağlamak için uygulanmıştır. Transformer modeli, LSTM ve GRU'ya benzer şekilde, kayıp fonksiyonu olarak kategorik çapraz entropi ve optimizasyon tekniği olarak Adam optimizasyonu kullanılarak derlenmiştir.

Tablo 3.2 Transformer modelinin hiperparametreleri ve tanımlamaları

Hiperparametreler	Değerler	Tanımlamalar
input_shape	(None,256,1) / (None,263,1)	Modelin girdi olarak kabul edeceği veri şeklidir.
n_classes	4	Modelin tahmin etmeye çalıştığı sınıf sayısıdır.
head_size, num_heads	256, 4	Dikkat mekanizmasının boyutlarını ve başlık sayısını tanımlar.
ff_dim	4	İleri-beslemeli ağın boyutunu tanımlar.

num_transformer_blocks	4	Modelde kullanılacak transformer kodlayıcı bloklarının sayısıdır.
mlp_units	[128]	Modelin sonundaki çok katmanlı algılayıcı (MLP) ağının birim sayısını tanımlar.
dropout, mlp_dropout	0.2, 0.2	Aşırı uyumu önlemek için kullanılan dropout oranlarıdır.

Tablo 3.2'de gösterilen Transformer modeline ait hiperparametreler deneysel çalışmalar sonucunda belirlenmiştir. Tabloda input_shape parametresi için iki farklı değer belirtilmiştir. Bunun sebebi, tek modlu ve çok modlu verilerine ait girdi boyutlarının farklı olmasıdır. 'None' olarak belirtilen değer, modelin farklı büyüklüklerdeki veri gruplarını işleyebileceğini göstermektedir. input_shape haricindeki diğer hiperparametreler, hem tek modlu hem de çok modlu yaklaşımlar için aynı kalmıştır. Modellemenin başarısı, bu hiperparametrelerin uygun bir şekilde ayarlanmasına, model parametrelerine, veri setinin kalitesine ve modelin eğitim sürecine bağlıdır.

Modellerin eğitimi esnasında ModelCheckpoint ve EarlyStopping callback'leri kullanılmıştır. ModelCheckpoint, eğitim sırasında en iyi modelin kaydedilmesini sağlamaktadır. Bu sayede, modeller daha sonra yeniden kullanılabilmiştir. EarlyStopping (Prechelt v.d.) ise model performansında belirli bir süre boyunca iyileşme olmaması durumunda eğitimi erkenden durdurarak gereksiz hesaplama sürecinin önüne geçmektedir. Her bir model 330 eğitim turu (epoch) boyunca eğitilmiştir ve her bir eğitim turunda eğitim veri setinin tamamı model tarafından işlenmiştir. Öğrenme oranı (learning rate) da her bir model için yapılan deneysel çalışmalar sonucunda 0.0001 olarak belirlenmiştir. Ek olarak modeller, önceden oluşturulmuş eğitim ve doğrulama veri setleri üzerinde eğitilmiş olup, eğitim süreci modellerin veri setindeki desenleri öğrenme yeteneklerini ve bu öğrenilen bilgilerin doğrulama seti üzerindeki performansını değerlendirmektedir.

Son olarak, her bir modelde hem tek modlu hem de çok modlu veriler için aynı model mimarisi kullanılmış, yalnızca girdi boyutlarında değişiklik yapılmıştır.

3.4 Müzik Öneri Sistemi

Öneri sistemleri, günümüzün bilgi yoğunluklu dünyasında giderek daha fazla öneme sahip olmaktadır. Bu sistemler, geniş veri setleri arasından kullanıcıların ilgi ve ihtiyaçlarına en uygun öğeleri seçerek, onlara özel önerilerde bulunur. Veri ile sürekli beslenen bu sistemler, e-ticaretten medya akış hizmetlerine, sosyal medyadan içerik sağlayıcılara kadar birçok alanda etkin bir şekilde kullanılmaktadır. Çok modlu duygu tanıma sistemleri üzerine yapılan çalışmalar, bu sistemlerin insanların duygularını anlama yeteneğini arttırarak daha kişiselleştirilmiş öneriler sunabileceğini göstermektedir. (Ayata v.d.)

Çalışmamızda kullanılan DEAP veri seti, bu tür çok boyutlu analizler için zengin bir kaynaktır. Ayrıca kullanıcıların müzik videoları izlerken alınan duyguları, öneri sistemlerinde kullanılmak üzere bir veri sağlamaktadır. Bu bölümde DEAP veri setindeki verilerin öneri sistemleri için nasıl girdi oluşturacak şekilde ön işlendiğini ve duygu tanıma işlemi sonucunda gelen duygulara göre öneri sisteminin nasıl oluştuğu detaylı bir şekilde açıklanmıştır.

Öneri sisteminin temel amacı, kullanıcının duygusal ihtiyaçlarına yanıt veren ve onların ruh halini destekleyen veya dönüştüren müzikler sunmaktır. Bu sistem de kullanıcıların mevcut duygusal durumlarını anlayarak, onlara en uygun müzikleri önermek üzere tasarlanmıştır. Böylece müzik, sadece bir eğlence aracı olmaktan çıkıp, kullanıcıların duygusal durumunu destekleyen bir araç haline gelebilir. Ayrıca duygu tanıma ve teknolojileri ve öneri sistemlerinin gücünü birleştirmek, kullanıcı deneyimini kişiselleştirme açısından büyük öneme sahiptir. Özellikle psikolojik araştırmalar açısından değerli bir katkıdır çünkü müzik terapisi insanların duygusal ve bilişsel durumlarını iyileştirmede önemli bir role sahiptir. (Witusik v.d. ; Landis-Shack v.d. ; McCaffrey v.d.) Çok modlu duygu tanıma sistemlerine dayalı müzik öneri sistemi, teknoloji ve insan duygusal zekâsı arasında köprü kurarak, kullanıcı deneyimini derinleştirme ve zenginleştirme potansiyeline sahiptir.

Yapılan çalışmada katılımcıların yaş, cinsiyet, duygusal tepkiler gibi bilgileri kullanılarak yapılan önerilere dair detaylar aşağıda sunulmuştur.

3.4.1 Veri İşleme

Veri işleme aşaması, müzik öneri sistemlerinin doğruluğu ve performansı üzerinde önemli bir etkiye sahiptir. Bu bölüm, veri ön işleme süreçlerinin detaylarını ele almakta olup, verilerin nasıl birleştirildiği, temizlendiği, yapılandırıldığı ve analize uygun hâle getirildiğini kapsamlı bir şekilde açıklar. Özellikle, verilerin doğru şekilde birleştirilmesi, gereksiz bilgilerin ayıklanması, veri setlerinin tutarlı ve anlamlı formatlara dönüştürülmesi öneri sisteminin başarısını doğrudan etkiler. Bu süreçler, algoritmanın daha doğru tahminler yapmasını sağlar.

Veri işleme sürecinin ilk aşamasında, EEG ve Yüz ifadelerinden gelen verilerden elde edilmiş .npy uzantılı veriler okunur. Toplamda 22 katılımcı için ayrı ayrı bulunan bu dosyaların veri boyutu (17360,263) iken bunlara karşılık gelen Etiketlerin boyutu ise (17360,) şeklindedir. Ardından, yüklenen EEG ve yüz ifadesi verileri create_dataframe adında özel bir fonksiyon kullanılarak işlenir. Bu fonksiyon, veri ve etiketleri alır ve belirli sütunları seçerek bir pandas DataFrame'i oluşturur. İlk olarak, veri seti bir DataFrame içine yerleştirilir ve ardından yalnızca EEG ve yüz ifadesi verilerine ait belirli özellikleri içeren sütunlar seçilir (256'dan 262'ye kadar). Bu sütunlar, özellikle duygusal durumları (kızgın, iğrenme, korku, mutlu, üzgün, şaşkın, nötr) temsil eder ve bu özellikler, yeni sütun adları ile güncellenir. Sonrasında, her veri noktasının karşılık gelen etiketi, DataFrame'e 'label' sütunu olarak eklenir ve etiketler tamsayıya çevrilir. Böylece, her bir özellik sütunu belirli bir duygusal ifadeyi temsil ederken, 'label' sütunu ise bu duygusal ifadelerin hangi kategoriye ait olduğunu gösterir. Bu işlem, veri setinin daha sonraki analiz ve modelleme aşamalarında kullanılmak üzere düzenli ve anlamlı bir biçimde organize edilmesini sağlar. Sonuç olarak, bu işlenmiş veri seti, duygusal durum tanıma modellerinin eğitilmesi için hazır hale getirilir ve veri setinin yapısı ve içeriği hakkında daha derinlemesine analiz yapılmasına olanak tanır. Bu fonksiyonun sözde kodu aşağıda Sözde Kod 3.6 ile sunulmuştur.

Function CreateDataFrame(features, labels):

Begin

1. Convert the input features into a DataFrame.

combined_df = Convert features array into a DataFrame

2. Define which columns to keep based on their indices.

columns_to_keep = List of indices [256, 257, 258, 259, 260, 261, 262]

3. Filter the DataFrame to include only the specified columns.

combined_df = Filter combined_df to include only columns_to_keep

4. Assign new names to the columns in the DataFrame.

new_column_names = Dictionary mapping old column indices to new names:

256: 'angry'

257: 'disgust'

258: 'fear'

259: 'happy'

260: 'sad'

261: 'surprise'

262: 'neutral'

Rename columns of combined_df using new_column_names

5. Convert the labels list to integer values and add it as a new column to the DataFrame.

combined_df['label'] = Convert each element in labels to integer and add as a new column

6. Return the processed DataFrame.

Return combined_df

End

Sözde Kod 3.6 Duygu durumları ve Etiketlerin kullanımı ile yeni dataframe oluşturan fonksiyon

Yüz ifadelerinin analizinden gelen duygusal tepkilerin yüzdelerini içeren sütunlarda en yüksek yoğunluğa sahip üç duyguyu belirlemek için bir fonksiyon yazılmıştır. ‘Label’ sütunu hariç ‘angry’, ‘disgust’, ‘fear’, ‘happy’, ‘sad’, ‘surprise’, ‘neutral’ sütunlarında bulunan değerler kıyaslanmış ve her satır için büyükten küçüğe doğru sıralama yapılarak ilk üç değer alınmıştır. Bu işlem neticesinde, her satır için duygusal yoğunluklar sıralanmış ve en baskın üç duygu belirlenmiştir. Fonksiyon bu üç duygunun sayısal ifadelerini değil, isimlerini döndürmektedir. Bu işlem neticesinde oluşan yeni sütunlar, ‘first’, ‘second’ ve ‘third’ olarak DataFrame’e eklenmiştir. Yapılan işlemler, veri setindeki duygusal tepkilerin analizini daha detaylı yapmak ve belirli duygusal durumların baskınlığını anlamak için kullanışlı bir yöntem sağlar.

Bir başka fonksiyon olan “get_most_frequent_label” ile, verilen DataFrame içindeki 'label' sütunundaki en sık rastlanan değer tespit edilir ve bu değer bir çıktı olarak döndürür. İkinci olarak, ‘calculate_emotion_averages’ fonksiyonu, en sık rastlanan etiketle ilgili olarak DataFrame'deki belirli duyguların ortalamalarını hesaplar. Bu fonksiyon, 'first', 'second', ve 'third' sütunlarındaki değerlerin frekansını hesaplar ve en yüksek üç frekansı bulur. Bu en sık rastlanan duygusal kategorilerin ortalama değerleri hesaplanır ve yuvarlanarak bir sözlük içinde döndürülür. Ardından, bir başka fonksiyon yardımıyla, DataFrame 434 satırlık parçalara bölünür ve her bir parça için yukarıdaki işlemler tekrarlanır. Son aşamada 22 katılımcı için bu işlemlerin her biri bir döngü içerisinde tekrarlanır ve bu sayede, her bir katılımcının her video için en sık rastlanan etiketi belirlenmiş olur. Bu işlemler neticesinde her bir katılımcının her bir videosuna karşılık yüz ifadelerinden gelen baskın üç duygunun ortalama değerleri ve en sık rastlanan etiket bilgisi elde edilmiş oldu.

Elde edilen değerler kullanılarak yeni bir DataFrame oluşturulmuştur. ‘Participant_ID’ ve ‘Trial’ bilgileri ile beraber yukarıdaki işlemler neticesinde elde edilen ilk üç baskın duygu, bu duyguların oranları ve label bilgisi bu DataFrame’de yer almaktadır. Oluşan DataFrame’e ait görsel Tablo 3.3’te sunulmuştur.

Tablo 3.3 Duygu durumları ve etiketler ile yapılan ön işlemler sonucu oluşan dataframe

Participant id	Trial	top1 emotion	top2 emotion	top3 emotion	top1 emotion rate	top2 emotion rate	top3 emotion rate	label
01	1	neutral	sad	angry	0.76	0.20	0.03	0
01	2	neutral	sad	angry	0.81	0.13	0.03	0
01	3	neutral	sad	angry	0.83	0.13	0.02	0
01	4	neutral	sad	angry	0.75	0.13	0.07	2
01	5	neutral	sad	angry	0.58	0.15	0.13	1
...	...	...	...	...	...	...	...	...
22	37	sad	neutral	happy	0.30	0.28	0.15	2
22	38	neutral	sad	fear	0.28	0.24	0.18	2
22	39	sad	neutral	fear	0.28	0.26	0.23	2
22	40	sad	fear	neutral	0.28	0.22	0.20	3

Oluşan DataFrame, bir csv dosyasına kaydedilmiştir. DEAP veri setinde bulunan participant_ratings dosyasında bulunan bilgiler de öneri sisteminde kullanılacaktır. Bu sebeple bu dosyadaki bilgiler ile oluşturulan DataFrame’de bulunan bilgiler ‘Participant_ID’ ve ‘Trial’ sütunları temel alınarak birleştirilir. Bu sayede DataFrame’e Experiment_ID ve ‘Liking’ bilgileri de eklenmiş olur. Yine DEAP veri seti ile birlikte verilen metadata dosyaları içerisinde yer alan video_list_modified dosyası da öneri sisteminde kullanılacaktır. Burada Sanatçı ismi, şarkı ismi, lastfm etiketi gibi bilgiler yer almaktadır. Bu dosyada yer alan ‘Artist’, ‘Title’ ve ‘Lastfm_tag’ bilgileri ‘Experiment_ID’ sütunu yardımıyla son oluşturduğumuz DataFrame ile birleştirilir. Bu işlem sonucunda elde edilen DataFrame ise Tablo 3.4’te sunulmuştur.

Tablo 3.4 Participant ratings ve video list dosyalarından gelen bilgilerin de eklenmesiyle oluşan dataframe

Participant id	Trial	Experiment id	top1 emotion	top2 emotion	top3 emotion	top1 emotion rate	top2 emotion rate	top3 emotion rate	Liking	label
1	1	5	neutral	sad	angry	0.76	0.20	0.03	6.05	0
1	2	18	neutral	sad	angry	0.81	0.13	0.03	8.01	0
1	3	4	neutral	sad	angry	0.83	0.13	0.02	8.06	0
1	4	24	neutral	sad	angry	0.75	0.13	0.07	8.22	2
1	5	20	neutral	sad	angry	0.58	0.15	0.13	8.13	1
...	...	...	...	...	...	...	...	...	...	...
22	36	10	sad	neutral	fear	0.32	0.27	0.19	5.00	2
22	37	33	sad	neutral	happy	0.30	0.28	0.15	1.00	2
22	38	2	neutral	sad	fear	0.28	0.24	0.18	7.12	2
22	39	25	sad	neutral	fear	0.28	0.26	0.23	7.05	2
22	40	35	sad	fear	neutral	0.28	0.22	0.20	1.00	3

Son olarak, DEAP veri setinde bulunan participant_questionnaire dosyasında, öneri sistemlerini etkileyebilecek yaş ve cinsiyet gibi önemli bilgiler bu DataFrame'e eklenmiştir. Cinsiyet bilgisi, sayısal bir değer olarak 0 ve 1 şeklinde ifade edilmiştir.

Top1_emotion, top2_emotion ve top3_emotion sütunlarında kızgın, iğrenme, korku, mutlu, üzgün, şaşkın, nötr gibi duygular yer almaktadır. Bu duygular, her bir satırda farklılık gösterebilmektedir. Bu nedenle, bu duygular birer sütun haline getirilmiş ve top1_emotion, top2_emotion, top3_emotion sütunları DataFrame'den çıkarılmıştır. Participant_ID'ye göre, her bir katılımcının izlediği videoya verdikleri en baskın üç duygu, mevcut DataFrame satırında 1 değerini almış, diğer duygular ise 0 değerini almıştır. Bu sayede, her bir duygu değişkeni nümerik forma dönüştürülmüştür. 1 olan duyguların oranları, top1_emotion_rate, top2_emotion_rate ve top3_emotion_rate sütunlarından alınmış ve bu sütunlar da DataFrame'den çıkarılmıştır. Baskın olmayan duyguların oranları ise 0 olarak ayarlanarak veri ön işleme süreci tamamlanmıştır. Oluşan nihai DataFrame Tablo 3.5'te sunulmuştur.

Tablo 3.5 Tüm ön işleme adımları sonrasında oluşan nihai dataframe

Participant id	Age	Female	Male	Experiment id	Lastfm tag	Artist	Title	neutral	...	surprise rate	Liking	label
1	31	0	1	5	exciting	Blur	Song 2	1	..	0.0	6.05	0
1	31	0	1	18	uplifting	Louis Armstrong	What A Wonderful World	1	..	0.0	8.01	0
1	31	0	1	4	fun	The B52'S	Love Shack	1	..	0.0	8.06	0
1	31	0	1	24	sad	James Blunt	Goodbye My Lover	1	..	0.0	8.22	2
1	31	0	1	20	romantic	Taylor Swift	Love Story	1	..	0.0	8.13	1
...	...	...	...	...	...	...	...	...	..	...	...	..
22	28	1	0	10	protest	Rage Against The Machine	Bombtrack	1	..	0.0	5.00	2
22	28	1	0	33	hate	Dead To Fall	Bastard Set Of Dreams	1	..	0.0	1.00	2
22	28	1	0	2	exciting	Lustra	Scotty Doesn't Know	1	..	0.0	7.12	2
22	28	1	0	25	depressing	A Fine Frenzy	Goodbye My Almost Lover	1	..	0.0	7.05	2
22	28	1	0	35	heavy	Napalm Death	Procrastination On The Empty Vessel	1	..	0.0	1.00	3

3.4.2 Öneri Sistemi Modelleme

Bu bölümde kullanıcı bazlı bir öneri sistemi geliştirilmiştir. Katılımcıların her bir müzik parçasına verdikleri “Liking” puanları ve EEG sinyalleri ve yüz analizi verilerinden füzyon yöntemiyle elde edilen “Label” bilgisi kullanılarak yapılmıştır. Bu etiketler, öneri sistemi için bir özellik olarak kullanılmıştır.

