



# Derin Öğrenme Tabanlı Gerçek Zamanlı Vücut Hareketlerinden Duygu Analizi Modeli

Mustafa YAĞCI<sup>1\*</sup> , Muhammed Erkam AYGÜL<sup>1</sup> 

<sup>1</sup>Ahi Evran Üniversitesi, Mühendislik-Mimarlık Fakültesi, Bilgisayar Mühendisliği, Kırşehir, Türkiye

## Makale Bilgisi

Araştırma makalesi  
Başvuru: 01/11/2023  
Düzeltilme: 15/02/2024  
Kabul: 04/04/2024

## Anahtar Kelimeler

Derin Öğrenme  
Makine Öğrenmesi  
Görüntü İşleme  
Duygu Analizi  
Nitelik Çıkarımı

## Article Info

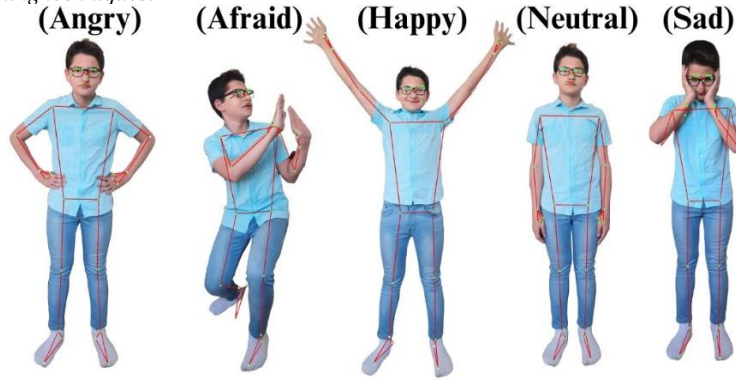
Research article  
Received: 01/11/2023  
Revision: 15/02/2024  
Accepted: 04/04/2024

## Keywords

Deep Learning  
Machine Learning  
Image Processing  
Emotion Analysis  
Attribute Extraction

## Grafik Özet (Graphical/Tabular Abstract)

Bu çalışmada insan pozlarını gerçek zamanlı olarak inceleyen ve duygu çıkarımı yapabilen bir makine öğrenmesi modeli, derin öğrenme teknikleri ile geliştirilmiştir. / In this study, a machine learning model that can analyze human poses in real time and infer emotions is developed with deep learning techniques.



Şekil A: Vücut pozu ve iskelet model / Figure A: Body pose and skeleton model

## Önemli noktalar (Highlights)

- Duygu çıkarımı, insan vücut hareketlerinin analiz edilmesi ile yapılmıştır. / Emotion inference is done by analyzing human body movements.
- Kızgın, korkmuş, mutlu, nötr ve üzgün duyguları tanımlanmıştır. / Angry, scared, happy, neutral and sad emotions were identified.
- Derin öğrenme ve yapay zekâ teknikleri ile teknolojinin insanlarla daha derin ve anlamlı şekilde etkileşime girmesi mümkündür. / With deep learning and artificial intelligence techniques, it is possible for technology to interact with people in a deeper and more meaningful way.

**Amaç (Aim):** Bu çalışmanın amacı, insan vücut hareketlerini derin öğrenme teknikleri ile analiz ederek bireyin duygusal durumlarını tahmin eden yeni bir makine öğrenmesi modeli geliştirmektir. / The aim of this research is to develop a new machine learning model that predicts an individual's emotional states by analyzing human body movements with deep learning techniques.

**Özgünlük (Originality):** Geleneksel duygu analizi yöntemleri genellikle yüz ifadelerine odaklanmaktadır. Bu çalışmada ise duygu analizi insan vücuduna iskelet modelin oturtulması ve eğitilen modelden çıkarım yapılması şeklindedir. / Traditional sentiment analysis methods usually focus on facial expressions. In this research, sentiment analysis is based on fitting a skeletal model to the human body and inferring from the trained model.

**Bulgular (Results):** Model, tanımlanan duyguları başarıyla sınıflandırarak, duygusal ifadelerin vücut dilinde belirgin olduğunu ve modelin bu ifadeleri doğru şekilde yorumlayabildiğini göstermiştir. / The model successfully classified the identified emotions, demonstrating that emotional expressions are evident in body language and that the model can correctly interpret these expressions.

**Sonuç (Conclusion):** Duygu çıkarımının insan pozlarından yapılabildiği gösterilmiştir. En yüksek doğruluk oranının %97.80 ile rastgele orman algoritması ile elde edildiği görülmüştür. Ayrıca gerçek zamanlı çıkarımda doğruluk oranı %94.50 olarak hesaplanmıştır. / It was shown that emotion extraction can be done from human poses. The highest accuracy rate was 97.80% with the random forest algorithm. In addition, the accuracy rate was calculated as 94.50% in real-time extraction.



## Derin Öğrenme Tabanlı Gerçek Zamanlı Vücut Hareketlerinden Duygu Analizi Modeli

Mustafa YAĞCI<sup>1\*</sup>, Muhammed Erkam AYGÜL<sup>1</sup>

<sup>1</sup>Ahi Evran Üniversitesi, Mühendislik-Mimarlık Fakültesi, Bilgisayar Mühendisliği, Kırşehir, Türkiye

### Makale Bilgisi

Araştırma makalesi  
Başvuru: 01/11/2023  
Düzeltilme: 15/02/2024  
Kabul: 04/04/2024

### Anahtar Kelimeler

Derin Öğrenme  
Makine Öğrenmesi  
Görüntü İşleme  
Duygu Analizi  
Nitelik Çıkarımı

### Öz

İletişim, canlılar arasında bilgi, duygu ve düşüncelerin aktarılması için kullanılan bir araçtır. Araştırmalar, insan iletişiminin yaklaşık %80'inin beden dili veya mimiklerle yapıldığını göstermektedir. Bir kişinin beden dilini yorumlamak ise o kişinin duygularından çıkarım yapılmasını sağlayabilir. Bu çalışmanın amacı, gerçek zamanlı olarak insan vücut hareketlerini derin öğrenme teknikleri ile analiz ederek bireyin duygusal durumlarını tahmin eden yeni bir makine öğrenmesi modeli geliştirmektir. İnsanların vücudunu tanıyan ve tanınan vücut pozuna iskelet modelin oturtulması işlemi ile o pozdan veri toplayan model geliştirilmiş, toplanan veri seti ile farklı modeller eğitilmiştir. Bununla birlikte gerçek zamanlı analiz yapılması için bir tahmin modeli geliştirilmiştir. Bu sayede insanların beden dili ve mimiklerinden duygu çıkarımı yapılabilmektedir. Yapılan farklı test modellerinin sonuçlarının doğruluk oranının en düşük %92.86, en yüksek %97.80 olduğu ve gerçek zamanlı analiz sonuçlarının doğruluk oranının %94.50 olduğu görülmüştür. Bu çalışmanın sonuçları daha iyi insan-makine etkileşimi, duygusal zekâ uygulamaları ve sosyal etkileşim alanlarında kullanılabilir.

## Deep Learning Based Real-Time Emotion Analysis Model from Body Movements

### Article Info

Research article  
Received: 01/11/2023  
Revision: 15/02/2024  
Accepted: 04/04/2024

### Keywords

Deep Learning  
Machine Learning  
Image Processing  
Emotion Analysis  
Attribute Extraction

### Abstract

Communication is a tool used to transfer information, feelings and thoughts between living beings. Research shows that about 80% of human communication is done through body language or gestures. Interpreting a person's body language can help infer their emotions. The aim of this study is to develop a new machine learning model that predicts an individual's emotional states by analyzing human body movements in real-time with deep learning techniques. A model was developed that recognizes the human body and collects data from that pose by fitting a skeletal model to the recognized body pose, and different models were trained with the collected data set. In addition, a prediction model was developed for real-time analysis. In this way, emotions can be inferred from people's body language and facial expressions. The accuracy rate of the results of the different test models was 92.86% at the lowest and 97.80% at the highest, and the accuracy rate of the real-time analysis results was 94.50%. The results of this study can be used in better human-machine interaction, emotional intelligence applications and social interaction.

### 1. GİRİŞ (INTRODUCTION)

Teknolojinin hızla evrildiği günümüz dünyasında, insan-makine etkileşimi gittikçe karmaşık hale gelmektedir. Bu karmaşıklık, kullanıcı deneyimini ve insanlarla yapay zekâ sistemleri arasındaki etkileşimi daha anlamlı hale getirmek için daha sofistike yaklaşımların arayışına yol açmaktadır.