Öncelikle, veri seti katılımcıların müzik tercihleri ve duygusal tepkilerini temsil eden özelliklerin seçilmesiyle hazırlanmıştır. Seçilen veriler, kullanıcılar bazında gruplandırılarak her bir duygusal tepki için ortalama değerler hesaplanmıştır. Bu süreç, kullanıcının genel müzik tercihlerini ve duygusal durumunu yansıtan bir profil oluşturulmuştur. Ardından, StandardScaler ile bu özellikler ölçeklendirilmiştir ve kullanıcıların birbirleriyle olan benzerliklerini tespit etmek için Cosine Similarity (Kosinüs Benzerliği) ve K-en yakın komşu (KNN) metrikleri kullanılmıştır. Cosine Similarity, iki vektör arasındaki açının kosinüsünü alarak, bu vektörlerin birbirine olan benzerlikleri o kadar yüksektir. Bu yöntem özellikle metin ve öneri sistemlerinde, kullanıcı veya öğe profillerinin sayısal ifadelerini karşılaştırmak için sıklıkla kullanılır. Örneğin, kullanıcıların müzikal tercihlerini ve duygusal tepkilerini vektörler olarak ifade edebiliriz. Her bir kullanıcı için, beğendikleri sanatçılar, müzik türleri veya belirli parçalar, bir vektör içinde ilgili ağırlıklarla temsil edilir. İki kullanıcının müzik tercihleri arasındaki kosinüs benzerliği, bu vektörlerin birbirine ne kadar yakın olduğunu gösterir, bu da öneri sistemlerinde kullanıcılar arası ilişkileri belirlemede kritik bir rol oynar.

KNN, bir kullanıcının veya öğenin, veri setindeki diğer kullanıcılar veya öğeler ile olan benzerliğine dayanarak, en yakın 'K' adet komşusunu belirler. Bu yöntem, belirlenen 'K' değerine bağlı olarak, kullanıcının zevklerine ve tercihlerine en yakın olan diğer kullanıcıları sıralar. Öneri sistemlerinde, bu en yakın komşuların tercihleri, belirli bir kullanıcıya müzik önerileri yapmak için kullanılır. KNN, özellikle büyük veri setlerinde, belirli bir kullanıcının tercihlerine benzer kullanıcıları hızlı bir şekilde bulmak için tercih edilen bir yöntemdir. (Singh v.d.)

Bu benzerlik yöntemleri, öneri sisteminde kullanıcıların zevklerine en uygun yeni müzikler keşfetmelerine yardımcı olur ve kişiselleştirilmiş müzik deneyimleri sunar. Böylelikle kullanıcı memnuniyeti ve sistem etkinliği artırılmış olur. Bu bağlamda, müzik öneri sistemleri için iki temel benzerlik hesaplama yöntemi olan Cosine Similarity ve KNN yöntemleri karşılaştırmalı olarak incelenmiştir. İncelenen yöntemlerin kullanıcı tercihlerini tahmin etme kapasitesini ölçmek amacıyla tutarlılık kontrolü ve çapraz doğrulama teknikleri uygulanmıştır.

Öncelikle, kullanıcı verilerini içeren veri seti üzerinden seçilen özniteliklerle (duygusal tepkiler ve müzikleri beğenme skoru gibi) kullanıcılar gruplandırılmış ve her kullanıcı için özniteliklerin ortalaması alınmıştır. Bu öznitelikler, her iki yöntemin de temelini oluşturacak şekilde düzenlenmiştir. Her iki yöntemin performansı, kullanıcılara önerilen şarkıların kullanıcıların gerçekte verdiği 'Liking' puanları ile karşılaştırılarak ölçülmüştür. Tutarlılık kontrolü, öneri sistemi tarafından önerilen içeriklerin kullanıcıların verdiği puanlarla ne kadar uyumlu olduğunu gösterirken, çapraz doğrulama ile modelin genelleştirme yeteneği test edilmiştir. Çapraz doğrulama sürecinde, veri seti rastgele eğitim ve test setlerine ayrılmış, eğitim seti üzerinde model eğitilmiş ve test seti üzerinden modelin başarısı değerlendirilmiştir.

Sonuçlar, Cosine Similarity yönteminin, KNN yöntemine göre hem tutarlılık kontrolünde hem de çapraz doğrulama testlerinde daha yüksek puanlar elde ettiğini göstermiştir. Bu sonuç, Cosine Similarity'nin kullanıcıların puanlama davranışlarını daha iyi yansıttığını ve bu yöntemin öneri sistemlerinde kullanıcı tercihlerini tahmin etme konusunda daha başarılı olabileceğini ortaya koymuştur. DEAP veri seti için Cosine Similarity yönteminin daha başarılı olması sebebiyle bu yöntem kullanılarak çalışmalara devam edilmiştir.

Bahsedilen yönteme ait fonksiyonların sözde kodları aşağıda sunulmuştur. Veri yükleme ve ön işleme için kullanılan kod sayesinde veriler bir pandas DataFrame'ine yüklenmiş ve belirlenen öznitelikler seçilmiştir. Seçilen özellikler arasında kullanıcı ID'si, müzik ile ilgili etiketler, beğeni puanı ve duygusal tepkiler yer almaktadır. Daha sonra bu veriler kullanıcı ID'sine göre gruplanmıştır. Fonksiyon hem orijinal veri setini hem de gruplanmış verileri döndürür. Bu işlem, sonradan analiz aşamasında kullanılacak verilerin ön hazırlığını sağlar. Böylece kullanıcı bazında öneriler yapmak ve kullanıcıların benzerliklerini hesaplamak için veri seti uygun formata getirilmiş olur.

Function LoadAndPreprocessData(file_path):

Begin

1. Load data from the specified file path into a DataFrame.

data = Load data from file_path as DataFrame

2. Define the required features for analysis.

features = List of 'Participant_id', 'Lastfm_tag', 'Liking', 'label', 'neutral_rate', 'happy_rate', 'angry_rate', 'sad_rate', 'fear_rate', 'surprise_rate'

3. Create a new DataFrame selecting only the defined features from the data.

selected_data = Select columns specified in features from data

4. Group the selected data by 'Participant_id' and aggregate it.

grouped_data = Group selected_data by 'Participant_id' and aggregate such that:

- 'Liking' and 'label' columns are aggregated into lists
- Other columns are aggregated by their mean

5. Return both the original and the grouped data.

Return data, grouped_data

End

Sözde Kod 3.7 Veri yükleme ve ön işleme fonksiyonu

Benzerliklerin hesaplanması için kullanılan fonksiyonda ise, önceden gruplandırılmış veriler üzerinden kullanıcılar arası benzerlik hesaplaması yapılır. İlk olarak, StandardScaler kullanılarak duygusal tepkileri içeren özelliklerin normalleştirilmesi gerçekleştirilir. Daha sonra, her kullanıcının beğeni puanlarının ortalaması hesaplanır ve bu ortalama ölçeklendirilir. Ölçeklendirilmiş beğeni ortalaması ve diğer duygusal tepkiler kullanılarak Kosinüs Benzerlik matrisi hesaplanır. Hesaplanan benzerlik matrisi, uygun etiketlerle bir pandas DataFrame'ine dönüştürülür. Bu işlem, öneri sistemlerinde kullanıcıları benzer tercih ve duygusal tepkilere sahip diğer kullanıcılarla eşleştirmek için bir temel oluşturur.

Function CalculateSimilarity(grouped_data):

Begin

1. Initialize a StandardScaler object for normalizing data.

 scaler = Create a StandardScaler object

2. Normalize emotional rate features in grouped_data using the scaler.

 For each feature in emotional rate features:

 Normalize the feature in grouped_data using scaler

3. Calculate the average 'Liking' score for each user and scale it.

 Calculate the mean of 'Liking' for each participant and add it as 'Liking_avg' column in grouped_data

 Scale 'Liking_avg' using scaler and add it as 'Liking_avg_scaled' column in grouped_data

4. Compute the cosine similarity matrix for the features including the scaled average liking and emotional rates.

 similarity_matrix = Calculate cosine similarity for the DataFrame columns 'Liking_avg_scaled' and emotional rates

5. Convert the similarity matrix into a DataFrame with appropriate labels.

 similarity_df = Convert similarity_matrix into a DataFrame with labels from grouped_data indices

6. Return the similarity DataFrame.

 Return similarity_df

End

Sözde Kod 3.8 Kullanıcılar arası benzerlik hesaplayan fonksiyon

```

Function MakeRecommendation(data, grouped_data, similarity_df, test_user_id):
    Begin
        1. Sort the similarities of the test_user_id and find the top similar users.
            user_similarity = Sort similarity_df column for test_user_id in descending order
            similar_users = Get indices of the top 5 entries from user_similarity excluding the first one
        2. Initialize a list to store recommended songs.
            recommended_songs = Initialize an empty list
        3. For each similar user, find and add their highly liked songs to the recommendation list.
            For each user in similar_users:
                user_songs = Filter data where 'Participant_id' is the current user
                high_liked_songs = Filter user_songs where 'Liking' is greater than 7
                Extend recommended_songs with the 'Artist' and 'Title' from high_liked_songs
        4. Ensure the list contains unique song recommendations and select the top 10.
            recommended_songs = Convert recommended_songs to a set of tuples to ensure
            uniqueness and select the first 10
        5. Return the list of recommended songs.
            Return recommended_songs
    End

```

Sözde Kod 3.9 Öneri Yapma Fonksiyonu

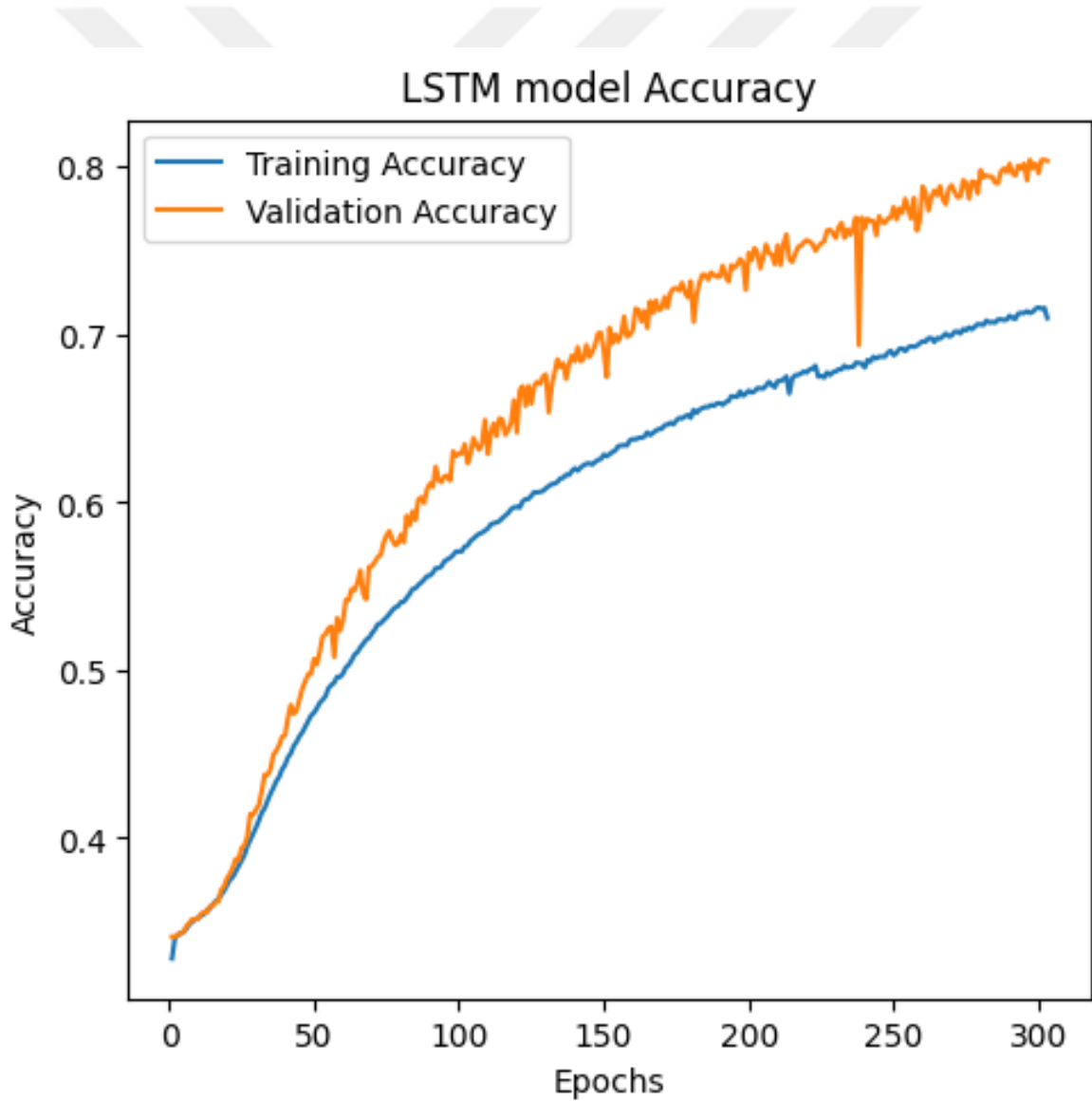
Öneri yapma fonksiyonunda ise belirli bir test kullanıcısı için müzik önerileri üretilir. İlk olarak belirtilen test kullanıcısına en benzer kullanıcıları bulmak için benzerlik DataFrame'i kullanılır. Benzerlik değerleri sıralanır ve en yüksek beş benzerlik skoruna sahip kullanıcılar seçilir. Ardından, bu benzer kullanıcılardan her birinin yüksek 'Liking' puanına sahip şarkıları belirlenir. Bu şarkılar, öneri listesine eklenir. İşlem sırasında şarkıların sanatçı ve şarkı ismi bilgileri kullanılarak eşsiz bir liste oluşturulur ve bu liste ilk on öneri ile sınırlanır. Son olarak bu liste fonksiyon tarafından döndürülür.

4. BULGULAR VE TARTIŞMA

Bu çalışmada, GRU, LSTM ve Transformer modellerinin sadece EEG verileri ve hem EEG hem de yüz ifadeleri verileri üzerindeki performansları incelenmiştir. Bulgular, her modelin doğruluk (accuracy) değerlerini, sınıflandırma raporlarını (classification report), karışıklık matrislerini (confusion matrix), kayıp (loss) ve doğruluk grafiklerini içermektedir.

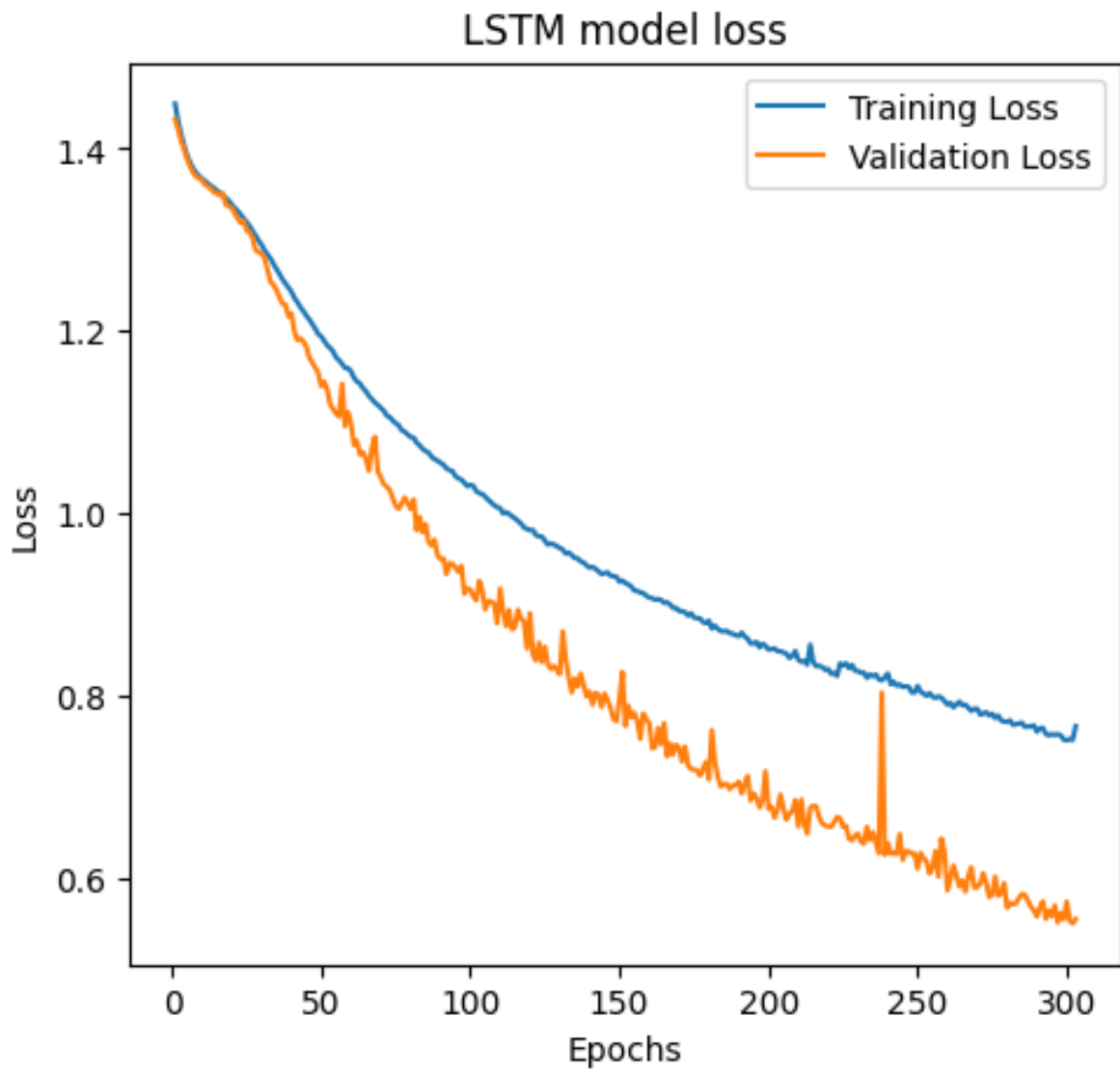
4.1 EEG Sinyalleri ile Duygu Sınıflandırma Sonuçları

Bu bölümde sadece EEG sinyalleriyle eğitilmiş olan LSTM, Transformer ve GRU modellerinin performansları değerlendirilmiş ve model sonuçları tartışılmıştır.



Şekil 4.1 Tek modlu LSTM modeline ait doğruluk(accuracy) grafiği

Şekil 4.1’deki doğruluk grafiği LSTM modelinin birleşik EEG ve yüz ifadelerinin eğitim ve validasyon veri setleri üzerindeki performansını her eğitim turu için göstermektedir. Başlangıçta düşük olan doğruluk, eğitim tur sayısı ilerledikçe artış göstermiştir. Eğitim (mavi çizgi) ve validasyon doğruluklarının birbirine yakın seyretmesi modelin aşırı uyuma gitmediğini belirtmektedir. Ancak, validasyonun doğruluk (turuncu çizgi) eğrisinde zaman zaman ani düşüslere sebep olan ani sıçramalar (spike) gözlemlenmiştir. Bu durum, modelin belirli eğitim turlarında validasyon seti üzerinde tutarsızlık gösterdiğine işaret edebilir. Genel olarak, model yüksek bir doğruluk oranı elde etmiş ve bu oran, kademeli bir şekilde artmıştır. Bu da modelin genelleştirme kapasitesinin iyi olduğunu göstermektedir.



Şekil 4.2 Tek modlu LSTM modeline ait kayıp(loss) grafiği

Şekil 4.2’deki kayıp grafiği, LSTM modelinin eğitim sürecinde nasıl geliştiğini ortaya koymaktadır. Kayıp, hem eğitim (mavi çizgi) hem de validasyon setleri için zamanla azalmaktadır, bu da modelin daha iyi tahminler yapmaya başladığını göstermektedir. Validasyon kaybı (turuncu çizgi) eğrisinde bazı ani sıçramalar mevcuttur. Bu tür ani sıçramalar, genellikle belirli veri noktalarının model tarafından iyi genellenemediğine veya eğitim sürecinin bazı aşamalarında modelin kararsızlık yaşadığına işaret etmektedir. Genel olarak, modelin kaybı düşmekte ve yüksek bir eğitim doğruluğuna ulaşmaktadır. Buradan modelin iyi bir öğrenme performansı sergilediği sonucuna ulaşılabilir.

Actual	HVHA	25299	895	1732	869
	HVLA	2231	13299	1026	878
	LVHA	2514	448	15463	912
	LVLA	2019	949	1804	12990
		Predicted			

Şekil 4.3 Tek modlu LSTM modeline ait karışıklık matrisi

Şekil 4.3, yalnızca EEG sinyalleri üzerinde eğitilen LSTM modelinin karmaşıklık matrisini sunmaktadır ve bu matris modelin test verileri üzerindeki tahminlerini göstermektedir. Bu matris, modelin tahmin ettiği sınıflar ile gerçek sınıflar arasındaki ilişkiyi yansıtmaktadır. Diyagonal üzerindeki değerler (matristeki daha koyu alanlar) doğru tahmin edilen sınıfları temsil ederken, diğer alanlar yanlış tahminleri temsil etmektedir. Model, HVHA sınıfı için 25299 doğru tahmin ile yüksek bir doğruluk oranı elde etmiştir. Bu sınıf için yanlış negatif sayısı nispeten düşük olmasına rağmen, dikkate değer yanlış pozitif tahminler bulunmaktadır. HVLA sınıfı için 13299 doğru tahmin yapılmıştır, ancak bu sınıfta HVHA sınıfına kıyasla yanlış negatif tahminlerde bir artış görülmektedir. LVHA sınıfı için 15463 tahmin ile iyi bir doğruluk oranı elde edilmiştir. Son olarak, LVLA sınıfı, diğer sınıflara kıyasla daha düşük bir doğruluk oranına ulaşmıştır ve 12990 doğru tahmin yapılmıştır. Bu sınıf için test verilerindeki örnek sayısının düşük olması, doğru tahmin sayısını etkileyen bir faktör olarak da düşünülebilir. LVLA sınıfı, test verilerinde HVLA sınıfından sonra en az örnek sayısına sahip ikinci sınıftır. Sınıflar arasındaki bu karmaşa, modelin bazı sınıfları birbirinden ayırt etmekte zorlandığını göstermektedir. Özellikle, LVHA ve LVLA sınıfları sıklıkla model tarafından birbirleriyle karıştırılmaktadır.

Tablo 4.1 Tek modlu LSTM modeline ait sınıflandırma raporu

	precision	recall	f1-score	support
HVHA	0.79	0.88	0.83	28795
HVLA	0.85	0.76	0.94	17434
LVHA	0.77	0.80	0.79	19337
LVLA	0.83	0.73	0.78	17762
accuracy			0.80	83328
macro_accuracy	0.81	0.79	0.80	83328
weighted_accuracy	0.81	0.80	0.80	83328

Tablo 4.1’de sadece EEG verileriyle eğitilmiş olan LSTM modelinin test verisi üzerindeki sınıflandırma raporu sunulmuştur. Precision (kesinlik) (*Denklemler IX*) (Buckland v.d.) metriği, modelin her bir sınıfı ne kadar doğru tahmin ettiğini oransal olarak göstermektedir.

Bu oranlar, HVHA için %79, HVLA için %85, LVHA için %77 ve LVLA için %83'tür. Recall (duyarlılık) (*Denklem X*) metriği (Buckland v.d.), gerçek pozitif durumların ne kadarının doğru tahmin edildiğini göstermektedir. Bu metrik değerleri HVHA için %88, HVLA için %76, LVHA için %80 ve LVLA için %73 olarak hesaplanmıştır. F1-skoru (*Denklem XI*) (Powers v.d.), precision ve recall metriklerinin harmonik ortalamasıdır ve modelin her iki metriktaki dengesini yansıtır. F1-skorları sırasıyla HVHA için %83, HVLA için %81, LVHA için %79 ve LVLA için %78'dir. Support ise her sınıfın veri setindeki örnek sayısını temsil eder ve her sınıf için değerlendirmenin temelini oluşturur.

$$Precision = \frac{True\ Positive(TP)}{True\ Positive(TP) + False\ Positive(FP)}$$

(*Denklem IX*)

$$Recall = \frac{True\ Positive\ (TP)}{True\ Positive(TP) + False\ Negative(FN)}$$

(*Denklem X*)

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

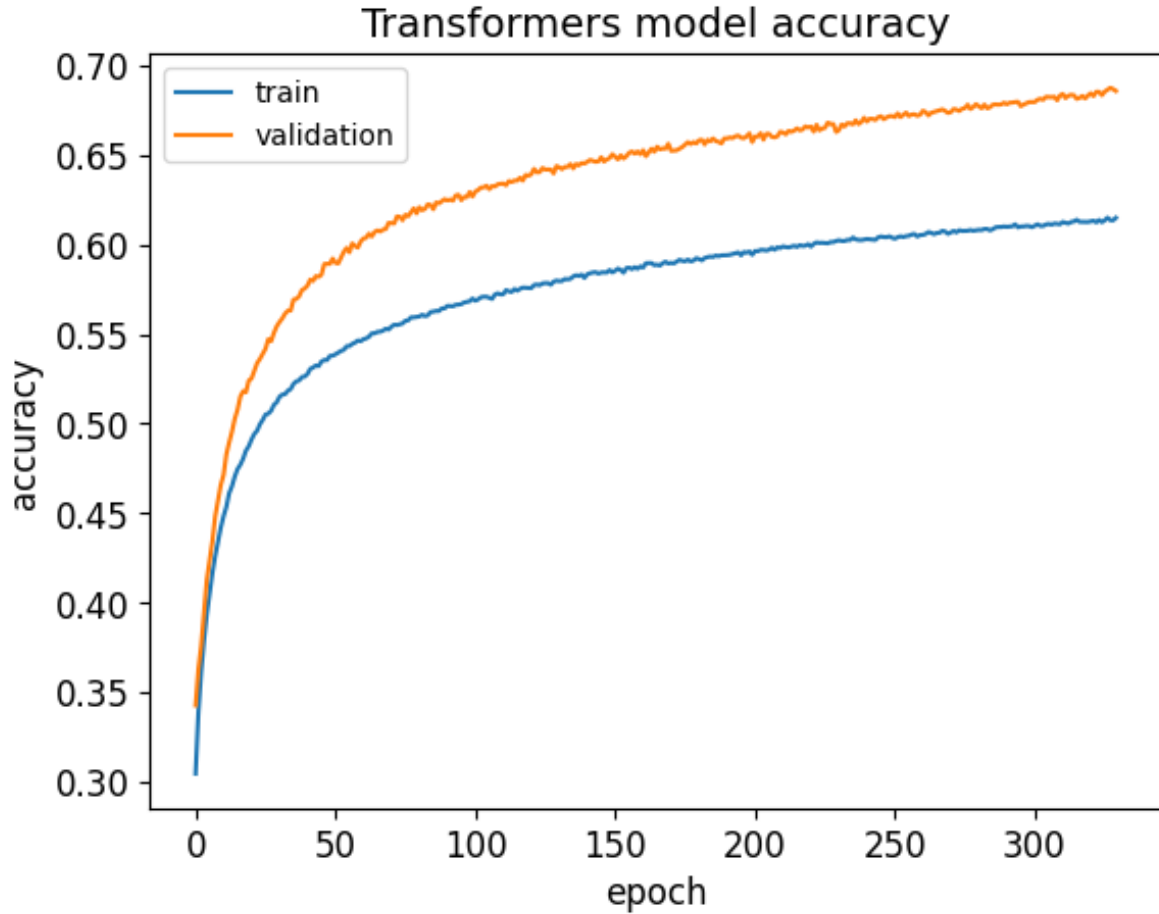
(*Denklem XI*)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

(*Denklem XII*)

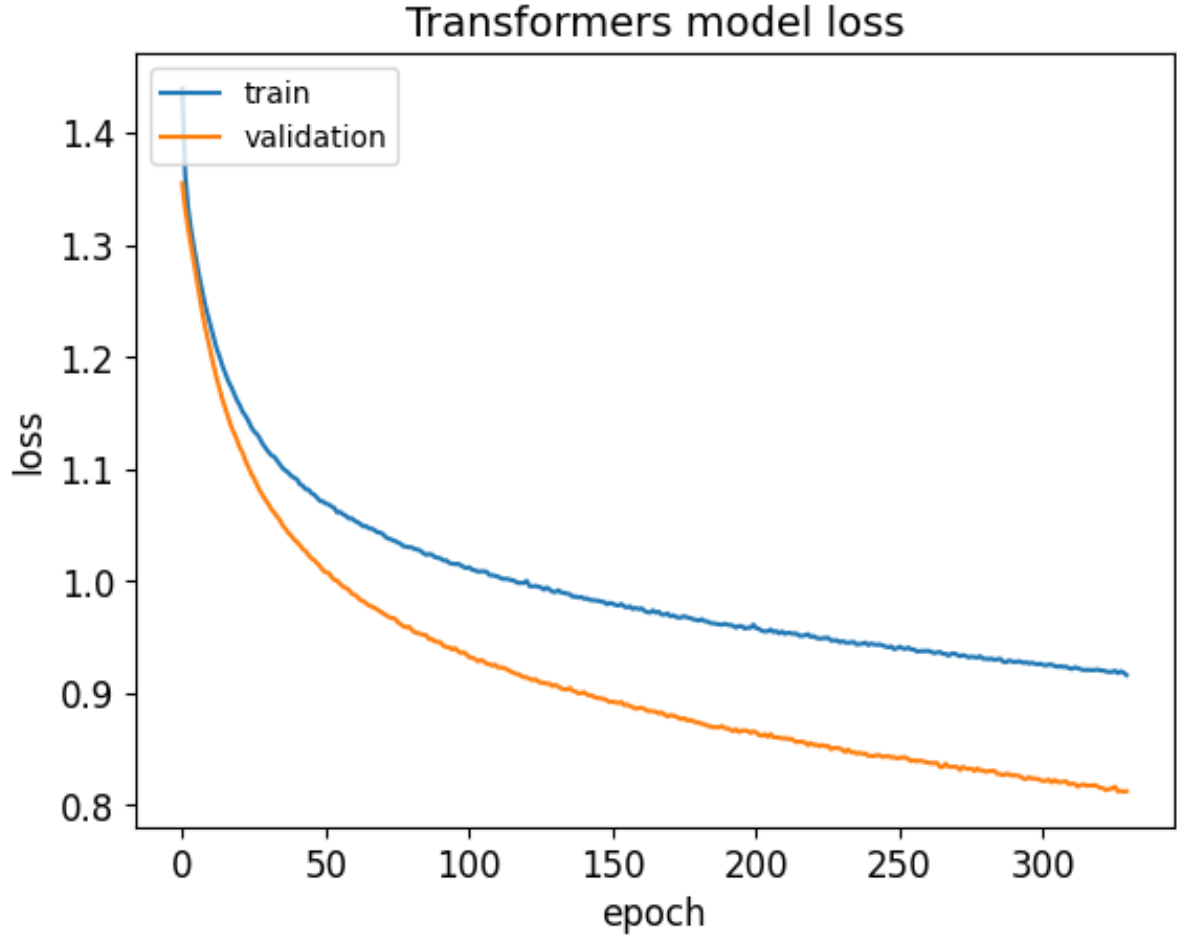
Genel olarak, %80 genel doğruluk (accuracy) (*Denklem XII*) (Tharwat) ile model etkili bir performans sergilemiştir. Makro (macro_accuracy) ve ağırlıklı doğrulukların (weighted_accuracy) ortalamaları %80 civarında olup, modelin tüm sınıflar üzerinde dengeli bir performans sergilediğini göstermektedir.

Bu sonuçlar, modelin başarılı olduğunu gösterse de bazı sınıflar arasındaki karışıklıkların ve düşük recall oranlarının giderilmesi için farklı algoritma yaklaşımları (Transformer ve GRU) denenmiştir.



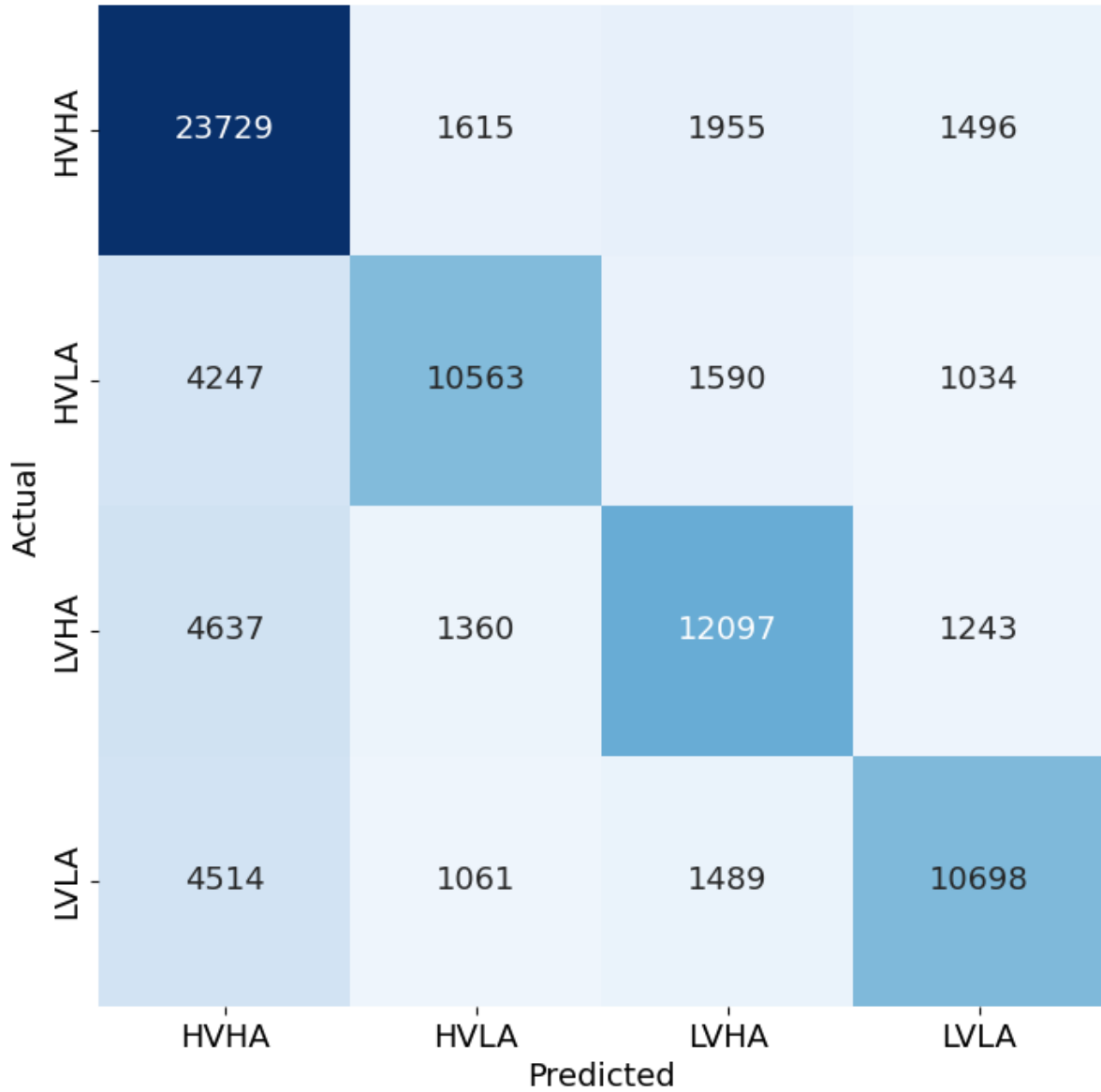
Şekil 4.4 Tek modlu Transformer modeline ait doğruluk grafiği

Şekil 4.4’te, sadece EEG sinyalleri üzerinde eğitilmiş olan Transformer modelinin doğruluk grafiği gösterilmektedir. Modelin eğitim doğruluğu başlangıçta keskin bir artış göstermiş ve 50’nci eğitim turundan itibaren daha yavaş bir artış eğilimine geçmiştir. Bu durum, modelin eğitim setindeki belirli örüntüleri hızla öğrendiğini ve belli bir seviyede doygunluğa ulaştığını ifade etmektedir. Doygunluk noktasından sonra eğitim doğruluğu az da olsa artmaya devam etmiştir. Bu da modelin veriden öğrenmeye devam ettiğini göstermektedir. Validasyon doğruluğu, eğitim doğruluğu ile benzer bir artış deseni göstermiş ve eğitim sonlarına doğru düzleşmiştir. Eğitim ve validasyon doğruluk oranları arasında küçük bir fark mevcuttur. Bu fark yüksek olmadığından modelin önemli ölçüde aşırı uyum göstermediğine işaret etmektedir.