İnsanların sözsüz ifadeleri, yani vücut hareketleri, genellikle duygusal durumların en doğal ifadelerinden biridir. Bu durum, özellikle etkileşimli robotlar, sanal gerçeklik deneyimleri ve sağlık sektöründeki uygulamalar gibi alanlarda duygusal zekâ ile donatılmış sistemlerin geliştirilmesine olan ilgiyi artırmıştır. Bu bağlamda, insanların duygusal durumlarını doğru bir şekilde

anlamak, insan-makine etkileşiminin kalitesini artırmanın ve sosyal etkileşimde bulunan yapay zekâ sistemlerinin daha insan benzeri hale gelmesinin temelidir.

İnsanlar, duygusal ifadelerini sadece kelimelerle ifade etmezler; aynı zamanda vücut hareketleri, jestler ve yüz ifadeleri gibi sözsüz iletişim araçlarını da kullanırlar. Bu sözsüz ifadeler, bir kişinin duygusal durumunu anlamının önemli bir parçasını oluşturur. Bu durum, özellikle insan-makine etkileşiminde, robot teknolojilerinde, oyun endüstrisinde, eğitimde ve zihinsel sağlık uygulamalarında önemli bir rol oynamaktadır. Bu bağlamda, insan vücut hareketlerinden duygusal durumları anlamak, yapay zekâ sistemlerinin insanlara daha duyarlı ve anlayışlı bir şekilde tepki vermesine imkân tanıyan bir anahtar haline gelmiştir. Derin öğrenme tekniklerini kullanarak insan vücut hareketlerinden duygusal durumlarını tahmin etmeye yönelik bazı çalışmalar şu şekildedir:

Gunawan vd. [1] derin öğrenme tekniklerini kullanarak video verilerinden duygusal durumları tanımak için bir model önermişlerdir. Evrişimli sinir ağlarını (CNN) ve uzun-kısa vadeli hafıza (LSTM) ağlarını kullanarak video görüntülerinden duygusal ifadeleri çıkarmışlardır. Ahmed vd. [2] vücut hareketlerinden duygu tanımayı amaçlamış, beş temel duyguyu (mutluluk, üzüntü, korku, öfke ve nötr) doğru bir şekilde tanımak için özellik seçim yöntemini kullanmışlardır. İlgisiz özellikleri ortadan kaldırmak için varyans analizi ve çok değişkenli varyans analizini kullanmışlardır. Duygu tanıma oranını maksimize eden ilgili özellikler listesinden bir özellik alt kümesi seçmek için ikili kromozom tabanlı bir genetik algoritma önermişlerdir. Chowdary vd. [3] yüz ifadeleri ve beyin dalgalarını (EEG) içeren veri setlerini kullanarak derin öğrenme modelleri oluşturmuşlardır. Yüz ifadeleri için önceden eğitilmiş bir derin öğrenme modeli (örneğin, VGGFace) ve EEG sinyalleri için özel bir derin öğrenme mimarisi kullanarak duygusal durumları tahmin etmişlerdir. Balti vd. [4] insan pozunu tahmin etmek, yürüyüş özelliklerini çıkarmak ve videoları analiz etmek için derin öğrenmeye dayalı bir model önermişlerdir. Normal ve patolojik yürüyüşü sınıflandırmak için geleneksel sinir ağlarını kullanmışlardır. Park vd. [5] spor analizi sonuçlarını geliştirmek için birden fazla rastgele orman türü kullanan bir insan pozu tahmin algoritması önermişler ve sporcunun anatomisindeki eklemlerini gösteren bir tür iskelet modeli kullanmışlardır. Wang vd. [6] 3 boyutlu poz tahmini için mevcut derin öğrenme tabanlı

çalışmaların kapsamlı bir incelemesini sunarak, bu yöntemlerin avantaj ve dezavantajlarını tartışmışlardır. Ota vd.[7] OpenPose kütüphanesinin eklem noktalarını tespit etme hassaslığını test etmek için 24 katılımcı ile bir çalışma yapmışlardır. Çalışmada katılımcıların görüntüleri alınmış ve eklem bölgelerini tespit ederek karşılaştırmışlardır. Si ve Liu [8] iskelet model verilerinin özelliklerinin anlaşılmasına dayanarak, insan vücut yapısıyla birleştirilmiş bir mesafe özelliği ve açı özelliği modeli önermektedir. Deneysel sonuçlarla mesafe özelliğinin ve açı özelliği değerinin temel olarak mesafe değişiminden etkilenmediğini tespit etmişlerdir. Verma vd. [9] vücut duruşunun analizini yaptığı çalışmalarında belirli bir görüntü için bir dizi vücut parçası çifti arasındaki en optimum mesafeleri hesaplamış ve karmaşık eylemlerin temsilini öğrenmek için derin sinir ağlarını kullanmışlardır. Hyppönen vd. [10] Progresif Miyoklonus Epilepsi Tip 1 (EPM1) hastalığının tespitini hasta insanların pozunu ve vücut hareketi analizleri, vücut kilit noktalarını belirleme, hareket düzgünlüğü ve hızını analiz ederek geliştirdikleri model ile yapmışlardır. Xu vd. [11] 2B algılamaya koşullanmış monoküler 3B poz tahmini için simetri kaybı ve bağlantı açısı sınırı gibi sorunları düzeltmek amacıyla bir insan pozunu uygulaması geliştirmişlerdir. 2B girdilerin ve 3B çıktılarının statik/dinamik yapısını eşzamanlı olarak dikkate alan derin bir kinematik analiz sistemi oluşturmuş ve derin modeller entegre etmişlerdir. Bu sayede 2B veride bulunan gürültülü verilerden kurtulmayı, kayıp paket oranını düşürmeyi amaçlamışlardır. Rohan vd. [12] hasta üzerine giydirilen sensörler ve hastayı izleyen kameralar ile hasta kas hareketleri, yürüyüş kuvveti gibi değerleri ölçmüşlerdir. CNN algoritmasını kullanarak bir yapay zekâ modeli geliştirmişlerdir.

Paudel vd. [13] endüstride çalışan insanların çalışma alanlarındaki yanlış ve sağlıklarına zarar veren duruşların tespiti amacıyla bir model geliştirmişlerdir. Model, 3B insan pozunu tahmini yoluyla çalışanların ergonomik duruşlarının risk analizini gerçekleştirmek için yeni bir model önermektedir. Hareketler esnasında insan vücudu eklemlerinin değerlerini analiz etmişlerdir. Bu sayede işçilere göre en güvenilir vücut bükülme açılarını hesaplamak için vücut açısı güvenilirlik kararı (BARD) yöntemini tanıtmışlardır. Stenum vd. [14] insanların yürüyüşlerini analiz etmek için insan pozunu tanıma modeli önermişlerdir. OpenPose ile ölçülen uzamsal-zamansal ve yönsel analizde kinematik yürüyüş parametrelerini, sağlıklı yetişkinlerin yürümesinden eşzamanlı olarak kaydedilen üç boyutlu hareketleri ile karşılaştırmışlardır. Kim vd. [15] evde yalnız

yaşayan yaşlıların düşme ve kendilerini yaralama riskiyle karşı karşıya olması nedeni ile duruşlarını otomatik olarak izleyen ve tanıyan bir mobil robota ihtiyaç duyabileceğini düşünmüşlerdir. Bu soruna yönelik 2B ve 3B insan pozlarının eklem açılarını kullanarak yapay zekâ destekli bir model önermişlerdir. Wang vd. [16] vücut hareketlerinden duyguları sürekli olarak tanıyan gerçek zamanlı bir model geliştirmişlerdir. Verileri sınıflandırmak için akış tabanlı yarı denetimli uyarlanabilir rastgele orman yöntemini kullanmışlardır. Santhoshkumar ve Geetha [17] insan vücut hareketlerinden ileri beslemeli derin evrişimli sinir ağı mimarisini kullanarak duygu tanıma ve duygusal durumların tespiti için bir model önermişlerdir.

Sonuç olarak, duygusal durumları tahmin etmek için derin öğrenme tekniklerinin insan vücut hareketlerini analiz etmedeki potansiyeli üzerine yapılan çalışmalar, araştırmacılara şu önerilerde bulunmaktadır:

**Çoklu Öznitelikli (Multimodal) Veri Kullanımı:** İnsan vücut hareketleri, ses tonu, yüz ifadeleri ve diğer çoklu öznitelikli veriler bir araya getirilerek daha doğru duygusal durum tahminleri yapılabilir. Bu, duygusal ifadelerin karmaşıklığını daha iyi yakalayabilir ve doğruluğu artırabilir [18].

**Derin Öğrenme Modellerinin Uyarlanabilirliği:** Derin öğrenme modelleri, büyük veri setlerinde eğitildiğinde daha iyi performans gösterir. Bu nedenle, geniş ve çeşitli veri setlerini kullanarak modelleri eğitmek, genel doğruluk düzeyini artırabilir [19].