Şekil 4.5 Tek modlu Transformer modeline ait kayıp grafiği

Şekil 4.5'teki kayıp grafiğinde, eğitim kaybı eğrisi ilk eğitim turlarında hızlı bir düşüş gerçekleştirmiştir. Bu durum, doğruluk grafiğine paralel olarak modelin başlangıçta eğitim verilerinden hızlı bir şekilde öğrendiğine işaret etmektedir. Sonrasında, kayıp eğrisi düşmeye devam etmiş, ancak kayıp hızı azalarak stabil bir eğilime bürünmüştür. Bu, modelin öğrenme sürecinin devam ettiğini, ancak daha yavaş bir hızda ilerlediğini göstermektedir. Validasyon kaybı da eğitim kaybına benzer bir düşüş eğilimi göstermiş ve eğitim kaybının biraz altında daha iyi bir azalma sergilemiştir. Genel olarak, her iki kayıp eğrisi de birbirine yakın ilerlemiş ve büyük bir ayrılma (divergence) görülmemiştir.



Şekil 4.6 Tek modlu Transformer modeline ait karışıklık matrisi

Şekil 4.6’da Transformer modelinin DEAP veri setindeki EEG sinyalleri test veri seti üzerindeki performansına ait karışıklık matrisi sunulmuştur. Karışıklık matrisinden, HVHA sınıfı için 23729 doğru tahminle en yüksek doğruluk oranına ulaşıldığı gözlemlenmektedir. Ancak bu sınıf için, yanlış negatif (gerçekte HVHA olup farklı olarak tahmin edilenler) ve yanlış pozitif sayıları (diğer sınıfların HVHA olarak yanlış sınıflandırılması) da dikkate değer seviyelerdedir.

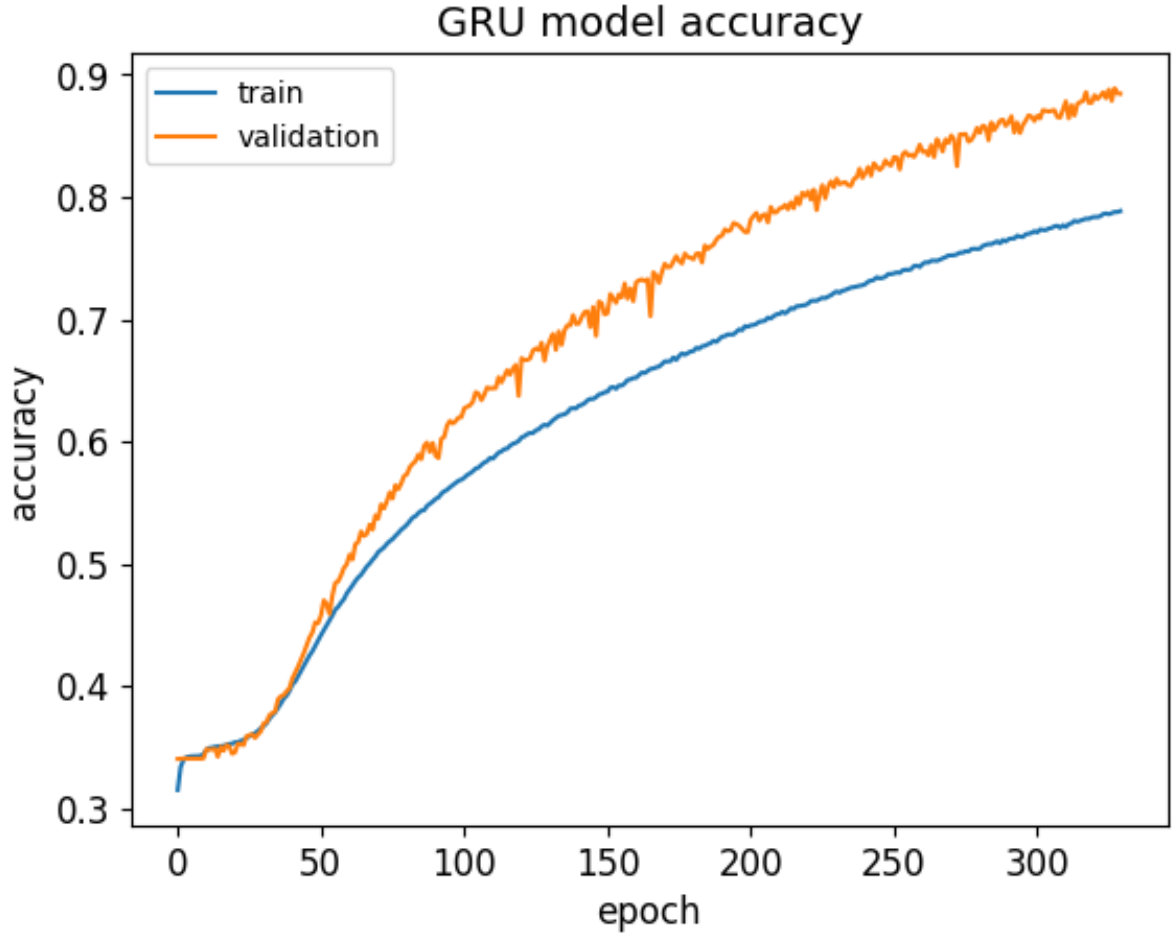
HVLA sınıfı için 10563 doğru tahmin ile önemli bir doğruluk oranı elde edilmiştir. Buna karşın, bu sınıf için de yanlış pozitif ve yanlış negatif sayıları Şekil 4.3'deki LSTM modelinin sonuçlarına kıyasla yüksektir. LVHA sınıfı için 12097 ve LVLA sınıfı için 10698 doğru tahmin yapılmıştır. Ancak, her iki sınıf için de yanlış pozitifler ve yanlış negatifler belirgin düzeydedir.

Tablo 4.2 Tek modlu Transformer modeline ait sınıflandırma raporu

	precision	recall	f1-score	support
HVHA	0.64	0.82	0.72	28795
HVLA	0.72	0.61	0.66	17434
LVHA	0.71	0.63	0.66	19337
LVLA	0.74	0.60	0.66	17762
accuracy			0.69	83328
macro_accuracy	0.70	0.66	0.68	83328
weighted_accuracy	0.69	0.69	0.68	83328

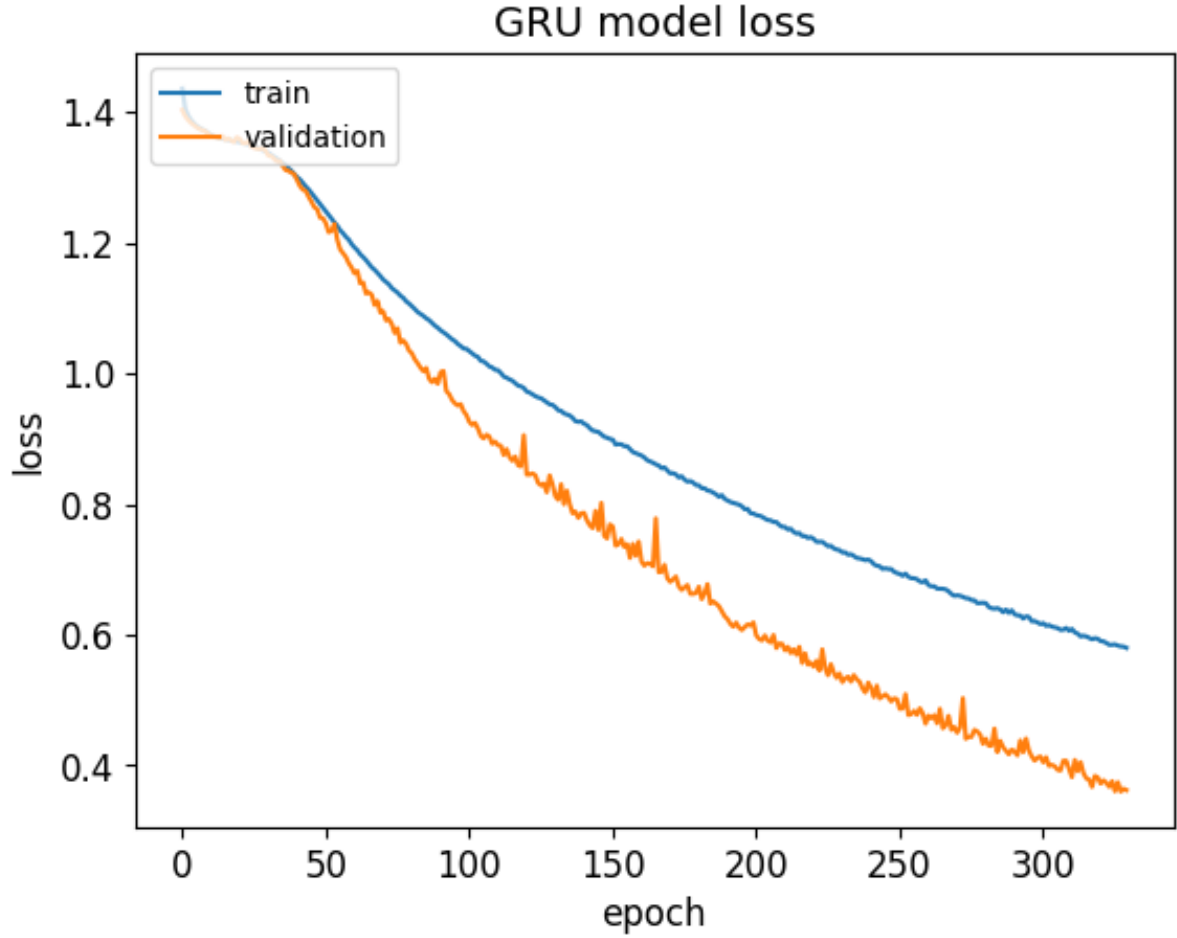
Tablo 4.2'de Transformer modelinin sınıflandırma raporuna göre precision değerleri HVHA sınıfı için %64, HVLA sınıfı için %72, LVHA sınıfı için %71 ve LVLA sınıfı için %74 olarak ölçülmüştür. Recall oranları, HVLA sınıfı için %82, HVHA sınıfı için %61, LVHA sınıfı için %63 ve LVLA sınıfı için %60 olarak hesaplanmıştır. F1-Skorları ise HVHA sınıfı için %72, HVLA sınıfı için %66, LVHA sınıfı için %66 ve LVLA sınıfı için %66'dır. Genel olarak, model %69'luk bir doğruluk ile orta düzeyde bir performans sergilemiştir. Makro ve ağırlıklı doğruluk değerleri %68 ile %70 arasında değişmektedir. Bu durum modelin tutarlı olduğunu ancak çok iyi bir performans sergilemediğini göstermektedir.

Özetle, modelin genel doğruluğu ve F1-skorları, Transformer mimarisinin EEG verilerinin karmaşıklığıyla başa çıkabileceğini göstermiştir. Buna karşın, modelin yüksek yanlış pozitif ve yanlış negatif oranları modelleme sürecinin optimize edilmesi gerektiğine işaret etmektedir.



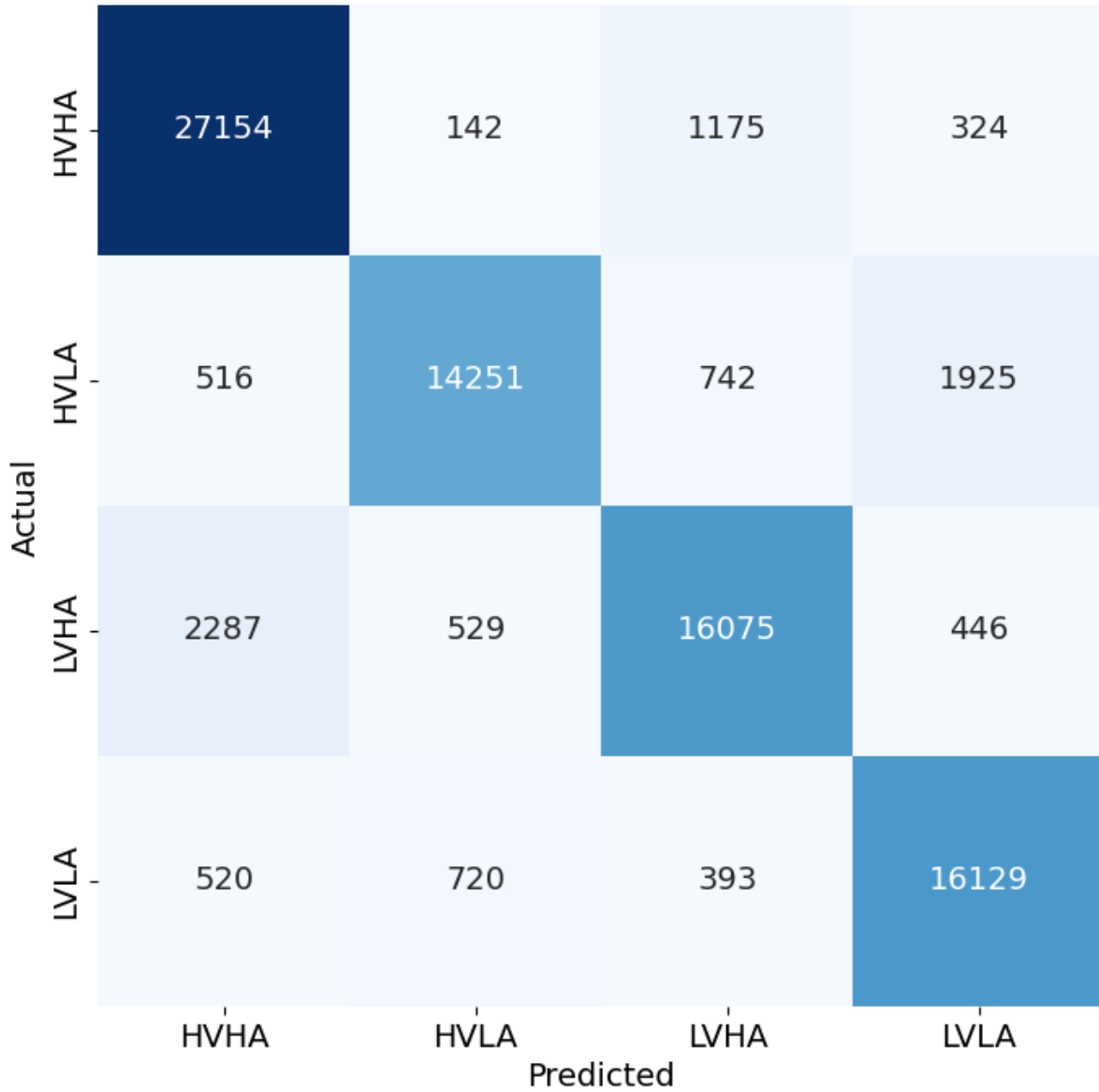
Şekil 4.7 Tek modlu GRU modeline ait doğruluk grafiği

Şekil 4.7'deki GRU modelinin doğruluk grafiği incelendiğinde, eğitim doğruluğu (mavi çizgi) başlangıçta hızlı bir şekilde artmaktadır. Özellikle ilk 100 eğitim turunda hızlı bir artış gözlemlenmiştir. 100'üncü eğitim turundan itibaren doğruluk artış hızı azalmış olsa da eğitim doğruluk oranı sürekli olarak artmış ve 300'üncü eğitim turunda %90 civarına erişmiştir. Bu durum, modelin eğitim verilerinden öğrenmeye devam ettiğini ve iyi bir uyum sağladığını göstermektedir. Bunun yanında, validasyon doğruluğu (turuncu çizgi) benzer bir eğilimle başlangıçta hızlı bir artış göstermiş ve 200'üncü eğitim turundan itibaren artış hızı azalmış olsa da dengeli bir şekilde eğitim doğruluğunun biraz üstünde artmaya devam etmiştir.



Şekil 4.8 Tek modlu GRU modeline ait kayıp grafiği

Şekil 4.8’de sunulan GRU modelinin kayıp grafiğinde, ilk 10 eğitim turunda modelin eğitim ve validasyon kaybında hızlı bir düşüş gözlemlenmiştir. Bu hızlı düşüş, modelin eğitim setindeki örüntüleri ve ilişkileri hızlı bir şekilde tespit ettiğini ve bunlara uyum sağladığını göstermektedir. 10’uncu ve 40’inci eğitim turları arasında kayıplarda bir plato olduğu görülmüştür. Bu da modelin bu aşamalarda belirli bir öğrenme doygunluğuna ulaştığını gösterir. 35’inci ve 80’inci eğitim turları aralığında, kayıplarda yeniden hızlı bir düşüş gerçekleşmiştir. Bu durum, modelin yeni örüntüler keşfettiğine veya mevcut öğrenmeyi derinleştirerek daha ince ayarlar yaptığına işaret etmektedir. 80’inci eğitim turu sonrası eğitim ve validasyon kayıplarındaki düşüş hızı azalmış, ancak kayıp düşmeye devam etmiştir. Bu da modelin artık öğrenme sürecinin daha incelikli bir aşamasına geçtiğini ve daha karmaşık veya daha az belirgin öğrenme kalıplarını öğrenmeye başladığını göstermektedir. Eğitim ve validasyon kayıp eğrileri arasındaki yakın fark ve stabil uyum, modelin aşırı uyum yapmadan veri setinin genel özelliklerini başarılı bir şekilde yakalayabildiğini ortaya koymaktadır.



Şekil 4.9 Tek modlu GRU modeline ait karışıklık matrisi

Şekil 4.9'da karışıklık matrisi GRU modelinin test veri seti üzerindeki sınıflandırma sonuçlarını göstermektedir. Bu matrise göre, HVHA sınıfı için 27154 doğru tahminle en yüksek doğruluk elde edilmiştir. HVLA sınıfı için 14251 doğru tahmin yapılmıştır. Bu sonuçlar, iyileştirme gereksinimine işaret etmekle birlikte yüksek bir doğruluk oranı sunmaktadır. LVHA ve LVLA sınıfları için sırasıyla 16075 ve 16129 doğru tahmin ile modelin her iki sınıfı da etkili bir şekilde tanıdığı gözlemlenmiştir.

Yanlış pozitiflerin ve yanlış negatiflerin LSTM ve Transformer modellerine kıyasla daha düşük sayıda olması, GRU modelinin test veri setindeki sınıfları etkili bir şekilde ayırt ettiğini göstermektedir.

Tablo 4.3 Tek modlu GRU modeline ait sınıflandırma raporu

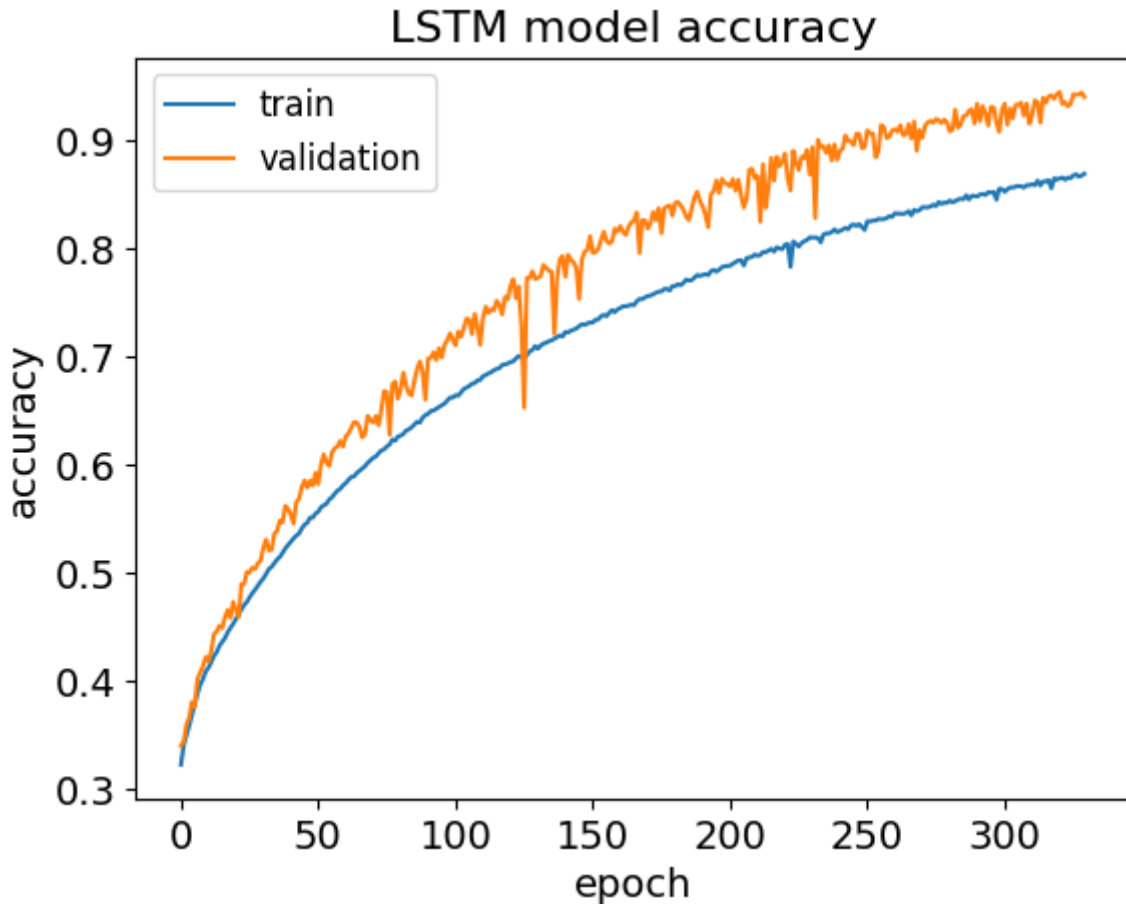
	precision	recall	f1-score	support
HVHA	0.89	0.94	0.92	28795
HVLA	0.91	0.82	0.86	17434
LVHA	0.87	0.83	0.85	19337
LVLA	0.86	0.91	0.88	17762
accuracy			0.88	83328
macro_accuracy	0.88	0.87	0.88	83328
weighted_accuracy	0.88	0.88	0.88	83328

Tablo 4.3’de GRU modelinin sınıflandırma raporuna göre precision değerleri HVHA için %89, HVLA için %91, LVHA için %87 ve LVLA için %86 olarak gerçekleşmiştir. Bu yüksek precision değerleri, modelin pozitif tahminlerinin büyük çoğunluğunun doğru olduğunu göstermektedir. Recall oranları ise HVHA için %94, HVLA için %82, LVHA için %83 ve LVLA için %91 olarak belirlenmiştir. Modelin F1-skorları HVHA için %92, HVLA için %86, LVHA için %85 ve LVLA için %88’dir. Bu F1-skorları, modelin recall ve precision metriklerinde iyi bir performans gösterdiğine işaret etmektedir. Modelin genel doğruluk, makro ve ağırlıklı doğruluk oranları her üç ölçüm metriği için de %88 olarak ölçülmüştür. Bu sonuçlar, modelin tüm sınıflarda yüksek performansla çalıştığını ortaya koymaktadır.

Sonuçlar, GRU modelinin genel anlamda karmaşık veri yapılarına sahip EEG sinyallerini başarılı bir şekilde işleyebildiğini göstermektedir. Sadece EEG verileriyle eğitilen GRU modeli, LSTM ve Transformer modelleriyle kıyaslandığında en iyi performansı sergilemiş ve EEG verileriyle çalışma konusundaki potansiyelini ortaya koymuştur. Ancak, daha iyi sonuçlara ulaşmak ve başarıyı artırmak amacıyla modeller çeşitli veri türleriyle yeniden eğitilmiş ve test edilmiştir. Bunlara ait sonuçlar, Bölüm 4.2’de daha detaylı bir şekilde açıklanmıştır.

4.2 Hem EEG Sinyalleri Hem Yüz İfadeleri ile Duygu Sınıflandırma Sonuçları

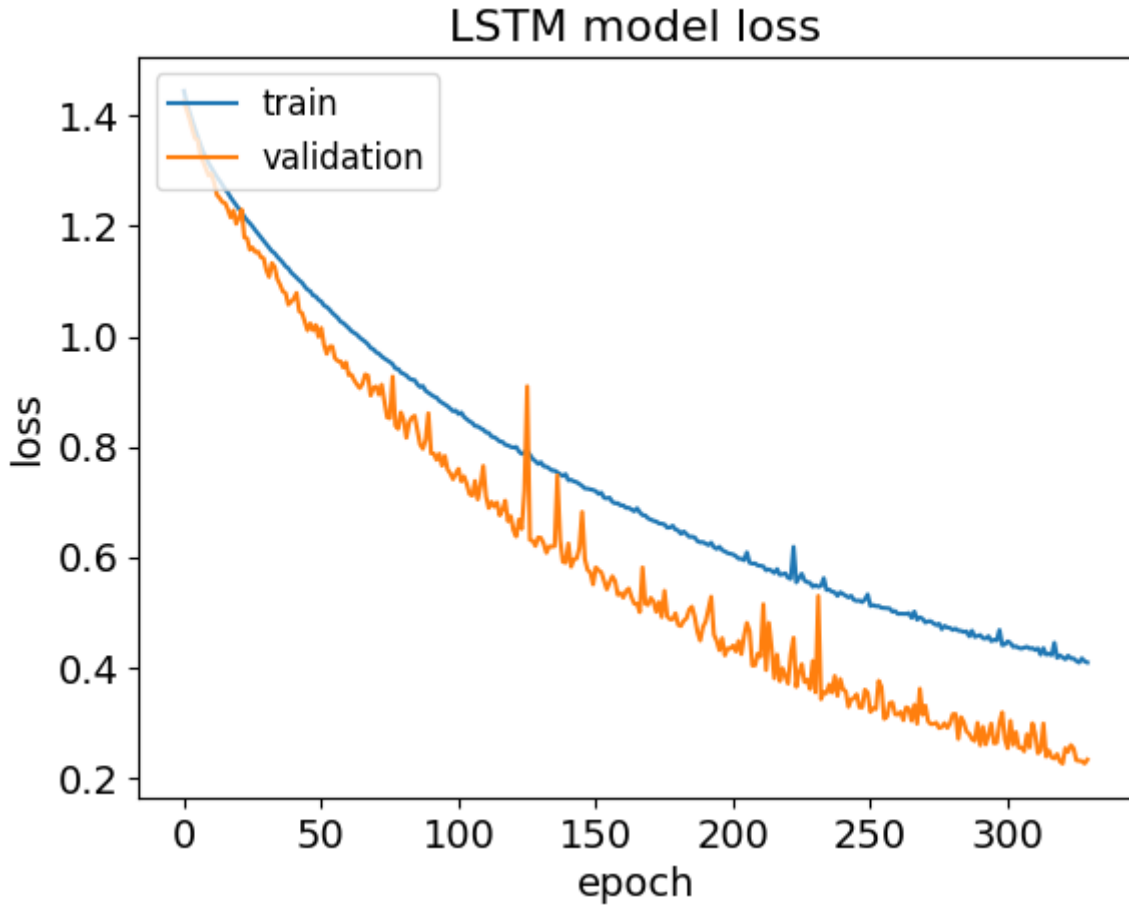
Bu bölümde birleştirilmiş EEG ve yüz ifadeleri verileriyle eğitilmiş olan LSTM, Transformer ve GRU modellerinin performansları değerlendirilmiş ve sonuçlar tartışılmıştır.



Şekil 4.10 Çok modlu LSTM modeline ait doğruluk grafiği

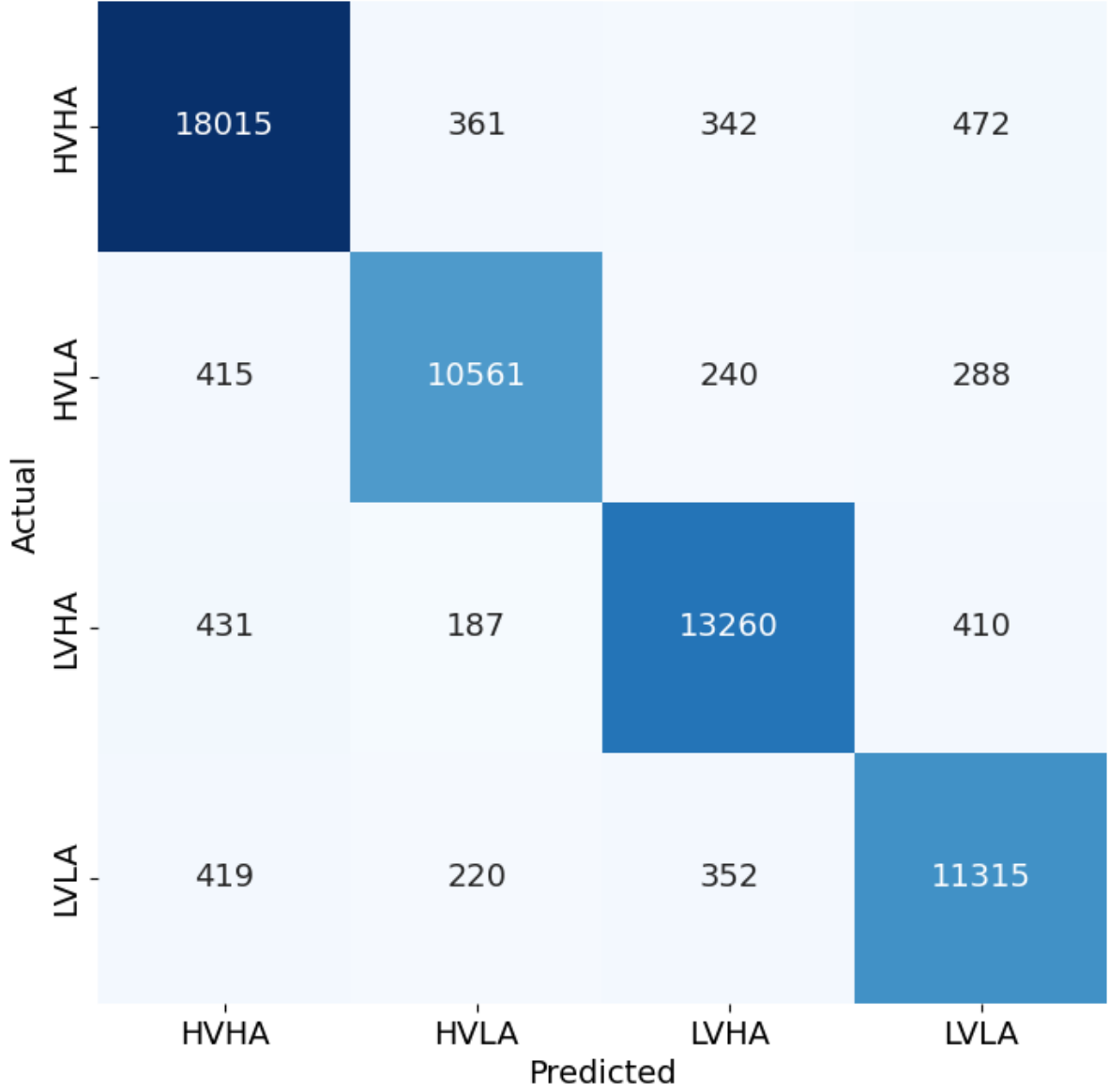
Şekil 4.10'daki grafik, LSTM modelinin birleşik EEG ve yüz ifadelerinin eğitim ve doğrulama veri setleri üzerindeki performansını her eğitim turu için göstermektedir. Başlangıçta düşük olan doğruluk, eğitim tur sayısı ilerledikçe artmıştır. Eğitim ve doğrulama doğruluklarının birbirine yakın seyretmesi modelin aşırı uyuma gitmediğini belirtmektedir. Ancak, validasyon doğruluğunun zaman zaman eğitimin doğruluğunun altına düştüğü bazı ani sıçramalar mevcuttur.

Bu durum, modelin belirli eğitim turlarında doğrulama seti üzerinde tutarsızlık gösterdiğine işaret edebilir. Genel olarak, model yüksek bir doğruluk oranı elde etmiş ve bu oran, eğitim sonuna kadar kademeli bir şekilde artmıştır.



Şekil 4.11 Çok modlu LSTM modeline ait kayıp grafiği

Şekil 4.11'deki kayıp grafiğine göre hem eğitim hem de validasyon setleri için kayıp zamanla azalmıştır. Bu da modelin daha iyi tahminler yapmaya başladığını göstermektedir. Validasyon kaybı eğrisinde modelin eğitim sürecinin bazı aşamalarında kararsızlık yaşadığını gösteren bazı ani sıçramalar mevcuttur. Genel olarak, model kaybı düşürmekte ve yüksek bir eğitim doğruluğuna ulaşmaktadır. Bu durum, LSTM modelinin iyi bir öğrenme performansı elde ettiğini gösterir.



Şekil 4.12 Çok modlu LSTM modeline ait karışıklık matrisi

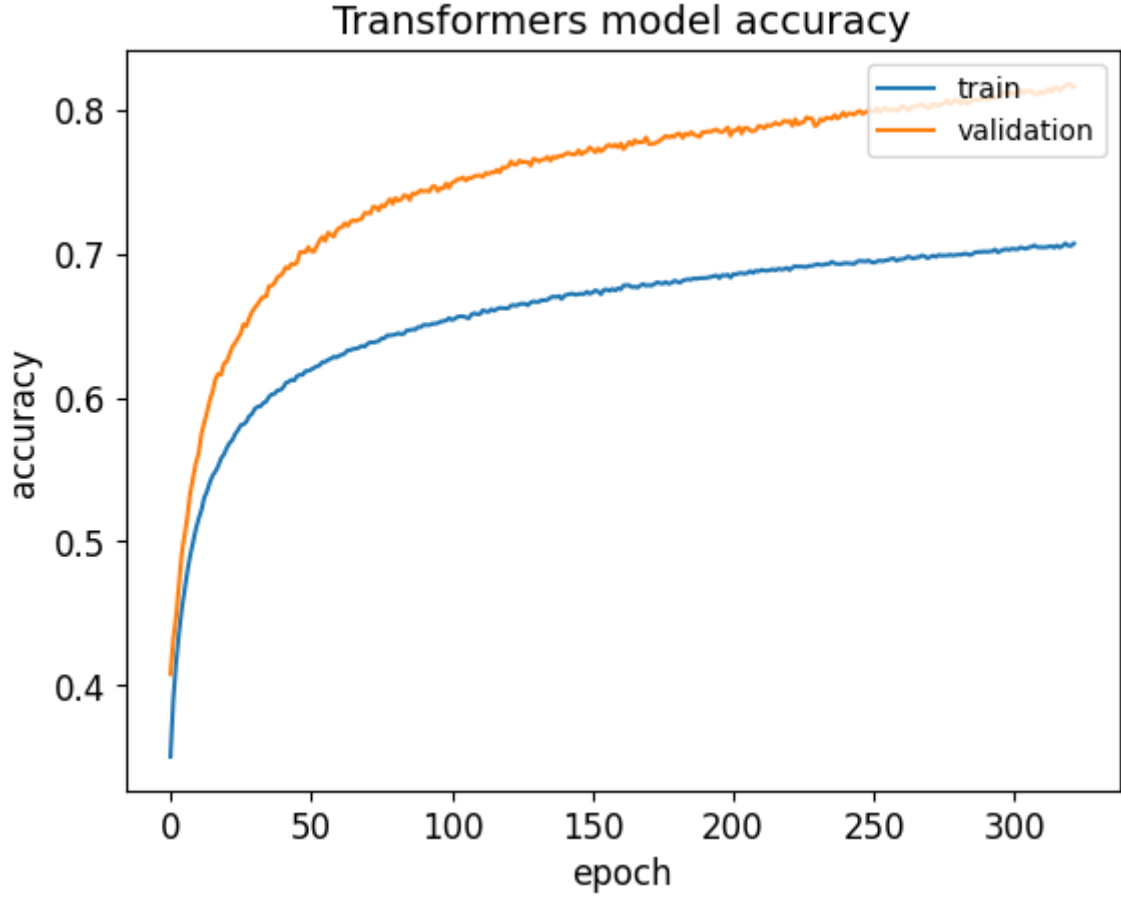
Şekil 4.12’de LSTM modelinin birleşik EEG ve yüz ifadeleri test verisi üzerinde yaptığı tahminlere ait karışıklık matrisi incelendiğinde, HVHA sınıfı için 18123 tahminle yüksek doğruluk gözlemlenmektedir. HVLA sınıfı için model 10,791 doğru tahminde bulunmuş ve bu sınıf için de yüksek bir doğruluk oranı elde etmiştir. LVHA ve LVLA sınıflarında da model benzer performans sergilemiş, sırasıyla 13,624 ve 11,602 doğru tahmin yapmıştır. Karışıklık matrisindeki yanlış pozitif ve yanlış negatif sayıları oldukça azdır. Bu durum, modelin her bir sınıfı dengeli bir şekilde tahmin ettiğini göstermektedir.