**Uygun Özellik Mühendisliği:** Verilerden uygun özelliklerin çıkarılması, derin öğrenme modelleri için kritiktir. Özellik mühendisliği, doğru özelliklerin seçilmesi ve modelin bu özellikleri doğru bir şekilde öğrenmesine yardımcı olabilir [20].

**Gerçek Zamanlı Uygulamalar:** Derin öğrenme modelleri, gerçek zamanlı duygusal analiz uygulamalarında kullanılacak hızda olmalıdır. Bu, interaktif uygulamalar ve insan-makine etkileşimi için önemlidir [21].

**Yeni Veri Toplama Yöntemleri:** Yeni ve daha hassas sensörler, insan vücut hareketlerini daha iyi yakalayabilir. Bu, daha zengin ve doğru veri setlerinin toplanmasına olanak tanır [22].

**Daha Derin ve Karmaşık Modeller:** Duygusal durumları tahmin etme konusunda derin öğrenme modellerinin karmaşıklığı artırılabilir. Örneğin, çoklu dalgaları (örneğin, EEG ve EKG) veya

mekânsal ilişkileri (örneğin, vücut parçaları arasındaki ilişkiler) yakalayabilen daha karmaşık mimariler geliştirilebilir. Bu, daha karmaşık duygusal ifadeleri anlayabilmek anlamına gelir [23].

**Transfer Öğrenme ve Uyumluluk:** Başka bir alanda eğitilmiş modellerin (örneğin, görüntü tanıma) duygusal analize uygulanabilmesi için transfer öğrenme teknikleri geliştirilmelidir. Ayrıca, farklı modellerin (örneğin, bir modelin ses analizi, diğerinin vücut hareket analizi) sonuçlarını birleştirmek için uyumluluk stratejileri de geliştirilmelidir [24].

**Yeni Veri Zenginleştirme Yöntemleri:** Var olan veri setlerini zenginleştirmek ve genişletmek için sentetik veri üretme yöntemleri kullanılabilir. Bu, modelin çeşitli durumları daha iyi anlamasına ve genelleştirmesine yardımcı olabilir [25].

**Hassasiyet ve Çeşitlilik:** Duygusal durumlar kültürel, bölgesel ve bireysel farklılıklara dayanabilir. Bu nedenle, modellerin hassasiyeti artırılmalı ve farklı kültürel bağlamlara uyum sağlama yetenekleri geliştirilmelidir. Bu, duygusal ifadelerin çeşitliliğini ve karmaşıklığını daha doğru bir şekilde yansıtabilir [26].

**İnsan-Makine Etkileşimi ve Rehabilitasyon:** Duygusal analiz, insana yönelik arayüzlerde (örneğin, konuşma tanıma sistemleri, öğrenme asistanları) ve rehabilitasyon uygulamalarında kullanılabilir. Modeller, duygusal ifadeleri anlayarak daha duyarlı ve etkili tepkiler verebilir, bu da insan-makine etkileşimini ve rehabilite edici süreçleri geliştirebilir [27].

**Eğitim ve Psikoloji Alanlarıyla İşbirliği:** Duygusal analiz çalışmaları, eğitim ve psikoloji alanlarındaki uzmanlarla işbirliği yaparak daha derinlemesine anlayışlar geliştirebilir. Bu iş birlikleri, özellikle eğitimde ve zihinsel sağlık uygulamalarında, duygusal analizin daha etkili bir şekilde kullanılmasını sağlayabilir [28].

Bu öneriler, duygusal durum tahmini üzerine yapılan derin öğrenme çalışmalarının gelecekteki yönelimlerini ve odak noktalarını belirlemek için önemli bir rehber sağlar. Bu alandaki ilerlemeler, daha iyi insan-makine etkileşimi, duygusal zekâ uygulamaları ve sosyal etkileşim alanlarında kullanılabilir. Ayrıca bu alandaki ilerlemeler, teknolojiyle insan etkileşimini daha anlamlı ve duygusal olarak zengin hale getirmeye yönelik büyük potansiyele sahiptir.

Literatürden ve yukarıdaki önerilerden derin öğrenme tekniklerini kullanarak insan vücut hareketlerinden doğru ve derinlemesine duygusal durumları çıkarabilen bir modelin geliştirilmesinin önemi anlaşılmaktadır. Geliştireceğimiz model, video verilerinden vücut hareketlerini hassas bir şekilde analiz edebilecek, jestleri, mimikler ile diğer vücut dilini yorumlayabilecek ve bu bilgileri kullanarak kullanıcının duygusal halini tahmin edebilecektir. Aynı zamanda geliştireceğimiz bu model gerçek dünya uygulamalarında, örneğin terapötik robotlar, interaktif eğitim sistemleri ve insan-makine etkileşimli oyunlar gibi alanlarda önemli bir etki yaratabilir.

### 1.1. Amaç (Purpose)

Bu çalışmanın amacı, insan vücut hareketlerini derin öğrenme teknikleri ile analiz ederek bireyin duygusal durumlarını tahmin eden yeni bir makine öğrenmesi modeli geliştirmektir. Geleneksel duyu analizi yöntemleri genellikle yüz ifadelerine odaklanmaktadır. Literatürden vücut hareketlerinin duygusal durumları yansıtabileceği anlaşılmaktadır. Bu çalışmada insan vücut tespiti, iskelet modelin oturtulması ve öznitelik çıkarımı gibi adımlarla vücut hareketleri belirlenmiştir. Eğitilmiş bir sınıflandırma algoritması kullanarak vücut hareketlerinden duygusal durumlar tahmin edilmiştir.

## 2. MATERYAL VE YÖNTEM (MATERIALS AND METHODS)

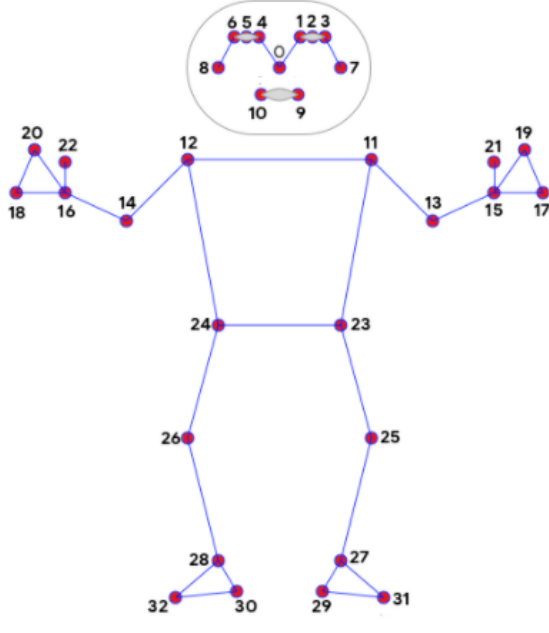
Bu çalışmada deneme ortamı olarak Jupyter Notebook, derleme ve çalıştırma ortamı olarak PyCharm kullanılmıştır. Kullanılan araçlar en genel olarak Python kütüphanelerinden OpenCV, NumPy, MediaPipe, Pandas, OS, Math, TensorFlow, Keras, Scikit-learn'dür. Sonuçları doğrulamak için evrişimli sinir ağları (CNN), destek vektör makineleri (DVM), K-en yakın komşular (k-NN), lojistik regresyon (LR) ve rasgele orman (RO) algoritmaları kullanılmıştır. Kızgın, korkmuş, mutlu, nötr ve üzgün olarak tanımladığımız beş ana duyguya ait resimler, internet ortamından elde edilmiştir. Her duygu için 200 olmak üzere

toplamda 1000 adet resim kullanılmıştır. İnsanların verdiği pozlardan çıkarım yapılmaktadır, eksik vücut bölümü bu verilerin alınmaması demektir ve bu istenmeyen bir durumdur. Bundan dolayı insan vücudunun tüm kısımları eksiksiz görünmelidir ve veriler toplanırken, insan vücudunun görünmeyen kısmının olmamasına dikkat edilmiştir.

Veri ön işleme sürecinde veriler ilk olarak klasörlerde toplanmıştır. Kullanılan resimler öznitelik çıkarımından önce MediaPipe kütüphanesinin en gelişmiş iskelet modeli olan "Pose" kütüphanesi ile ön işlemeye tabi tutulmuştur. Bu kütüphane kendi içinde modelin türüne göre 256x256 veya 320x320 piksel olacak şekilde yeniden boyutlandırma yapar ardından resmi RGB'ye dönüştürerek pikselleri 0-1 arası ölçeklendirir. Ayrıca parametreler ile bu değerler özelleştirilebilir. Resimler bu sayede standartlaştırılır. Doğru çıkarılan öznitelikler istenen ve hedeflenen tahmin doğruluğuna ulaşmak için önemlidir.

### 2.1. İskelet Model ile Özniteliklerin Çıkarılması (Extracting Features with Skeleton Model)

İskelet model, insan pozlarındaki uç noktaları birbirine bağlayarak koordinat noktaları ile bir düzleme oturtan yapıdır. MediaPipe kütüphanesinin iskelet modelinde bir insan pozu için 33 nokta tanımlanmıştır. Bu noktalar ve ne oldukları Görsel 1'de gösterilmiştir. Bu 33 nokta, iskelet modelden alınacak özniteliklerin ana sınıfıdır. Buradaki her bir noktayı 3 boyutlu bir düzlem gibi düşünüp, değerleri alınmaktadır. Bundan dolayı bu 33 noktanın her birinin x,y,z ve vis değerleri hesaplanmaktadır. Koordinat düzlemindeki 3 yönü belirten x,y ve z için, bu noktaların 3B düzlemindeki değeri hesaplanır ve alınır. Buradaki 4.öznitelik türü olan vis (visible) yani görünürlük değeridir. 0 ile 1 arasında bir değer alır ve ilgili vücut noktası hiç görünmüyor ise 0'a yakın bir değer alır. Eğer görülen bir nokta ise 1'e yakın bir değer alır. 33 noktanın her biri için bu x,y,z ve vis değerleri çarpazlanır ve 132 öznitelik oluşturulur. Bu, her bir insan pozundan 132 öznitelik çıkarıldığı anlamına gelir.



- |                 |                        |
|-----------------|------------------------|
| 0. burun        | 17. sol_serçe_parmağı  |
| 1. sol_göz_içi  | 18. sağ_serçe_parmağı  |
| 2. sol_göz      | 19. sol_işaret_parmağı |
| 3. sol_göz_dışı | 20. sağ_işaret_parmağı |
| 4. sağ_göz_içi  | 21. sol_başparmak      |
| 5. sağ_göz      | 22. sağ_başparmak      |
| 6. sağ_göz_dışı | 23. sol_kalça          |
| 7. sol_kulak    | 24. sağ_kalça          |
| 8. sağ_kulak    | 25. sol_diz            |
| 9. ağız_sol     | 26. sağ_diz            |
| 10. ağız_sag    | 27. sol_ayak_bileği    |
| 11. sol_omuz    | 28. sağ_ayak_bileği    |
| 12. sağ_omuz    | 29. sol_topuk          |
| 13. sol_dirsek  | 30. sağ_topuk          |
| 14. sağ_dirsek  | 31. sol_ayak_ucu       |
| 15. sol_bilek   | 32. sağ_ayak_ucu       |
| 16. sağ_bilek   |                        |

**Şekil 1.** Eşleşme noktalarının iskelet modelde gösterimi (Representation of matching points in the skeleton model) [29]

Şekil 1’de bir iskelet modelde bulunan eşleşme noktaları gösterilmiştir. Bunlar, insan vücudu üzerine oturtulan bir koordinat sistemi noktalarıdır. Her bir noktanın doğru şekilde insan vücuduna oturtulması ile öznitelikleri alınabilir. Aksi takdirde, resim kapsama uygun olsa da iskelet model öznitelikleri kapsama uymaz. Bu da uzun vadede modelin doğruluk oranını düşürecektir. Resimlerin doğru bir şekilde modele oturtulabilmesi için piksel boyutları birbirlerine yakın değerlerde olmalıdır. Ayrıca arka plan ve insan pozunu arasında belirgin bir fark olması gerekmektedir. Sıradaki işlem insan pozunun tanınması iskelet modelin tam olarak pozun üzerine oturtulmasıdır.

İnsan pozunun tanınması için ilk olarak poz üzerindeki vücut noktaları kütüphane tarafından tespit edilerek ilgili noktalar yerleştirilmiştir. İnsan iskeletine dönüştürmek için bu noktalar çizgilerle birleştirilmiştir. İki nokta arasındaki mesafe “Öklidyen mesafe” formülü ile hesaplanmıştır. Bu formül, iki nokta arasındaki doğru çizginin uzunluğunu hesaplar ve 3 boyutlu uzayda da kullanılabilir. Bir noktanın bir çizgiye olan konumunu belirlemek için “Noktanın çizgiye olan konumu” hesaplanmıştır. Bu hesaplama, bir noktanın bir çizgi üzerinde mi yoksa çizginin solunda veya sağında mı olduğunu belirler. Bu sayede noktalar arasındaki açılar ve çizgiler yerleştirilebilir. Bir çizginin ortak noktaları arasındaki açıyı hesaplamak için ise “İki çizgi arasındaki açı” formülü kullanılmıştır. Bu formül,

iki çizginin vektörlerini kullanarak kosinüs teoremiyle açıyı hesaplar. Daha sonra bir çizginin yatayla yaptığı açıyı hesaplamak için “Çizgi ve yatay arasındaki açı” formülü kullanılmıştır. Bu açı, çizginin eğimini yani yatay eksene olan açısını ifade eder. Bu sayede pozun iskeleti çıkartılmış olunur. Son olarak, çıkarılan öznitelikler bir CSV dosyasına kaydedilmiştir. Bu CSV dosyasındaki her bir satır, öznitelikleri çıkarılmış bir resimdir. İlk satır ise, özniteliklerin tanımlarının yani duyguların bulunduğu satırdır.

## 2.2. Özniteliklerin Etiketlenmesi (Tagging Attributes)

Öznitelikleri çıkarılan resimlerin, hangi duygu pozuna ait olduğunun makine tarafından öğrenilebilmesi için etiketlenmesi gerekmektedir. Etiket, o özniteliğin hangi duygu pozunu taşıdığını gösterir. Etiketlerin sayısının belirlenmesi, veri setindeki verilerin sayısı ile ilişkilidir. Tüm etiketlerin ilgili CSV satırına kaydedilmesi gerekir. Bunun için özniteliklerin kaydedildiği CSV dosyasının en sonuna yeni bir sütun eklenir. Bu sütun baştan sona yazısal değişkendir. Makineler, sadece sayısal veriler üzerine işlem yapar. Sırası ile kızgın, korkmuş, mutlu, nötr, üzgün etiketleri için 0, 1, 2, 3, 4 sayıları sütunlarda bu yazısal değerler yerine yazılırsa, makine sayıların büyüklüğüne göre yanlış çıkarım yapacaktır. Örneğin kızgın duygusunun sayısal değeri 0 iken, üzgün duygusunun sayısal değeri 4 olmuş olur. Bu

durumda üzgün çıkarımını, kızgın çıkarımından daha önde tutabilir. Bu ve benzeri durumların önüne geçmek için ise Sckit-Learn kütüphanesinin içerisinde olan OneHotEncoder ve LabelEncoder sınıfları kullanılmıştır.

OneHotEncoder, çok sınıflı kategorik verileri sayısallaştırmak için kullanılan bir metod'dur. Bu metod'da her bir kategori ayrı bir özellik olarak temsil edilir ve bu özelliklerin değerleri 0 veya 1 olur. Örneğin kızgın, korkmuş, mutlu, nötr, üzgün gibi duygu etiketlerinin OneHotEncoder ile sayısallaştırılmış hali Tablo 1'de verilmiştir.

**Tablo 1.** OneHotEncoder ile sayısallaştırılmış duygu etiketleri (Digitized emotion labels with OneHotEncoder)

Duygu	kızgın	korkmuş	mutlu	nötr	üzgün
Kızgın	1	0	0	0	0
Korkmuş	0	1	0	0	0
Mutlu	0	0	1	0	0
Nötr	0	0	0	1	0
Üzgün	0	0	0	0	1

LabelEncoder ise tek sütunlu kategorik verileri sayısal etiketlere dönüştürmek için kullanılmıştır. Her bir kategoriye benzersiz bir sayı atanmıştır. 132 sütun bulunan CSV dosyasına etiket

sütunları olan 5 sütun daha eklenerek 137 sütuna çıkarılmıştır. Şekil 2'de veri setinin ilk 4 kaydı gösterilmiştir.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	NOSE_x,NOSE_y,NOSE_z,NOSE_vis,LEFT_EYE_INNER_x,LEFT_EYE_INNER_y,LEFT_EYE_INNER_z,LEFT_EYE_INNER_vis,LEFT_EYE_x,LEFT_EYE_y,LEFT_EYE_z,LEFT_EYE_vis,LEF															
2	0.5480905771255493,0.2401376366615295,-0.9046703577041626,0.9999902248382568,0.5611100792884827,0.2250009179115295,-0.8805035352706909,0.9999676942825316															
3	0.4835946261882782,0.0982447117567062,-0.8734963536262512,0.9999909996986388,0.5032166838645935,0.0829728618264198,-0.8343377709388733,0.999970555305481,															
4	0.5256025195121765,0.1147461459040641,-1.047596096992493,0.9999905228614808,0.5438736081123352,0.0965046286582946,-0.9985254406929016,0.9999712705612184,															
5	0.5439168214797974,0.1091045811772346,-1.0460264682769775,0.9999913573265076,0.5656952857971191,0.0939799472689628,-0.9741354584693908,0.9999738931655884															
6																

**Şekil 2.** CSV dosyasındaki veriler (Data in CSV file)

### 2.3. Model Eğitimi (Model Training)

Bir model eğitimi için öncelikle model eğitim fonksiyonlarının, parametrelerinin ve hiper parametrelerinin bilinmesi gerekmektedir. Kullanılan kütüphanelere göre bu parametre ve fonksiyonlar değişiklik gösterebilir. Ancak her model için kullanılan hiper parametreler aynıdır.