Tablo 4.4 Çok modlu LSTM modeline ait sınıflandırma raporu

	precision	recall	f1-score	support
HVHA	0.96	0.94	0.95	19190
HVLA	0.93	0.94	0.94	11504
LVHA	0.94	0.95	0.95	14288
LVLA	0.94	0.94	0.94	12306
accuracy			0.95	57288
macro_accuracy	0.94	0.94	0.94	57288
weighted_accuracy	0.95	0.95	0.95	57288

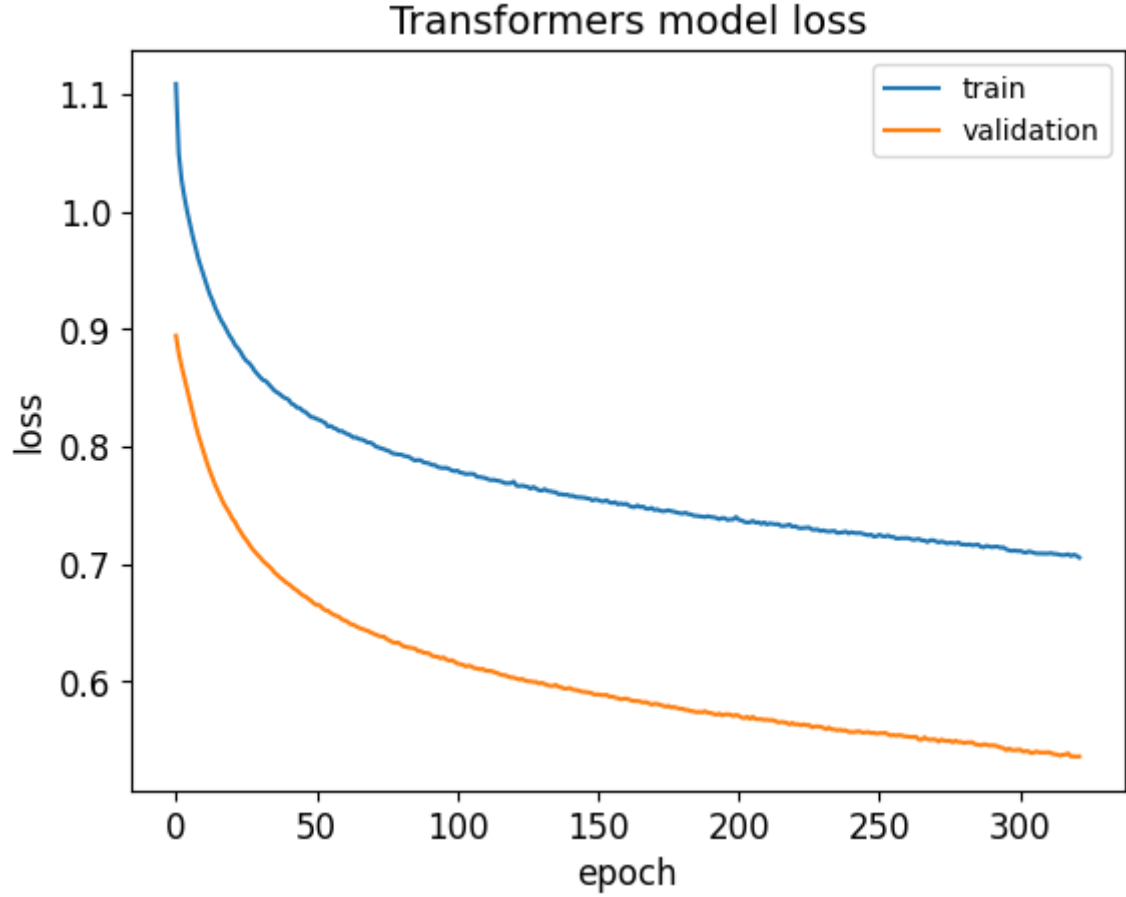
Tablo 4.4'te, birleşik EEG ve yüz ifadeleri ile eğitilmiş LSTM modeline ait sınıflandırma raporu sunulmaktadır. Precision metriği tüm sınıflar için %93 ile %96 arasında değişmekte ve modelin yanlış pozitif oranının düşük olduğunu göstermektedir. Recall oranları %94 ile %95 arasında değişmektedir. Bu da, modelin çok sayıda gerçek pozitif değeri başarıyla tahmin ettiğini göstermektedir. Tüm sınıflar için F1-Skoru ise %94 ile %95 arasında seyretmiş ve modelin yüksek performanslı olduğunu ortaya koymuştur.

Model %95'lik bir genel doğruluk ile başarılı bir performans sergilemiştir. Makro ve ağırlıklı doğruluklar da %94 ve %95 ile tutarlı ve yüksek değerlerdir. Bu sonuçlar, modelin bireysel sınıflar arasında büyük bir performans farkının olmadığını göstermektedir. Bölüm 4.1'deki Tablo 4.1'de sunulan tek modlu LSTM yaklaşımının sınıflandırma raporu ile kıyaslandığında, çok modlu verilerle çalışan LSTM modelinin accuracy, precision, recall ve F1-score metriklerindeki başarı oranlarının arttığı görülmektedir. Yüksek kesinlik, duyarlılık ve F1-skorları, modelin duygusal durumları sınıflandırmada tutarlı sonuçlar ürettiğini ortaya koymuştur. Bu da çok modlu verilerin duygu sınıflandırmada model başarımına olumlu yönde etki ettiğini göstermektedir.



Şekil 4.13 Çok modlu Transformer modeline ait doğruluk grafiği

Şekil 4.13’de Transformer modelinin doğruluk grafiği, modelin eğitime başlamasıyla doğruluğun hızlı bir şekilde arttığını göstermektedir. Model ilk eğitim turlarının ardından eğitim ve validasyon setleri üzerinde %50’nin üzerinde bir doğruluk değerine ulaşmıştır. Eğitim doğruluk oranı, eğitim süresince sürekli artmış ve yaklaşık %70 doğrulukla eğitimi tamamlamıştır. Bu durum, modelin eğitim verilerinden öğrenmeye devam ettiğini belirtmektedir. Validasyon doğruluk oranı da eğitim doğruluğuna benzer bir artış göstermiş ve yaklaşık %82 civarında bir doğruluk oranına ulaşmıştır. Eğitim süresince, eğitim ve doğrulama doğrulukları arasındaki boşluğun (gap) belirli bir seviyede olması modelin aşırı uyuma gitmiş olabileceğine işaret etmektedir. Buna karşın, her iki doğruluk eğrisi de birbirine yakın ilerlemiş ve büyük bir ayrışma görülmemiştir. Bu da aşırı uyumun olmadığını veya önemsiz seviyelerde kaldığını gösterir.



Şekil 4.14 Çok modlu Transformer modeline ait kayıp grafiği

Şekil 4.14'teki Transformer modelinin kayıp grafiğine göre eğitim kaybı başlangıçta hızlı bir düşüş göstermiş ve sonraki eğitim turlarında kayıp hızı yavaşlasa da azalmaya devam etmiştir. Kaybın azalmaya devam etmesi, modelin örüntüleri tanıdığını ve bunları daha iyi tahmin etmeye başladığını göstermektedir. Kaybın düşüş hızının zamanla azalması ise modelin başlangıçtaki hızlı iyileşmenin ardından eğitim verisinden öğrenilecek daha az yeni bilgi bulduğuna işaret etmektedir. Validasyon kaybı da eğitim kaybına paralel olarak azalmaya devam etmiş ve eğitim kaybının biraz altında eğitim sürecini tamamlamıştır.

Actual	HVHA	16645	584	892	1069
	HVLA	1175	8997	482	850
	LVHA	1405	410	11633	840
	LVLA	1544	567	697	9498
		HVHA	HVLA	LVHA	LVLA
		Predicted			

Şekil 4.15 Çok modlu Transformer modeline ait karışıklık matrisi

Şekil 4.15'teki Transformer modelinin karışıklık matrisine göre HVHA sınıfı için 16645 doğru tahmin ile en yüksek doğru sınıflandırma sayısına ulaşılmıştır. Bu sınıf için en yüksek yanlış negatif sayısı LVLA sınıf etiketinde kaydedilmiştir. HVLA sınıfı için 8997 doğru sınıflandırma ile HVHA sınıfına kıyasla daha düşük bir değer elde edilmiştir. Bu düşüklüğe test verisindeki bu sınıfa ait örnek sayısının azlığı da bir faktör olarak gösterilebilir. LVHA etiketi için 11633 doğru tahmin ile dikkate değer bir doğruluğa ulaşılmıştır. Ancak, bu sınıftaki yanlış negatif sayısı HVLA sınıfına nazaran daha yüksektir.

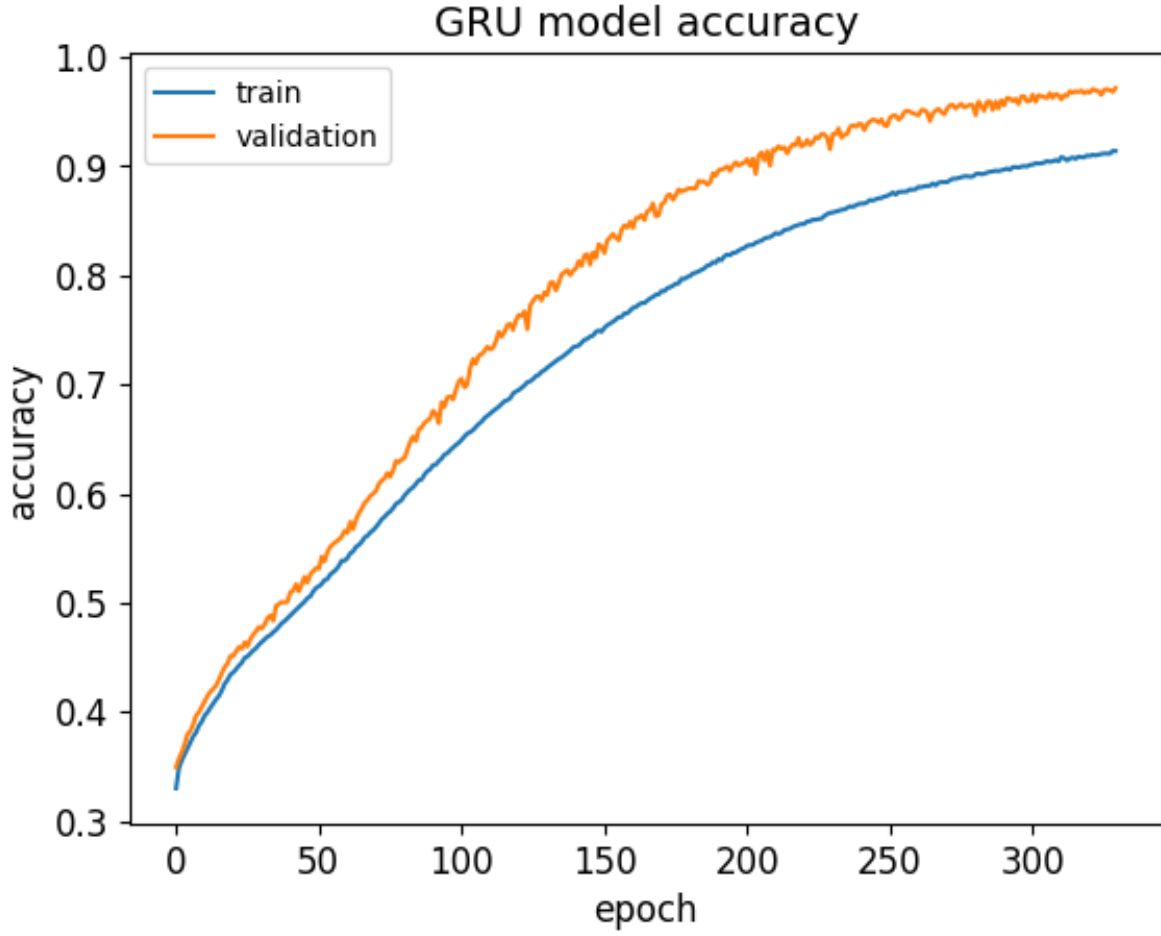
Son olarak, model LVLA sınıfı için 9498 doğru sınıflandırma ile yüksek bir sınıflandırma doğruluğu elde etmiştir. Ek olarak, modelin bazı sınıfları ayırt etmekte zorlandığı tespit edilmiştir. Özellikle, HVHA ile LVLA sınıfları arasında belirgin bir karışıklık görülmektedir.

Tablo 4.5 Çok modlu Transformer modeline ait sınıflandırma raporu

	precision	recall	f1-score	support
HVHA	0.80	0.87	0.83	19190
HVLA	0.85	0.78	0.82	11504
LVHA	0.85	0.81	0.83	14288
LVLA	0.77	0.77	0.77	12306
accuracy			0.82	57288
macro_accuracy	0.82	0.81	0.81	57288
weighted_accuracy	0.82	0.82	0.82	57288

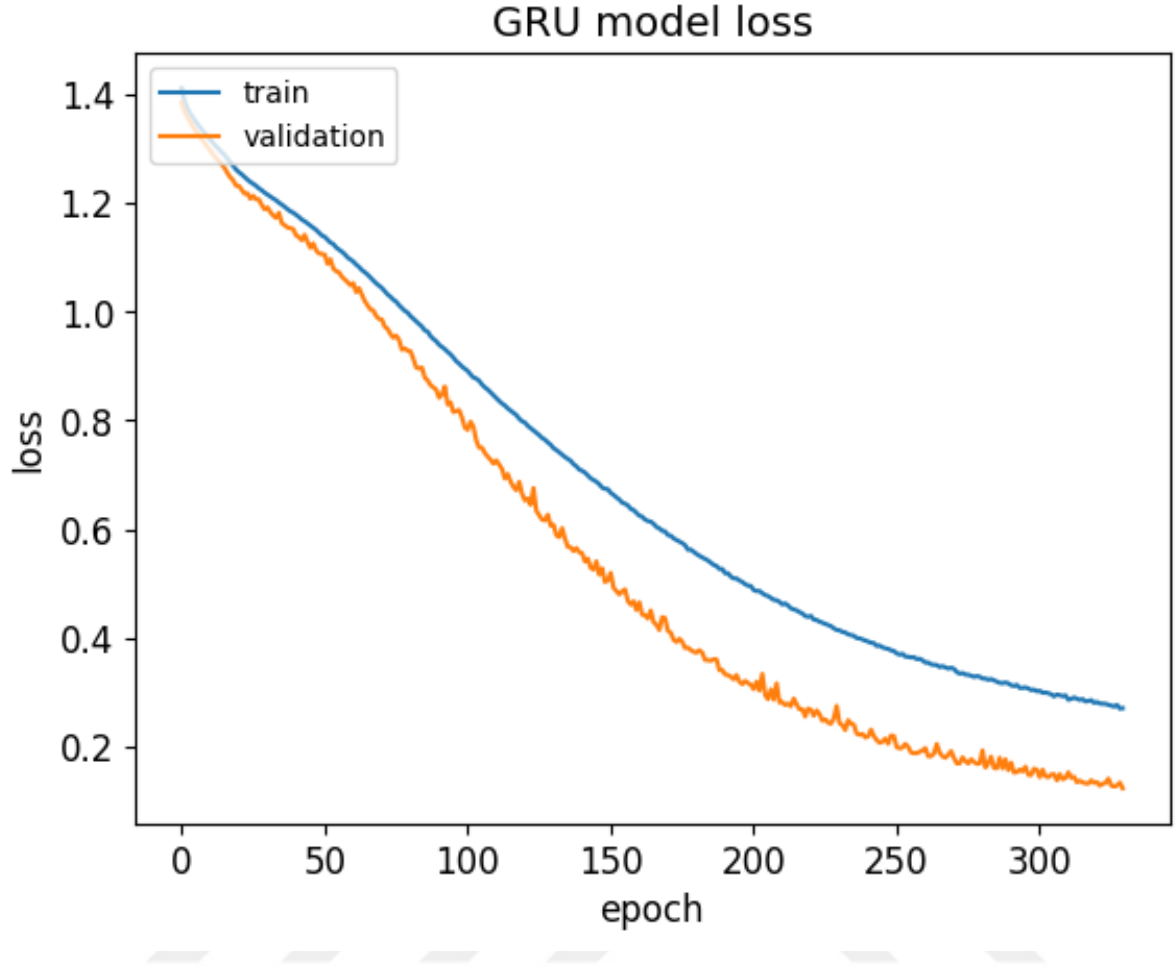
Tablo 4.5'te Transformer modelinin sınıflandırma raporundaki precision sütunu incelendiğinde, modelin HVHA sınıfını %80, HVLA sınıfını %85, LVHA sınıfını yine %85 ve LVLA sınıfını ise %77 precision ile tahmin ettiği gözlemlenmiştir. Recall sonuçları HVHA için %87, HVLA için %78, LVHA için %81 ve LVLA için %77 olarak gerçekleşmiştir. Modelin F1-skorları ise HVHA sınıfı için %83, HVLA sınıfı için %82, LVHA sınıfı için %83 ve LVLA sınıfı için %77'dir.

Model, %82'lik genel doğruluk oranı ile iyi bir performans göstermiştir. Makro ve ağırlıklı ortalamalar ise %81 ve %82 olarak ölçülmüştür. Sonuçlar, modelin belirli sınıflar arasında küçük performans farklılıkları olduğunu göstermekle birlikte, çok modlu verilerin dahil edilmesi ile tek modlu Transformer modeline (Bölüm 4.1 Tablo 4.1) kıyasla duyguları sınıflandırmada daha yüksek bir başarı elde edildiğini ortaya koymuştur.