Bunlardan birisi batch\_size hiper parametresidir. Batch\_size, bir eğitim döngüsünde aynı anda işlenen örneklerin sayısını belirleyen bir parametredir. Batch\_size, bir trade-off oluşturan bir parametredir. Trade-off ise, bir seçim veya karar verme sürecinde birbirine zıt olan iki veya daha fazla öğe arasında denge kurma veya uzlaşma anlamına gelir. Trade-off, bir öğenin artışı veya iyileştirilmesiyle diğer öğenin azalması veya bozulması arasındaki ilişkiyi ifade eder. Bu çalışmada batch\_size değeri 64 olarak belirlenmiştir. Bir diğer hiperparametre learning\_rate (öğrenme oranı)'dir. Learning\_rate, modelin her bir güncelleme adımında ağırlıklarının ne kadar değiştireceğini belirler. Bir makine

öğrenme modeli, hedef fonksiyonunu minimize etmek için ağırlıklarının güncellerken, learning\_rate bu güncellemelerin büyüklüğünü kontrol eder. Bu çalışma için learning\_rate değeri 0.001'dir. Bir diğer hiper parametre ise epoche sayısıdır. Epoch, makine öğrenmesi ve derin öğrenme modellerinde eğitim verilerinin tamamının model tarafından geçirilmesi anlamına gelir. Bir epoch, modelin bir kez tüm eğitim verilerini kullanarak öğrenme ve güncelleme işlemini tamamladığı bir iterasyonu ifade eder. Her epoch, modelin ağırlıklarını veya parametrelerini güncellediği bir eğitim döngüsünü temsil eder. Bu çalışmada epoch sayısı karşılaştırma yapılması amacı ile 10,20,30,50,75 ve 100 olarak belirlenmiştir. Her bir epoch değeri için ayrı model çıktıları oluşturulmuştur. Tüm batch\_size değerleri için epoch sayısı arttıkça hata yüzdesinin düştüğü görülmüştür.

Ayrıca modelin doğruluk oranının epoch sayısı ile doğru orantılı bir şekilde arttığı görülmüştür. Kayıp paket miktarı ise düşmüştür. Her iki grafikte de bir zaman sonra duraksama meydana gelmiştir. Bu da

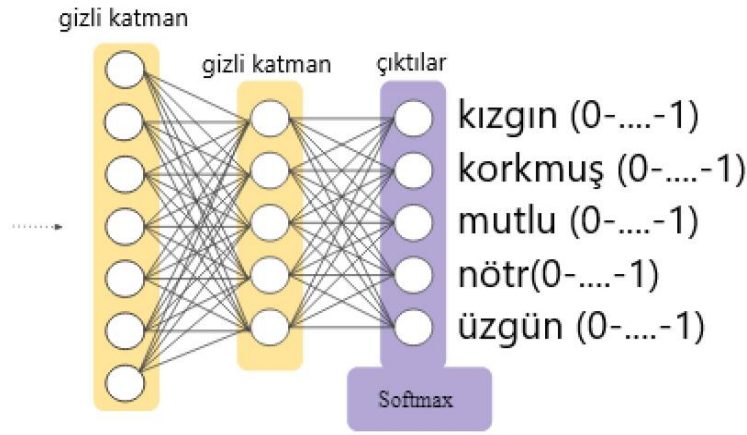
epoch sayısının çok yüksek değil, optimum düzeyde ayarlanması gerektiği anlamına gelmektedir.

Bu çalışmada eğitim modeli olarak sequential sınıfı seçilmiştir. Sequential, Keras kütüphanesinde yer alan bir sınıftır. Bu sınıf, yapay sinir ağlarını oluşturmak için kullanılan bir model türünü temsil eder. Modelin her katmanı ardışık olarak bir önceki katmanın çıktısını alır ve kendisinden sonraki katmana girer olarak verir.

Bu aşamada sequential bir model oluşturulup, yoğun (dense) katman eklenmiştir. Bu katmanda her bir nöron, bir önceki katmandaki tüm nöronların çıktılarından etkilenir. Giriş katmanının boyutunu ayarlamak için, özniteliklerin toplam sayısı alınmıştır. Aktivasyon fonksiyonu olarak bu yoğun katmana ReLU uygulanmıştır. Bu fonksiyon, modelin yapılandırılması sırasında ilk katman olarak kullanılır ve giriş verilerini doğru boyutta kabul edebilecek bir yoğun katmanı modelin yapısına ekler. Böylece, model, verileri giriş katmanından alarak daha karmaşık işlemler yapabilen diğer katmanlara aktarabilir. Bu 132 nöronlu yoğun katmanın üstüne bir yoğun (dense) katman daha eklendi ve bu katman sadece 1 adet nörona sahip olacak şekilde ayarlanmıştır. Bu, çıkış katmanının yalnızca bir çıktı üreteceği anlamına

gelmektedir. Ayrıca, aktivasyon fonksiyonu olarak da sigmoid kullanılmıştır.

Artık yukarıdaki tanımlama ile 132 nöronlu giriş ve 1 nöronlu çıkışa sahip modelin ilk aşaması tamamlanmış oldu. Sigmoid fonksiyonunun çıkışının 1 olması, sınıfların toplam tahmin değerlerinin 1 olduğu anlamına gelir. Yani ileride tanımlanacak olan 5 sınıfa ait 5 çıkışın değerlerinin toplamının 1 olacağını ifade eder. Sonraki aşama için, öznitelik sayılarına sahip 132 nöron girişli, aktivasyon fonksiyonu ReLU olan bir model daha oluşturulmuştur. Bu modele, bir yoğun (dense) katmanı daha eklenmiş ve bu katman 5 adet nörona sahip olacak şekilde ayarlanmıştır. Bu, çıkış katmanının 5 farklı sınıf için, yani duygu etiketlerimiz için olasılık değerleri üreteceği anlamına gelmektedir. Bu aşamanın son adımı olarak da softmax fonksiyonu uygulanmıştır. Softmax fonksiyonu, çıktıları normalleştirerek her sınıf için bir olasılık dağılımı üretmektedir. Bu da, çok sınıflı çıkışın olduğu bir sınıflandırma problemindeki tüm sınıfların olasılıklarını temsil ettiği anlamına gelmektedir. Bu kod, modelin yapılandırılması sırasında son katman olarak kullanılır ve modelin çok sınıflı sınıflandırma problemlerindeki çıkışını tanımlar. Böylece, toplam değerleri 1 olacak şekilde oranları dağıtılacak olan 5 çıkış nöronlu modelimiz tanımlanmıştır.



**Şekil 3.** Softmax 5 nöronlu çıkış şekli (Softmax 5 neuron output shape)

Şekil 3’de Softmax’ın 5 nöronlu çıkış şekli verilmiştir. Sigmoid ile softmax beraber kullanıldığında, değerleri 0 ile 1 arasında olan 5 çıkış nöronu oluşur. Bu nöron çıktılarının toplam değeri 1’dir.

#### 2.4. Modelden Gerçek Zamanlı Çıkarım Yapılması (Real-Time Inference from the Model)

Gerçek zamanlı çıkarım, oluşturulan ve çıktısı alınan modelin öğrendiği veriler ile kameradan anlık olarak alınan görüntü üzerinde tahmin yapması ve çıktı vermesi anlamına gelir. Model eğitiminde resimler üzerinde iskelet model uygulama işlemi yapılmıştı. Daha sonra ise çıkarılan iskelet model öznitelikleri kaydedilmişti.