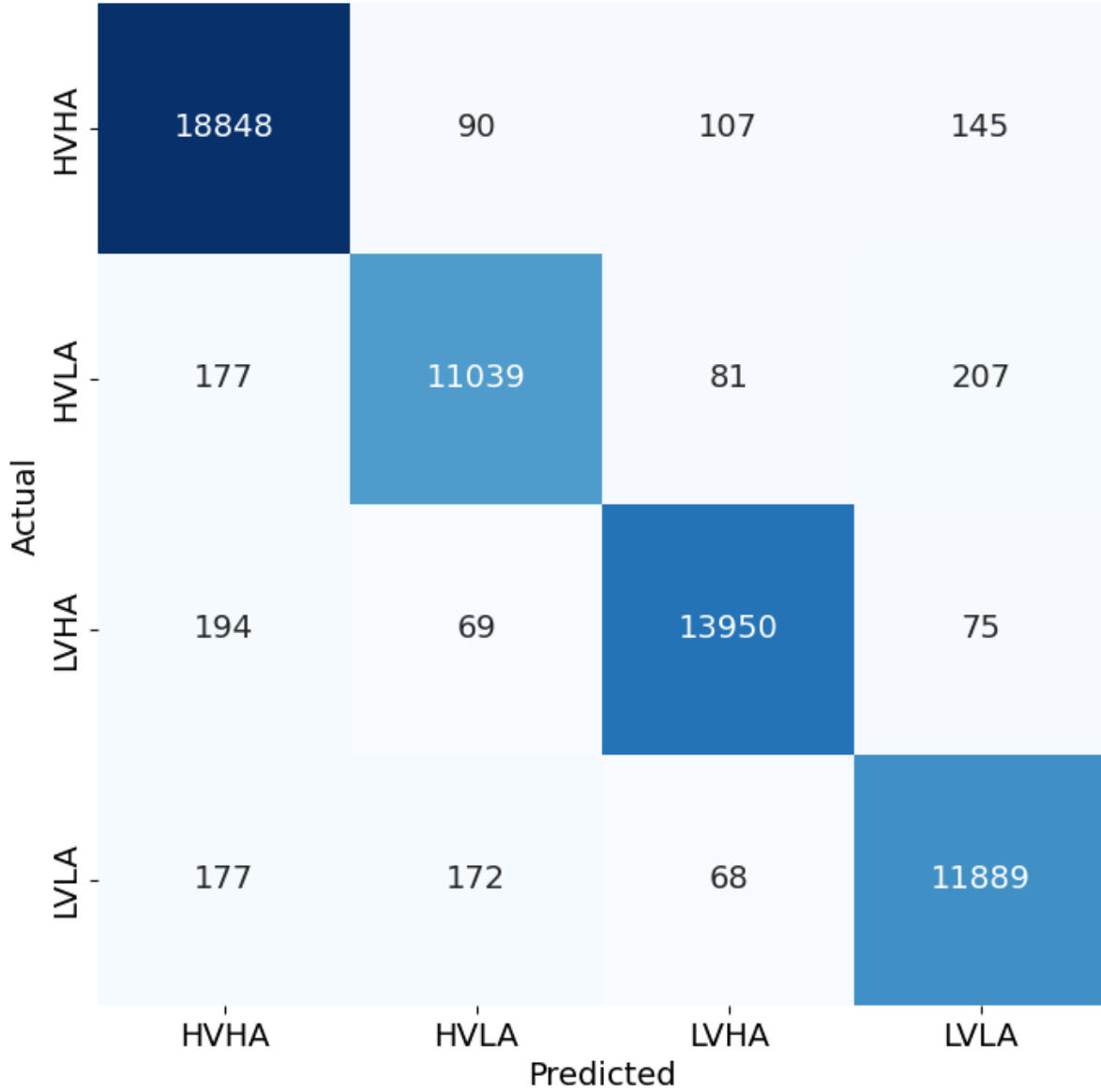
Şekil 4.16 Çok modlu GRU modeline ait doğruluk grafiği

Şekil 4.16'daki birleşik EEG ve yüz ifadeleri veri seti ile eğitilmiş GRU modelinin doğruluk grafiğinden, eğitim doğruluğunun başlangıçtan itibaren düzenli bir artış eğilimi gösterdiği anlaşılmaktadır. Eğitim doğruluk eğrisi 300'üncü eğitim turunda neredeyse %90 seviyesine ulaşmıştır. Bu durum, modelin eğitim setindeki verilerden yüksek kapasitede öğrendiğini göstermektedir. Validasyon doğruluğu başlangıçta eğitim doğruluğuna benzer bir artış göstermiş ve belirli bir noktadan itibaren eğitim doğruluğu ile uyumlu bir şekilde artmıştır. Validasyon doğruluğundaki artış, eğitim turlarının sonuna kadar devam etmiş ve yüksek seviyelere ulaşmıştır. Bu, modelin sadece eğitim verilerinde değil, aynı zamanda görmediği veriler üzerinde de iyi bir genelleme performansına sahip olduğunu gösterir. Eğitim ve validasyon doğruluğu arasındaki dengeli boşluk (gap) ve validasyon doğruluğunun yüksek bir seviyede sabitlenmesi, modelin aşırı uyuma gitmediğini ve veri setinin genel özelliklerini yakalayabildiğini ortaya koymaktadır.



Şekil 4.17 Çok modlu GRU modeline ait kayıp grafiği

Şekil 4.17'deki GRU modelinin kayıp grafiği hem eğitim hem de validasyon setleri için modelin kayıp değerlerinin nasıl değiştiğini göstermektedir. Eğitim kaybı, başlangıçtan itibaren düşüş göstermiş ve bu düşüş eğitim turları boyunca devam etmiştir. Grafiğin sonuna doğru, kaybın azalma hızı yavaşlamış ve bir plato oluşumu gözlemlenmiştir. Validasyon kaybı da başlangıçtan itibaren eğitim turları boyunca düşmüş ve eğitim kaybına benzer bir trend izleyerek eğitim sonlarına doğru plato oluşturmuştur. Eğitim süresince kaybın azalması, modelin eğitim verileri üzerindeki öğrenme kapasitesinin arttığına işaret etmektedir.



Şekil 4.18 Çok modlu GRU modeline ait karışıklık matrisi

Şekil 4.18'deki karışıklık matrisinde, GRU modelinin test veri seti üzerinde HVHA, HVLA, LVHA ve LVLA sınıfları için sınıflandırma sonuçları sunulmuştur. Model, HVHA sınıfı için 18848 doğru tahmin ile yüksek bir doğruluk elde etmiş ve bu sınıf için oldukça düşük yanlış negatif tahminler üretmiştir. HVLA sınıfı için 11039 doğru tahmin ile yüksek doğruluk gözlemlenmiştir. LVHA sınıfı için model, 13950 doğru tahmin ile benzer şekilde güçlü bir doğruluk oranı ortaya koymuştur. LVLA sınıfında ise 11889 doğru tahmin ile kayda değer bir doğruluk oranı elde edilmiştir. Sınıf etiketleri arasında oldukça az sayıda yanlış sınıflandırma mevcuttur. Bu düşük yanlış sınıflandırma oranıyla GRU modeli, farklı duygusal durumları tahmin etmede yüksek etkinlik sergilediğini ortaya koymuştur.

Tablo 4.6 Çok modlu GRU modeline ait sınıflandırma raporu

	precision	recall	f1-score	support
HVHA	0.97	0.98	0.98	19190
HVLA	0.97	0.97	0.97	11504
LVHA	0.98	0.98	0.98	14288
LVLA	0.97	0.97	0.97	12306
accuracy			0.97	57288
macro_accuracy	0.97	0.97	0.97	57288
weighted_accuracy	0.97	0.97	0.97	57288

Tablo 4.6’da GRU modeline ait sınıflandırma raporu sunulmuştur. Bu rapor incelendiğinde, her bir sınıf için %97 ile %98 arasında değişen precision oranları, modelin pozitif olarak tahmin ettiği durumların çoğunun gerçekten pozitif olduğunu göstermektedir. Recall değerleri de %96 ile %98 arasında olup, gerçek pozitif durumların büyük oranda doğru tahmin edildiğini belirtmektedir. %97 ile %98 arasında değişen F1-skorları hem precision hem de recall metriklerinde modelin dengeli bir performans sergilediğini gösterir. Modelin genel doğruluğu %97 olarak ölçülmüş ve bu oran, çok modlu LSTM ve Transformer modellerinin genel doğruluk oranlarından daha yüksektir. Makro ve ağırlıklı doğruluk değerleri de %97 ile modelin sınıflar arasında tutarlı bir performans sergilediğini göstermektedir. Bu sonuçlar, modelin duyguları pratikte de güvenilir ve etkili bir şekilde sınıflandırabileceğini göstermiştir.

Genel olarak, yüz ifadeleriyle birleştirilmiş EEG verileri kullanılarak yapılan sınıflandırmalarda her bir model, sadece EEG verilerine kıyasla daha iyi sonuçlar elde etmiştir. GRU modeli hem tek modlu hem de çok modlu yaklaşımlarda en yüksek doğrulukla duyguları sınıflandırmış ve bu bağlamda hem Transformer hem de LSTM modellerini geride bırakmıştır. Diğer modellere kıyasla daha yüksek performansı, GRU modelini bu veri setindeki duygusal durumları tahmin etmek için daha uygun hale getirmektedir.

4.3 Genel Değerlendirme ve Genel Mimari

Her tek modlu hem de çok modlu yaklaşımlar için LSTM, Transformer ve GRU olmak üzere üç farklı model test edilmiş ve GRU modeli duygu sınıflandırma konusunda en yüksek performansı göstermiştir. Karışıklık matrisleri, sınıflandırma raporları, doğruluk ve kayıp grafiklerinin analizi, modelde iyileştirilmesi gereken alanların belirlenmesine olanak tanımıştır. Özellikle, çok modlu veri setlerinin kullanılması, modellerin doğruluk oranını iyileştirmede önemli bir unsur olarak değerlendirilmiştir.

Tablo 4.7 Modellerin genel doğruluk oranları

Model – Genel Doğruluk Oranları	Tek Modlu (Sadece EEG)	Tek Modlu (Sadece Yüz İfadeleri)	Çok Modlu (Hem EEG hem yüz)
GRU	0.88	0.49	0.97
LSTM	0.82	0.49	0.94
Transformer	0.70	0.40	0.82

Tablo 4.7’de, hem tek modlu hem de çok modlu veriler üzerinde modellerin elde ettiği doğruluk oranları yer almaktadır. Tek modlu yaklaşımlar, sadece EEG ve sadece yüz ifadeleri verileriyle duyguların sınıflandırılmasını kapsamaktadır. EEG verilerini kullanan tek modlu yaklaşımlar için GRU, LSTM ve Transformer modelleri sırasıyla %88, %82 ve %70 doğruluk oranları ile duyguları sınıflandırma konusunda iyi başarı göstermişlerdir. Ancak, yalnızca yüz ifadeleri verilerinin kullanıldığı tek modlu yaklaşımlarda modeller sırasıyla %49, %49 ve %40 gibi daha düşük doğruluk oranları elde etmiştir. Bu durum, yüz ifadeleri verilerinin tek başına duyguları sınıflandırmada yeterli olmadığını göstermektedir. Ancak, çok modlu yaklaşımlarda modellerin genel doğruluğu artmıştır. Bu artış, yüz ifadeleri verilerinin, EEG verilerine takviye olarak kullanıldığında duygu sınıflandırmadaki başarıyı artırabileceği yönünde bir çıkarım yapılmasını sağlar.

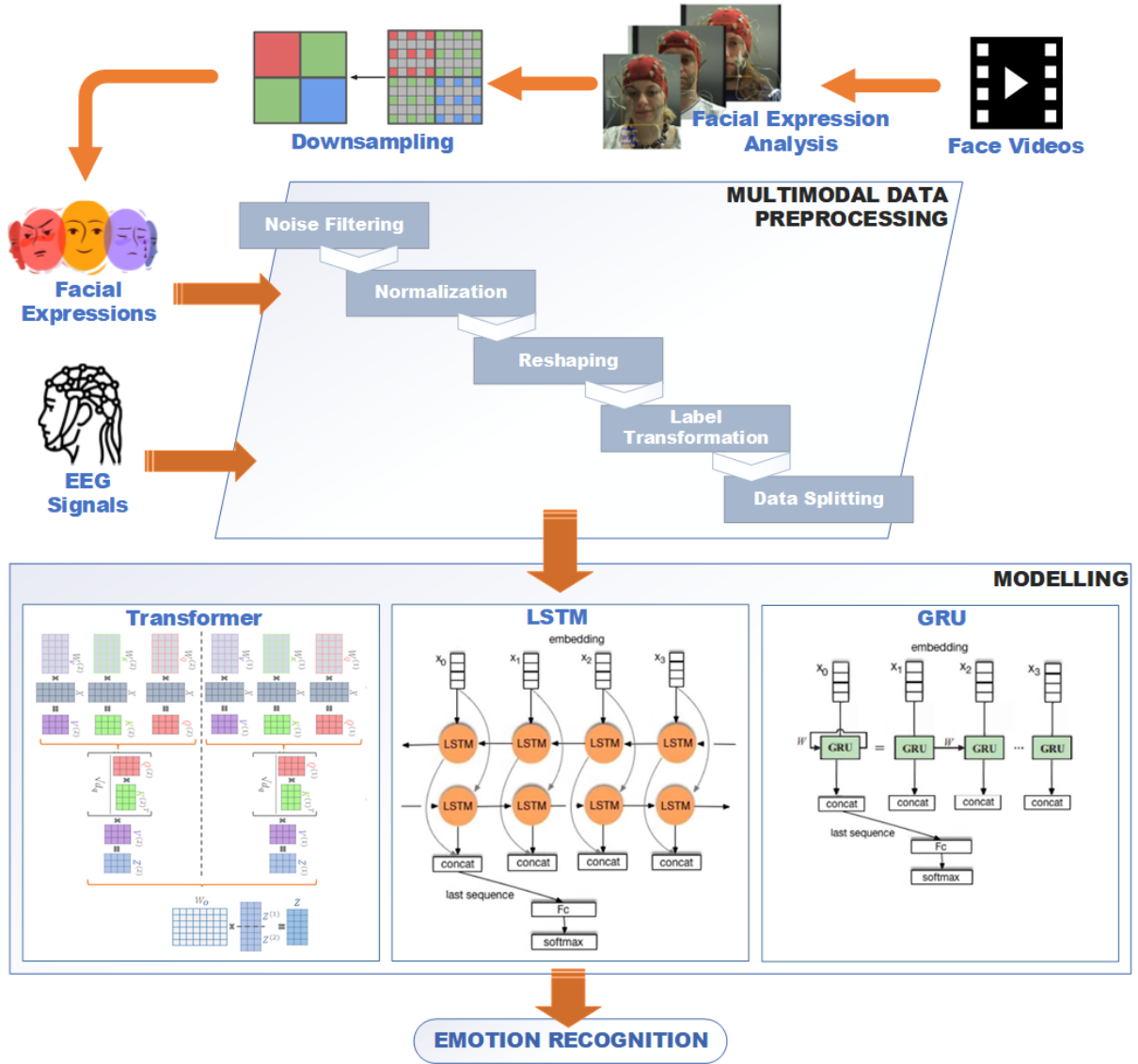
GRU modeli hem tek modlu (sadece EEG ve sadece yüz ifadeleri) hem de çok modlu veri setlerinde sırasıyla %88, %49 ve %97 ile en yüksek doğruluk oranlarına ulaşmıştır. Buna karşın, Transformer modeli, her iki yaklaşımda da diğer modellerle kıyaslandığında %70, %40 ve %82 ile en düşük doğruluk oranlarını elde etmiştir. GRU modelinin basit yapısı ve az parametre içermesi, EEG verilerinin boyutuyla ve zaman yapısıyla iyi uyum sağlamasına olanak tanımıştır.

Öte yandan, Transformer modelinin daha düşük performans göstermesinin nedenleri arasında karmaşık yapısının daha büyük ve kapsamlı veri setlerine daha uygun olması gösterilebilir.

Tablo 4.8 Modellerin genel kayıp performansları

Model – Kayıp İstatistiği	Tek Modlu (Sadece EEG)	Tek Modlu (Sadece Yüz İfadeleri)	Çok Modlu (Hem EEG hem yüz)
GRU	0.36	1.14	0.12
LSTM	0.54	1.14	0.21
Transformer	0.80	1.26	0.52

Tablo 4.8’de GRU, LSTM ve Transformer modellerinin hem tek modlu hem de çok modlu veriler kullanılarak test setlerinde ulaştıkları kayıp skorları sunulmuştur. GRU modeli 0.36, 1.14 ve 0.12 değerleri ile her iki yaklaşımda da en düşük kayıp skorlarına ulaşmıştır. LSTM modeli, tek modlu yaklaşımda sadece EEG sinyalleri ile 0.54, sadece yüz ifadeleriyle 1.14 ve çok modlu yaklaşımda ise 0.21’lik kayıp skorlarıyla iki modelin arasında iyi bir performans sergilemiştir. Öte yandan, Transformer modeli tek modlu verilerde 0.80 ve 1.26, çok modlu verilerde ise 0.52 ile diğer iki modele göre daha kötü bir kayıp performansı sergilemiştir. Sonuç olarak, GRU modelinin düşük kayıp skorları tek modlu ve çok modlu veri setlerinde kullanımı açısından güçlü bir argüman oluşturmaktadır.



Şekil 4.19 Sistemin genel mimarisi

Şekil 4.19'da özetlenen genel mimari DEAP veri setinden elde edilen EEG sinyalleri ve yüz ifadeleri kullanılarak duygu sınıflandırmasını amaçlayan tek modlu ve çok modlu deneyleri göstermektedir. Bu çalışmada, tek modlu deneyde, EEG verileri öncelikle gürültü azaltma, normalizasyon, ölçeklendirme ve yeniden şekillendirme gibi çeşitli ön işleme adımlarından geçirilmekte, ardından bu veriler Transformer, LSTM ve GRU gibi farklı derin öğrenme modellerine giriş olarak kullanılmaktadır. Modellerin performansı, doğruluk, kesinlik, geri çağırma ve F1-skoru gibi metrikler kullanılarak değerlendirilmektedir.

Çok modlu deneyde ise, yüz videolarından elde edilen veriler, yüz ifadelerini analiz eden bir algoritma kullanılarak işlenmektedir. Bu analiz sonuçları, EEG downsampling yöntemiyle işlenmiş ve ardından EEG sinyalleri ile birleştirilmiştir. Bu şekilde elde edilen çok modlu veri seti, duyguları sınıflandırmak için kullanılan modellere giriş verisi olarak sunulmaktadır. Her iki deney tipinde de elde edilen sonuçlar üzerinden modellerin duygu tanıma performansları detaylı bir şekilde analiz edilmektedir.

4.4 Öneri Sistemi Sonuçları

Müzik öneri sisteminin anlatıldığı Bölüm 3.4'te elde edilen bulguların detaylı analizi bu bölümde sunulmuştur. Modelin test kullanıcısı için sunduğu önerilerin bir görseli Şekil 4.20'de sunulmuştur. Elde edilen çıktıdan, öneri sisteminin çeşitlilik ve kişiselleştirme yönleri değerlendirilmiştir. Model tarafından önerilen müziklerin listesi, sanatçılar ve şarkı isimleriyle birlikte incelenmiştir. Önerilen parçalar arasında Marilyn Manson'dan "The Beautiful People", James Blunt'ın "Goodbye My Lover" ve Enya'nın "May It Be" (Saving Private Ryan film müziği) gibi farklı tarzları ve duygusal tonları temsil eden şarkılar bulunmaktadır. Ayrıca listede "Song 2" ile Blur ve "Breathe Me" ile Sia gibi çeşitli dönemlerden ve farklı müzik türlerinden örnekler de mevcuttur. Bu liste, öneri sisteminin geniş bir yelpazedeki müzikleri keşfedebildiğini ve potansiyel olarak kullanıcı memnuniyetini artırabilecek öneriler sunduğunu göstermektedir.

```
[('Marilyn Manson', 'The Beautiful People'),  
 ('The Submarines', 'Darkest Things'),  
 ('Emiliana Torrini', 'Jungle Drum'),  
 ('James Blunt', 'Goodbye My Lover'),  
 ('Enya ', 'May It Be (Saving Private Ryan)'),  
 ('Blur ', 'Song 2 '),  
 ('Sia', 'Breathe Me'),  
 ('Air', 'Moon Safari'),  
 ('Madonna', 'Rain'),  
 ("The B52'S ", 'Love Shack')]
```

Şekil 4.20 Öneri sisteminin test kullanıcısı için yaptığı öneriler

Mevcut sistem, kullanıcılar arası benzerlikleri hesaplamakta ve benzer kullanıcıların beğenilerine göre öneriler sunmaktadır. Sistemin kişiselleştirilebilmesi için kullanıcıların etkileşimleri ve tepkileri daha ayrıntılı olarak analiz edilebilir. Örneğin, bir kullanıcının bir şarkıyı kaç kez dinlediği, şarkıları dinleme süresi, atladığı şarkılar ve kullanıcıların oylama sistemine verdikleri yanıtlar gibi davranışsal veriler incelenerek öneri algoritması daha kişisel hale getirebilir. Bu tür veriler, kullanıcı tercihlerinin daha derinlemesine anlaşılmasını sağlayarak, daha doğru önerilerde bulunulmasına imkân tanır.

Ayrıca, öneri sistemi içerik tabanlı filtreleme ve işbirliğine dayalı filtreleme yöntemlerini hibrit bir yaklaşım ile birleştirebilir. İçerik tabanlı filtreleme, öğelerin (şarkıların) özelliklerini (tür, ritim, enstrümanlar, sözlerin duygusal içeriği vb.) analiz ederek benzer özelliklere sahip öğeleri önerirken, işbirliğine dayalı filtreleme kullanıcılar arası benzerlikleri temel alır. Hibrit sistemler, her iki metodun avantajlarını bir araya getirerek daha isabetli öneriler yapabilir. Öte yandan, kullanıcının konumu, günün saati, mevsim ve hatta hava durumu gibi bağlam verileri, kullanıcının muhtemel ruh halini ve dolayısıyla müzik tercihlerini etkileyebilir. Kullanıcıların sosyal medyada beğendikleri içerikler, takip ettikleri sanatçılar ve arkadaşlarının tercihleri gibi sosyal veriler analiz edilerek, sosyal çevreden gelen önerilerin gücünden yararlanılabilir. Bu, kullanıcıların keşfetmediği yeni müzikleri veya trendleri keşfetmesine yardımcı olabilir.

Son olarak, yapay zekâ modellerinin sürekli öğrenme kabiliyetleri, öneri sistemlerinin sürekli olarak gelişmesine olanak tanır. Kullanıcı verileri ile beslenen ve sürekli olarak güncellenen modeller, zamanla kullanıcı tercihlerini daha iyi anlamayı ve kişiselleştirmeyi öğrenebilir. Bu, modelin esnekliğini artırır ve kullanıcı tatminini maksimize ederken, öneri sisteminin değişen eğilimlere ve tercihlere uyum sağlamasını sağlar.

Bu çalışmada, kullanıcıların EEG sinyalleri ve yüz ifadelerinin birleşiminden elde edilen duygusal durumları da bir özellik olarak sisteme verilerek bir öneri sistemi oluşturulmuştur. Sunulan model, kullanıcı verilerinden yararlanarak ve kullanıcılar arasındaki benzerliklere dayalı olarak çeşitli ve özelleştirilmiş müzik önerileri sağlamıştır. Öneriler, deney sırasında katılımcılara sunulan müzik yelpazesinden seçilmiş ve kullanıcıların genel anlamda beğenilerini yansıtacak şekilde tasarlanmıştır. Ancak, bu ilk adımın ötesine geçmek ve kullanıcı deneyimini daha da zenginleştirmek için modelin kişisel tercihlerle sürekli beslenmesi gerektiği ve daha detaylı analizlere ihtiyaç duyduğu anlaşılmaktadır.

5. SONUÇ VE ÖNERİLER

Bu çalışma, duygu tanıma alanında önemli bir katkı sunarak, EEG sinyalleri ve yüz ifadeleri gibi çok modlu veri kaynaklarının entegrasyonunun, duyguların sınıflandırılmasında nasıl bir fark yaratabileceğini detaylı bir şekilde inceler. Gelişmiş derin öğrenme modelleri olan GRU, LSTM ve Transformer'ın performanslarını hem tek modlu hem de çok modlu veri setlerinde karşılaştırmak suretiyle, GRU modelinin tutarlı olarak yüksek doğruluk oranları elde ettiği ve duygu sınıflandırma görevlerinde en iyi performansı sergilediği görülmüştür. Bu bulgu, çok modlu veri entegrasyonunun, derin öğrenme modellerinin duygusal durumları tanıma kabiliyetini önemli ölçüde artırabileceğini vurgular.

Çalışmanın yürütüldüğü süreç, EEG sinyallerinin ve yüz ifadelerinin nasıl toplandığı, işlendiği ve birleştirildiğiyle ilgili detaylı metodolojik bilgiler sunar. Bu süreç, derin öğrenme modellerinin eğitimi için gerekli olan zengin ve çeşitlendirilmiş veri setinin oluşturulmasında kritik bir rol oynar. GRU modelinin diğer modellere göre üstün performansı, zaman serisi verileriyle çalışırken gösterdiği etkili öğrenme kapasitesine ve karmaşık duygusal durumların tanınmasında daha iyi genelleme yapabilmesine dayandırılabilir.

Bu araştırmanın teorik katkıları, duygu tanıma çalışmalarının yanı sıra insan-bilgisayar etkileşimi ve yapay zeka sistemlerinin anlayışına yeni bir boyut getirir. Çok modlu veri entegrasyonunun önemini vurgulayarak, bu yaklaşımın, derin öğrenme modellerinin daha hassas ve doğru bir şekilde duygusal ifadeleri sınıflandırmasına olanak tanıdığını ortaya koyar. Ayrıca, duygu tanıma teknolojilerinin gelişimine katkıda bulunarak, bu teknolojilerin günlük yaşamda, özellikle de insan-makine etkileşiminde oynayabileceği roller üzerine önemli düşünceler sunar.

Pratik uygulamalara gelince, bu çalışma, etkileşimli sistemler, eğitim teknolojileri, sağlık hizmetleri ve güvenlik sistemleri gibi birçok alanda potansiyel kullanım senaryoları sunmaktadır. Örneğin, duygu tanıma sistemlerinin gelişimi, öğrencilerin veya kullanıcıların duygusal durumlarını daha iyi anlayarak, onlara daha kişiselleştirilmiş deneyimler sunabilecek

eđitim yazılımlarının tasarımıını mmkn kılar. Benzer Őekilde, sađlık hizmetlerinde, duygusal durumların otomatik olarak tanınması, depresyon gibi durumların erken teŐhisinde yardımcı olabilir.

AraŐtırmanın sınırlılıkları da dikkate alınarak, gelecekteki alıŐmaların daha byk ve eŐitli veri setleri zerinde yrtlmesi, farklı duygusal durumları daha geniŐ bir yelpazede kapsayacak Őekilde geniŐletilmesi nerilmektedir. Ayrıca, yeni derin đrenme mimarilerinin keŐfi ve farklı modalitelerin entegrasyonu, duygu tanıma alanındaki alıŐmaları daha da ileriye taŐıyabilecek potansiyel alanlardır.

Son olarak, bu alıŐma, duygu tanıma teknolojilerinin gelecekteki geliŐimine ynelik sađlam bir temel oluŐtururken, insan-makine etkileŐiminin daha derinlemesine anlaŐılmasına ve geliŐtirilmesine olanak tanıyan deđerli iđrler sunmaktadır. Teknolojinin, insan duygularını daha iyi anlama ve yorumlama yeteneđinin geliŐtirilmesi, yapay zekanın daha duyarlı ve anlayıŐlı hale gelmesine ynelik nemli adımlardan biridir. Bu bađlamda, gelecekteki araŐtırmaların, bu alıŐmanın bulgularını temel alarak, duygu tanıma teknolojilerinin ve insan-makine etkileŐiminin sınırlarını geniŐletmeye devam etmesi beklenmektedir.

6. KAYNAKÇA

Aboalsamh, H., Hussain, M., ve Muhammad, F. "A bimodal emotion recognition approach through the fusion of electroencephalography and facial sequences." *Diagnostics*, cilt 13, no. 5, 2023.

Ali, A. A., Hammad, A., ve Houssein, E. H. "Human emotion recognition from EEG-based brain-computer interface using machine learning: a comprehensive review." *Neural Computing and Applications*, cilt 34, no. 15, ss. 12527-12557, 2022.

Akbulut, F. P., ve Dogan, G. "Multi-modal fusion learning through biosignal, audio, and visual content for detection of mental stress." *Neural Computing and Applications*, cilt 35, no. 34, ss. 24435-24454, 2023.

Akbulut, F. P. "Evaluating the effects of the autonomic nervous system and sympathetic activity on emotional states." *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, cilt 21, no. 41, ss. 156-169, 2022.

Ayata, D., Yaslan, Y., & Kamasak, M. E. "Emotion based music recommendation system using wearable physiological sensors." *IEEE Transactions on Consumer Electronics*, cilt 64, no. 2, ss. 196-203, 2018.

Buckland, M., ve Gey, F. "The relationship between recall and precision." *Journal of the American Society for Information Science*, cilt 45, no. 1, ss. 12-19, 1994.

Chakravarthy, G. K., Suchithra, M., & Thatavarti, S. (2023). EEG-based multimodal emotion recognition with optimal trained hybrid classifier. *Multimedia Tools and Applications*, 1-24.

Chaparro, V., ve ark. "Emotion recognition from EEG and facial expressions: a multimodal approach." *Proc. IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, ss. 530-533.

Chen, J., Liu, Y., Xue, W., Hu, K., & Lin, W. (2022). Multimodal EEG emotion recognition based on the attention recurrent graph convolutional network. *information*, 13(11), 550.

Chen, P., Yin, Z., Zhang, J., ve Nichele, S. "Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review." *Information Fusion*, cilt 59, ss. 103-126, 2020.

Cimtay, Y., Ekmekcioglu, E., & Caglar-Ozhan, S. (2020). Cross-subject multimodal emotion recognition based on hybrid fusion. *IEEE Access*, 8, 168865-168878.

Derdiyok, S., ve Akbulut, F. P. "Biosignal based emotion-oriented video summarization." *Multimedia Systems*, cilt 29, no. 3, ss. 1513-1526, 2023.

Dey, R., ve Salem, F. M. "Gate-variants of gated recurrent unit (GRU) neural networks." *Proc. 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2017, ss. 1597-1600.

Hochreiter, S., ve Schmidhuber, J. "Long short-term memory." *Neural Computation*, cilt 9, no. 8, ss. 1735-1780, 1997.

Huang, W., Wu, W., Lucas, M. V., Huang, H., Wen, Z., ve Li, Y. "Neurofeedback training with an electroencephalogram-based brain-computer interface enhances emotion regulation." *IEEE Transactions on Affective Computing*, 2021.

Huang, Y., ve ark. "Combining facial expressions and electroencephalography to enhance emotion recognition." *Future Internet*, cilt 11, no. 5, 2019.

Huang, Y., ve ark. "Fusion of facial expressions and EEG for multimodal emotion recognition." *Computational Intelligence and Neuroscience*, cilt 2017, 2017.

Ioffe, S., ve Szegedy, C. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." *Uluslararası Makine Öğrenimi Konferansı*, 2015, ss. 448-456. PMLR.

Keshari, T., ve Palaniswamy, S. "Emotion recognition using feature-level fusion of facial expressions and body gestures." *Proc. International Conference on Communication and Electronics Systems (ICCES)*, 2019, ss. 1184-1189.

Koelstra, S., ve ark. "Deap: A database for emotion analysis; using physiological signals." *IEEE Transactions on Affective Computing*, cilt 3, no. 1, ss. 18-31, 2011.

- Landis-Shack, N., Heinz, A. J., & Bonn-Miller, M. O. "Music Therapy for Posttraumatic Stress in Adults: A Theoretical Review." *Psychomusicology*, cilt 27, no. 4, ss. 334-342, 2017. <https://doi.org/10.1037/pmu0000192>
- Leong, S. C., Tang, Y. M., Lai, C. H., ve Lee, C. K. M. "Facial expression and body gesture emotion recognition: A systematic review on the use of visual data in affective computing." *Computer Science Review*, cilt 48, 2023.
- Liu, W., Zheng, W. L., ve Lu, B. L. "Emotion recognition using multimodal deep learning." *Neural Information Processing: ICONIP 2016, Kyoto, Japan, 2016*, Springer International Publishing, ss. 521-529.
- Luz Junior, J. D. A., ve Rodrigues, M. A. F. "Comparative analysis of facial expression recognition systems for evaluating emotional states in virtual humans." *Proc. Symposium on Virtual and Augmented Reality*, 2023, ss. 38-47.
- Ma, J., ve ark. "Emotion recognition using multimodal residual LSTM network." *Proc. ACM International Conference on Multimedia*, 2019, ss. 176-183.
- McCaffrey, T., Edwards, J., & Fannon, D. "Is there a role for music therapy in the recovery approach in mental health?". *The Arts in Psychotherapy*, cilt 38, no. 3, ss. 185-189, 2011.
- Ng, A. Y. "Feature selection, L1 vs. L2 regularization, and rotational invariance." *Yirmi Birinci Uluslararası Makine Öğrenimi Konferansı İşlemleri*, 2004, s. 78.
- Pessoa, L. "Neural dynamics of emotion and cognition: From trajectories to underlying neural geometry." *Neural Networks*, cilt 120, ss. 158-166, 2019.
- Powers, D. M. "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation." *arXiv ön baskı arXiv:2010.16061*, 2020.
- Prechelt, L. "Early stopping-but when?." *Neural Networks: Tricks of the Trade içinde*, ss. 55-69, Berlin, Heidelberg: Springer Berlin Heidelberg, 2002.
- Rayatdoost, S., Rudrauf, D., ve Soleymani, M. "Multimodal gated information fusion for emotion recognition from EEG signals and facial behaviors." *Proc. International Conference on Multimodal Interaction*, 2020, ss. 655-659.
- Russell, J. A. "A circumplex model of affect." *Journal of Personality and Social Psychology*, cilt 39, no. 6, ss. 1161, 1980.

Singh, R. H., Maurya, S., Tripathi, T., Narula, T., & Srivastav, G. "Movie recommendation system using cosine similarity and KNN." *International Journal of Engineering and Advanced Technology*, cilt 9, no. 5, ss. 556-559, 2020.

Sharma, S., Sharma, S., ve Athaiya, A. "Activation functions in neural networks." *Towards Data Science*, cilt 6, no. 12, ss. 310-316, 2017.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., ve Salakhutdinov, R. "Dropout: a simple way to prevent neural networks from overfitting." *The Journal of Machine Learning Research*, cilt 15, no. 1, ss. 1929-1958, 2014.

Tan, Y., Sun, Z., Duan, F., Solé-Casals, J., & Caiafa, C. F. (2021). A multimodal emotion recognition method based on facial expressions and electroencephalography. *Biomedical Signal Processing and Control*, 70, 103029.

Tharwat, A. "Classification assessment methods." *Applied Computing and Informatics*, cilt 17, no. 1, ss. 168-192, 2020.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ve ark. "Attention is all you need." *Advances in Neural Information Processing Systems*, cilt 30, 2017.

Viola, P., ve Jones, M. "Rapid object detection using a boosted cascade of simple features." 2001 IEEE Bilgisayar Topluluğu Konferansı'nda Bilgisayar Görüşü ve Desen Tanıma İşlemleri, CVPR 2001, Cilt 1, 2001.

Wang, S., Qu, J., Zhang, Y., & Zhang, Y. (2023). Multimodal emotion recognition from EEG signals and facial expressions. *IEEE Access*, 11, 33061-33068.

Witusik, A., & Pietras, T. "Music therapy as a complementary form of therapy for mental disorders." *Polski Merkuriusz Lekarski: Organ Polskiego Towarzystwa Lekarskiego*, cilt 47, no. 282, ss. 240-243, 2019.

Wu, D., Zhang, J., ve Zhao, Q. "Multimodal fused emotion recognition about expression-EEG interaction and collaboration using deep learning." *IEEE Access*, cilt 8, ss. 133180-133189, 2020.

Zhang, K., Zhang, Z., Li, Z., ve Qiao, Y. "Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks." *IEEE Signal Processing Letters*, cilt 23, no. 10, ss. 1499-1503, Ekim 2016. DOI: 10.1109/LSP.2016.2603342.

Zheng, W. L., Dong, B. N., ve Lu, B. L. "Multimodal emotion recognition using EEG and eye tracking data." Proc. IEEE Engineering in Medicine and Biology Society, 2014, ss. 5040-5043.