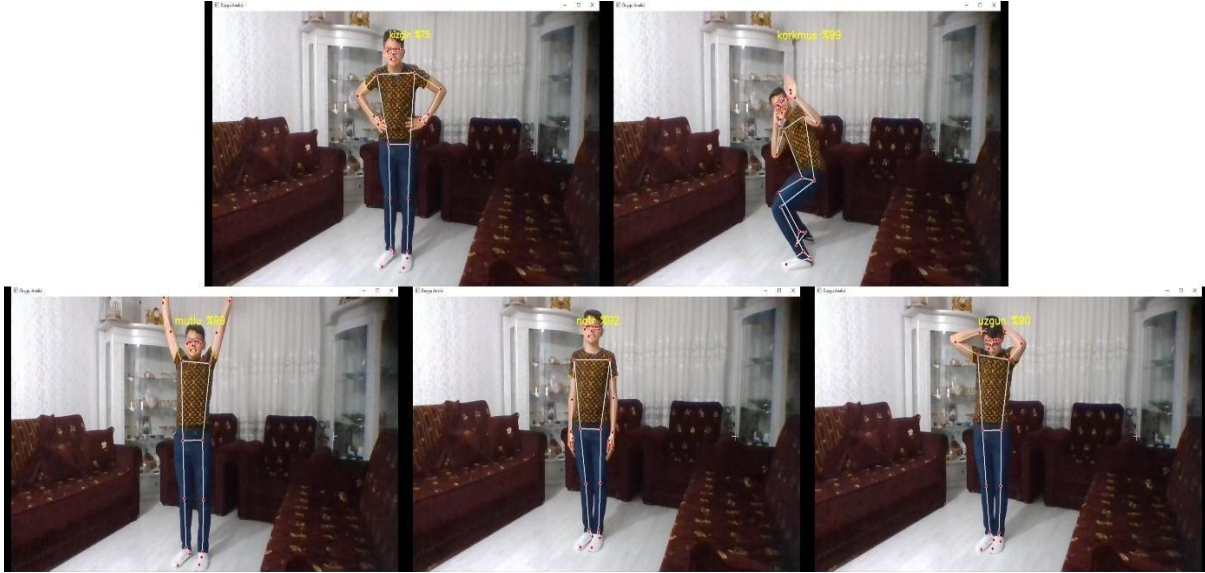
Burada da her bir çerçeve üzerinde iskelet modeli işlenir, alınan öznitelikler ile çağrılan model üzerinde tahmin işlemi gerçekleştirilir.

**Tablo 2.** Gerçek zamanlı duygu çıkarımı (Real-time emotion extraction)

Görsel 4'ten alınan verilerin çıkarım oranları ve çıkarım sonucu %						
Satır	Kızgın	Korkmuş	Mutlu	Nötr	Üzgün	Çıkarım Sonucu
1	%0	%37.8	%0	%62.1	%0	Nötr
2	%0	%41.0	%0	%58.9	%0	Nötr
3	%0	%47.4	%0	%52.5	%0	Nötr
4	%0	%68.7	%0	%31.2	%0	Korkmuş
5	%0	%91.2	%0	%08.7	%0	Korkmuş

Tablo 2'de gerçek zamanlı işlenen insan pozundan yapılan öznitelik çıkarımında bazı duygu etiketlerinin 0'dan küçük değer almasından dolayı 0 değerini aldığı, bazı duygu etiketlerinin ise olasılık değerleri çerçevesinde 0'dan büyük değerler aldığı görülmektedir. İlk satırda kızgın, mutlu ve üzgün

duyguları 0 değerini almışken, korkmuş duygusu %37.8 oranında, nötr duygusu %62.1 oranında değer almıştır. İlk satır, o çerçeve için yapılan işlemlerden çıktı olarak "nötr" duygusunun çıkarımını yapmıştır. Bu sayede insan pozlarından gerçek zamanlı çıkarım yapılmış olunur.



**Şekil 5.** Gerçek zamanlı çıkarım (Real-time emotion extraction)

Şekil 5'de gerçek zamanlı çıkarımın insan vücudu üzerinde nasıl durduğu ve pozdan ürettiği tahmin değeri görülmektedir. Modelden gerçek zamanlı yapılan çıkarımların görüntülerinin derlenmiş halidir. Duygulara ait çıkarımlar ve yüzdeler tespit edilen insan yüzünün üstüne konulmuştur. İskelet modeli ise isteğe bağlı olarak konulup kaldırılabilir.

### 3. BULGULAR (FINDINGS)

Bu çalışmada seçilen batch-learning tekniği, modelin sürekli öğrenmeye kapalı olduğu ve tüm eğitim verilerinden bir kere eğitimin yapıldığı, eğitimin bir süreç değil sonuç olduğu yaklaşımdır.

Bu yaklaşımın kullanılmasının nedeni, insan pozlarını anlama zorluğundan kaynaklıdır. Çünkü insan pozları sınırsız hareket kombinasyonlarına sahiptir. Aykırı verilere yakalanılmaması için ise etiketlere göre incelenmesi gerekir. Bu nedenle, modelin eğitimi aktif bir süreç yerine toplu olarak gerçekleştirilmiş, insan pozlarının anlaşılmasındaki zorluk ve etiketlere göre incelenmesi gerekliliği göz önünde bulundurularak, modelin doğruluk oranının yükseltilmesi amaçlanmıştır.

### 3.1. Performans Ölçüm Modelleri ve Sonuçları (Performance Measurement Models and Results)

edilen duygu tespiti sınıflandırma sonuçları verilmiştir.

Tablo 3, 4, 5, 6 ve 7’de CNN, DVM, KNN, Lojistik Regresyon ve Rastgele Orman algoritmaları ile elde

**Tablo 3.** CNN sınıflandırması (CNN classification)

CNN Sınıflandırması		
Duygu Durumu	Sınıflandırma Doğruluğu	Yüzde
Kızgın	0.983	%98.30
Korkmuş	0.989	%98.90
Mutlu	0.994	%99.40
Nötr	0.950	%95.00
Üzgün	0.978	%97.80

Tablo 3 incelendiğinde en düşük doğruluğa sahip duygunun %95.00 ile “nötr” duygusu olduğu, en yüksek doğruluğa sahip duygunun %99.40 ile

“mutlu” duygusu olduğu görülmektedir. CNN ile doğrulamanın ortalama doğruluk değeri ise %97.88’dir.

**Tablo 4.** DVM sınıflandırması (DVM classification)

DVM Sınıflandırması		
Duygu Durumu	Sınıflandırma Doğruluğu	Yüzde
Kızgın	0.989	%98.90
Korkmuş	0.984	%98.40
Mutlu	1.0	%100
Nötr	0.995	%99.50
Üzgün	0.978	%97.80

Tablo 4 incelendiğinde en düşük doğruluğa sahip duygunun %97.80 ile “üzgün” duygusu olduğu, en yüksek doğruluğa sahip duygunun %100 ile “mutlu” duygusu olduğu görülmektedir. DVM ile

doğrulamanın ortalama doğruluk değeri ise %98.92’dir.

**Tablo 5.** KNN sınıflandırması (KNN classification)

KNN Sınıflandırması		
Duygu Durumu	Sınıflandırma Doğruluğu	Yüzde
Kızgın	1.0	%100
Korkmuş	0.934	%93.40
Mutlu	0.989	%98.90
Nötr	1.0	%100
Üzgün	0.945	%94.50

Tablo 5 incelendiğinde en düşük doğruluğa sahip duygunun %93.40 ile “korkmuş” duygusu olduğu, en yüksek doğruluğa sahip duyguların %100 ile

“kızgın” ve “nötr” duyguları olduğu görülmektedir. KNN ile doğrulamanın ortalama doğruluk değeri ise %97.36’dır.

**Tablo 6.** Lojistik regresyon sınıflandırması (Logistic regression classification)

Lojistik Regresyon Sınıflandırması		
Duygu Durumu	Sınıflandırma Doğruluğu	Yüzde
Kızgın	0.995	%99.50
Korkmuş	0.972	%97.20
Mutlu	1.0	%100
Nötr	0.995	%99.50
Üzgün	0.989	%98.90

Tablo 6 incelendiğinde en düşük doğruluğa sahip duygunun %97.20 ile “korkmuş” duygusu olduğu, en yüksek doğruluğa sahip duygunun %100 ile

‘mutlu’ duygusu olduğu görülmektedir. Lojistik Regresyon ile doğrulamanın ortalama doğruluk değeri ise %99.02’dir.

**Tablo 7.** Rastgele orman sınıflandırması (Random forest classification)

Rastgele Orman Sınıflandırması		
Duygu Durumu	Sınıflandırma Doğruluğu	Yüzde
Kızgın	1.0	%100
Korkmuş	0.989	%98.90
Mutlu	1.0	%100
Nötr	0.983	%98.30
Üzgün	0.989	%98.90

Tablo 7 incelendiğinde en düşük doğruluğa sahip duygunun %98.30 ile “nötr” duygusu olduğu, en yüksek doğruluğa sahip duyguların %100 ile “kızgın” ve “mutlu” duyguları olduğu görülmektedir. Rastgele Orman ile doğrulamanın ortalama doğruluk değeri ise %99.22’dir.

Sonuç olarak Rastgele Orman algoritması tüm duyguları yüksek bir doğrulukla tahmin etme yeteneği göstermiştir. Kızgın ve Mutlu duygularının sınıflandırmada tüm algoritmalar çok başarılı olmuştur. Bu da kızgın ve mutlu duygularının diğerlerine kıyasla daha belirgin ve kolay tanımlanabilir bir yapıya sahip olduklarını göstermektedir. Nötr ve korkmuş duyguları ise diğer duygulara kıyasla daha düşük sonuç vermiştir. Özellikle korkmuş duygusu hiçbir modelde en yüksek doğruluğa erişememiştir. Bunun nedeni, diğer duygulara göre aykırı yapıda olmasıdır. Diğer duygular genelde bir insanın dik bir şekilde poz vermesinden oluşurken korkmuş duygusu, kendini küçülterek korumaya çalışan bir insan poz şeklindedir.

### 3.2. Resim Üzerinde Test Sonuçları (Test Results on Picture)

Aynı resim üzerinde farklı epoch sayılarının etkisi Tablo 8’de gösterilmiştir. Tablo 8 incelendiğinde, kızgın, üzgün ve mutlu duygularının doğruluk oranının epoch sayısı ile doğru orantılı olarak arttığı, korkmuş duygusunun ilk 3 epoch modelinde pasif kaldığı ancak son 3 epoch modelinde yüksek doğruluk gösterdiği, nötr duygusunun ilk 3 epoch modelinde yarı pasif kaldığı ve son 3 epoch modelinde arttığı görülmüştür. Korkmuş duygusunun doğruluk oranının yüksek çıkmasının nedeni, diğer duygu pozlarına kıyasla çok daha geniş bir öznitelik korelasyonuna sahip olması olabilir. Korkmuş duygusu bu yüzden model eğitimlerinin düşük epoch değerlerinde pasif kalmıştır. Nötr duygusu ise diğer ana duygular ile benzerlikler taşıdığı için, makine öğrenimi düşük epoch sayılarında başarısız denilebilecek sonuçlar vermektedir.

Tablo 8’den 50 epoch değerinin 3 duygu için çok başarılı, 1 duygu için başarılı, 1 duygu için ise yetersiz kaldığı görülmektedir. Epoch 75 değerine çıktığında ise 5 duygu için de çok başarılı sonuçlar alındığı görülmüştür. Epoch değeri 100’e yükseltildiğinde, duyguların çoğunda doğruluk oranında bir artış gözlemlenmiştir. Fakat bu aşamada, modelin aşırı öğrenme sorununa karşı

dikkatli olunması gerekir. Bu sonuçlara göre epoch değeri için 100 ideal bir değerdir.

**Tablo 8.** Epoch doğruluk oranları (Epoch accuracy rates)

Epoch sayısı / Sınıflandırma doğruluğu %						
Duygu / Epoch sayısı	10	20	30	50	75	100
Kızgın	%46	%77	%85	%98	%99	%99
Korkmuş	%16	%18	%26	%95	%97	%99
Mutlu	%78	%96	%99	%99	%99	%99
Nötr	%44	%44	%50	%70	%91	%95
Üzgün	%58	%78	%80	%82	%93	%95

### 3.3. Gerçek Zamanlı Test ve Sonuçları (Real Time Testing and Results)

Modelin gerçek zamanlı test edilmesinin önemi, uygulama alanına en yakın test alanı olmasıdır. Modelin gerçek zamanlı performansını göstermekte ve ileriye dönük geliştirme işlemleri için ipuçları sunmaktadır. Buradaki veriler, gerçek zamanlı tahmin işleminde alınan insan pozlarından çıkarılan verilerin doğruluk oranlarının aralığını barındırır. Bu aralık pozlara göre değişen doğruluk oranlarının en düşük ve en yüksek değerleridir. Bant aralıklarına ve epoch değerine bakıldığında, epoch sayısının arttıkça gerçek zamanlı testin doğruluk aralıklarının da yükseldiği görülmektedir.

Tablo 9 incelendiğinde, kızgın ve nötr duygusunun başlangıçta pasif kaldığı görülmektedir. Epoch değeri 100 iken doğruluk oranı 90'ın altında olan tek duygunun kızgın olmasının nedeni, genel olarak nötr duygusuna çok yakın değerlere sahip

olmasındandır. Bununla birlikte 100 epoch değeri ile öğrenmede %80 olan en düşük doğruluk oranının iyi bir sonuç olduğu söylenebilir. Korkmuş duygusunun hızlı bir şekilde öğrenilmesinin nedeni korkmuş duygusunun özneliklerinin diğer duygulardan oldukça farklı olmasıdır. Korkmuş hariç diğer 4 duyguda gövde ve bacak yapısı birbirine çok benzemekle beraber, korkmuş hariç 4 duygu arasındaki öğrenme farkı kollar, omuzlar ve baş bölgesi ile yapılır. Buna karşın korkmuş pozunu veren kişi, görünür vücudunu küçültme ve elleri ile kendini koruma içgüdüleri sergiler. Bu doğruluk değerleri, eğitim setindeki insan pozlarına özenle uyum sağlayarak verilen pozlardan elde edilmiştir. Eğitim setindeki pozlara olabildiğince çok benzer pozlar verilmesinin nedeni, tıpkı doğrulama modellerinin yaptığı türden bir doğruluk testi yapma amacı taşımaktadır. Modelin gerçek zamanlı olarak verebileceği en yüksek değerlerin tespiti amaçlanmıştır.

**Tablo 9.** Gerçek zamanlı epoch doğruluk oranları (Real-time epoch accuracy rates)

Epoch sayısı / Sınıflandırma doğruluğu %						
Duygu / Epoch sayısı	10	20	30	50	75	100
Kızgın	%28-%31	%34-%40	%33-%35	%66-%75	%75-%80	%80-%87
Korkmuş	%35-%42	%51-%61	%76-%86	%98-%99	%98-%99	%98-%99
Mutlu	%70-%80	%80-%86	%91-%96	%91-%98	%98-%99	%98-%99
Nötr	%30-%40	%54-%58	%77-%80	%90-%92	%94-%97	%93-%98
Üzgün	%57-%60	%60-%65	%63-%73	%84-%91	%92-%97	%98-%99

### 4. SONUÇ VE TARTIŞMA (CONCLUSION AND DISCUSSION)

Bu çalışmada, gerçek zamanlı olarak vücut pozlarına dayalı ifade durumlarının tespiti için bir insan pozunu tanıma sistemi geliştirilmiş ve veri setindeki görüntüler ile farklı sınıflandırma yöntemleri, derin evrimsel sinir ağlarının

doğruluk oranları ve eğitim süreleri karşılaştırılmıştır. Bu çalışmada Keras, Tensorflow, MediaPipe ve OpenCV kütüphaneleri kullanılmıştır. Sınıflandırma algoritmalarından CNN, k-NN, DVM, LR ve RO algoritmaları kullanılmıştır. Ayrıca, dâhili kamera ile yüz tespiti için OpenCV kütüphanesinin derin sinir ağları (DSA) sınıfından yüz tanıma algoritması

kullanılmıştır. Veri depolama alanı olarak CSV uzantılı dosya depolama türü tercih edilmiştir. Bu dosyanın tercih edilmesinin nedeni ise, barındırdığı verilerin kolaylıkla incelenebilmesidir. Model çıktısını kaydetmek için ise “.h5” uzantılı yani “Hiyerarşik Veri Formatı 5” kullanılmıştır. Sonuç olarak en yüksek doğruluk oranının %97.80 ile rastgele orman ile elde edildiği görülmüştür. Çalışmadaki gerçek zamanlı doğruluk oranı ortalaması ise %94.50 olarak hesaplanmıştır.

Doğruluk sonucunun artması için veri seti büyüklüğü ve duygulara ait resimlerin sayısal dağılımının eşit olmasına, duygusal ifadelerin spesifik özelliklerinin kapsamları dışına çıkmamasına ancak kapsam dışına çıkmadan olabileceği her durum için veri setinde örneğinin bulunmasına, hiper parametrelerin veri seti büyüklüğüne göre özenle seçilmesine ve modelin gerçek zamanlı kullanımı için performansı yüksek bir bilgisayara ihtiyaç duyduğu görülmüştür. Çalışmada kullanılan veri seti büyüklüğü 1000 resim olmakla birlikte gerçek zamanlı testte en yüksek doğruluğun learning\_rate değeri 0.001, batch\_size değeri 64, epoch değeri 100 iken olduğu görülmüştür.

Bu çalışma, özellikle insan-makine etkileşimi, duygusal zekâ, psikoloji ve yapay zekâ alanlarında önemli katkılar sağlayabilir. Bu çalışmanın önemi, sadece teknolojik bir ilerleme olarak değil, aynı zamanda sosyal bağlamda da büyük bir anlam taşımaktadır. Örneğin, otizm spektrum bozukluğu gibi durumlarla ilgili bireylere daha iyi yardım sağlama, psikolojik rahatsızlıkları olan bireylerin tedavi ve destek süreçlerini geliştirme ve etkileşimli eğitim uygulamalarını daha kişiselleştirme gibi bir dizi alanda önemli etkileri olabilir. Ayrıca bu çalışma, duygusal analiz alanındaki gelecekteki araştırmalar için bir temel oluşturarak, teknolojinin insanlarla daha derin ve anlamlı bir şekilde etkileşime girmesini sağlayabilecek yeni yöntemlerin geliştirilmesine de katkıda bulunacaktır.

#### **ETİK STANDARTLARIN BEYANI** (DECLARATION OF ETHICAL STANDARDS)

Bu makalenin yazarları çalışmalarında kullandıkları materyal ve yöntemlerin etik kurul izni ve/veya yasal-özel bir izin gerektirmediğini beyan ederler.

The author of this article declares that the materials and methods they use in their work do not require ethical committee approval and/or legal-specific permission.

#### **YAZARLARIN KATKILARI** (AUTHORS' CONTRIBUTIONS)

Yazarlar makalenin fikir aşamasından yazım aşamasına kadar tüm aşamalarında birlikte çalışmıştır.

The authors worked together at all stages of the article, from idea to writing.

#### **ÇIKAR ÇATIŞMASI** (CONFLICT OF INTEREST)

Bu çalışmada herhangi bir çıkar çatışması yoktur.

There is no conflict of interest in this study.

#### **KAYNAKLAR** (REFERENCES)

- [1] Gunawan, T. S., Ashraf, A., Riza, B. S., Haryanto, E. V., Rosnelly, R., Kartiwi, M., & Janin, Z., Development of video-based emotion recognition using deep learning with Google Colab. TELKOMNIKA (Telecommunication Computing Electronics and Control), 18 (5), 2463-2471, 2020.
- [2] Ahmed, F., Bari, A. H., & Gavrilova, M. L., Emotion recognition from body movement. IEEE Access, 8, 11761-11781, 2019.
- [3] Chowdary, M. K., Nguyen, T. N., & Hemanth, D. J., Deep learning-based facial emotion recognition for human-computer interaction applications. Neural Computing and Applications, 35(32), 23311-23328, 2023.
- [4] Balti, A., Khelifa, M. M. B., Hassine, S. B., Ouazaa, H. A., Abid, S., Lakhoua, M. N., & Sayadi, M., Gait Analysis and Detection of Human Pose Diseases. In 2022 8th International Conference on Control, Decision and Information Technologies (CoDIT), 1 (IEEE), 1381-1386, 2022.
- [5] Park, S., Yong Chang, J., Jeong, H., Lee, J. H., & Park, J. Y., Accurate and efficient 3d human pose estimation algorithm using single depth images for pose analysis in golf. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 49-57, 2017.
- [6] Wang, J., Tan, S., Zhen, X., Xu, S., Zheng, F., He, Z., & Shao, L., Deep 3D human pose estimation: A review. Computer Vision and Image Understanding, 210, 103225, 2021.
- [7] Ota, M., Tateuchi, H., Hashiguchi, T., Kato, T., Ogino, Y., Yamagata, M., & Ichihashi, N., Verification of reliability and validity of motion analysis systems during bilateral squat using human

- pose tracking algorithm. *Gait & posture*, 80, 62-67. 2020.
- [8] Si, L., & Liu, B., Multifeature Fusion Human Pose Tracking Algorithm Based on Motion Image Analysis. *Wireless Communications and Mobile Computing*, 2022.
- [9] Verma, A., Meenpal, T., & Acharya, B., Human body pose distance image analysis for action recognition. *International Journal of Pattern Recognition and Artificial Intelligence*, 36 (07), 2255012. 2022.
- [10] Hyppönen, J., Hakala, A., Annala, K., Zhang, H., Peltola, J., Mervaala, E., & Kälviäinen, R., Automatic assessment of the myoclonus severity from videos recorded according to standardized Unified Myoclonus Rating Scale protocol and using human pose and body movement analysis. *Seizure*, 76, 72-78, 2020.
- [11] Xu, J., Yu, Z., Ni, B., Yang, J., Yang, X., & Zhang, W., Deep kinematics analysis for monocular 3d human pose estimation. In *Proceedings of the IEEE/CVF Conference on computer vision and Pattern recognition*, 899-908, 2020.
- [12] Rohan, A., Rabah, M., Hosny, T., & Kim, S. H., Human pose estimation-based real-time gait analysis using convolutional neural network. *IEEE Access*, 8, 191542-191550, 2020.
- [13] Paudel, P., Kwon, Y. J., Kim, D. H., & Choi, K. H., Industrial Ergonomics Risk Analysis Based on 3D-Human Pose Estimation. *Electronics*, 11 (20), 3403, 2022.
- [14] Stenum, J., Rossi, C., & Roemmich, R. T., Two-dimensional video-based analysis of human gait using pose estimation. *PLoS computational biology*, 17 (4), e1008935, 2021.
- [15] Kim, J. W., Choi, J. Y., Ha, E. J., & Choi, J. H., Human pose estimation using mediapipe pose and optimization method based on a humanoid model. *Applied Sciences*, 13 (4), 2700, 2023.
- [16] Wang, W., Enescu, V., & Sahli, H., Adaptive real-time emotion recognition from body movements. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 5 (4), 1-21, 2015.
- [17] Santhoshkumar, R., & Geetha, M. K., Deep learning approach: emotion recognition from human body movements. *Journal of Mechanics of Continua and Mathematical Sciences (JMCMs)*, 14(3), 182-195, 2019.
- [18] Samadiani, N., Huang, G., Cai, B., Luo, W., Chi, C. H., Xiang, Y., & He, J., A review on automatic facial expression recognition systems assisted by multimodal sensor data. *Sensors*, 19 (8), 1863, 2019.
- [19] Al Ghaithi, A., A robust approach for shear log predictions using deep learning on big data sets from a carbonate reservoir for integrated reservoir characterization projects. In *Second International Meeting for Applied Geoscience & Energy* (pp. 1840-1844). Society of Exploration Geophysicists and American Association of Petroleum Geologists, 2022.
- [20] Zhang, X., Hu, Y., Deng, J., Xu, H., & Wen, H., Feature engineering and artificial intelligence-supported approaches used for electric powertrain fault diagnosis: a review. *IEEE Access*, 10, 29069-29088, 2022.
- [21] Mo, Y., Wu, Y., Yang, X., Liu, F., & Liao, Y., Review the state-of-the-art technologies of semantic segmentation based on deep learning. *Neurocomputing*, 493, 626-646, 2022.
- [22] Šumak, B., Brdnik, S., & Pušnik, M., Sensors and artificial intelligence methods and algorithms for human-computer intelligent interaction: A systematic mapping study. *Sensors*, 22 (1), 20, 2021.
- [23] Yang, C. Y., Chen, P. C., & Huang, W. C., Cross-Domain Transfer of EEG to EEG or ECG Learning for CNN Classification Models. *Sensors*, 23 (5), 2458, 2023.
- [24] Tamil Priya, D., & Divya Udayan, J., Transfer learning techniques for emotion classification on visual features of images in the deep learning network. *International Journal of Speech Technology*, 23, 361-372, 2020.
- [25] Mukhopadhyay, A., Rajshekar Reddy, G. S., Mukherjee, I., Kumar Gopa, G., Pena-Rios, A., & Biswas, P., Generating Synthetic Data for Deep Learning using VR Digital Twin. In *Proceedings of the 2021 5th International Conference on Cloud and Big Data Computing*, August, 52-56, 2021.
- [26] Rudovic, O., Utsumi, Y., Lee, J., Hernandez, J., Ferrer, E. C., Schuller, B., & Picard, R. W., CultureNet: A deep learning approach for engagement intensity estimation from face images of children with autism. In *2018 IEEE/RSJ*

International Conference on Intelligent Robots and Systems (IROS), October, 339-346, 2018.

[27] Yao, R., Hou, L., Yang, L., Gui, J., & Wu, O., Deep human answer understanding for natural reverse QA. *Knowledge-Based Systems*, 254, 109625, 2022.

[28] Wang, T., Zhao, Y., & Yin, M., Analysis and research on the influence of music on students' mental health under the background of deep learning. *Frontiers in Psychology*, 13, 998451, 2022.

[29] YOLOv7. Pose vs MediaPipe in Human Pose Estimation. <https://learnopencv.com/yolov7-pose-vs-mediapipe-in-human-pose-estimation/>. Accessed: Feb. 12, 2024.