

Makine Öğrenmesi Kullanarak Futbol Yıldızlarının Yüz Tespiti ve Tanınması

Yiğit Görkem Özer, Ahmet Kurt, Seyit Ali Arslan

Bilgisayar Mühendisliği Bölümü

Ders: Denetimli Öğrenmenin Temelleri

Danışman: Dr. Öğr. Üyesi Resul Tugay

Özet—Bu çalışmada, futbol yıldızlarına ait yüz görüntülerinin otomatik olarak tespit edilmesi ve tanınması amacıyla makine öğrenmesi tabanlı bir yüz tanıma sistemi geliştirilmiştir. Yüz tanıma sistemleri, biyometrik kimlik doğrulama, güvenlik sistemleri, sosyal medya uygulamaları ve eğlence sektörü gibi birçok alanda yaygın olarak kullanılmaktadır. Geliştirilen sistem; yüz tespiti, özellik çıkarımı ve sınıflandırma olmak üzere üç ana aşamadan oluşmaktadır. Yüz tespiti için Haar Cascade yöntemi, özellik çıkarımı için yüz embedding teknikleri ve sınıflandırma için Destek Vektör Makineleri (SVM) kullanılmıştır.

Index Terms—Yüz Tanıma, Makine Öğrenmesi, SVM, Futbolcular, Görüntü İşleme

I. GİRİŞ

Yüz tanıma sistemleri, günümüzde biyometrik kimlik doğrulama, güvenlik sistemleri, sosyal medya uygulamaları ve eğlence sektörü gibi birçok alanda yaygın olarak kullanılmaktadır. Özellikle görüntü işleme ve makine öğrenmesi alanındaki gelişmeler sayesinde, insan yüzlerinin otomatik olarak tespit edilmesi ve kimliklendirilmesi mümkün hale gelmiştir.

Spor dünyasında ise ünlü futbolcuların yüzlerinin otomatik olarak tanınması; medya arşivleme, maç analizleri, istatistiksel çalışmalar ve eğlence amaçlı uygulamalar açısından önemli bir problem olarak karşımıza çıkmaktadır. Geleneksel yöntemlerde bu tür işlemler manuel olarak yapılmakta, bu durum zaman kaybına ve hatalara yol açabilmektedir.

Bu çalışmada, futbolculara ait yüz görüntülerinin otomatik olarak tespit edilmesi ve tanınması amacıyla makine öğrenmesi tabanlı bir yüz tanıma sistemi geliştirilmiştir. Projede, klasik derin öğrenme yöntemleri yerine, yüzlerden çıkarılan ayırt edici özelliklerin Destek Vektör Makineleri (Support Vector Machine – SVM) ile sınıflandırıldığı hibrit bir yaklaşım benimsenmiştir.

Geliştirilen sistem; yüz tespiti, özellik çıkarımı ve sınıflandırma olmak üzere üç ana aşamadan oluşmaktadır. Çalışmanın amacı, farklı futbolculara ait yüz görüntülerini yüksek doğruluk oranı ile taniyabilen, güven skoru üretebilen ve gerçek hayatı uygulanabilir bir yüz tanıma modeli ortaya koymaktır.

II. VERİ SETİ

Bu projede kullanılan veri seti, farklı futbol yıldızlarına ait yüz görüntülerinden oluşmaktadır. Veri seti görüntü tabanlı olup, her futbolcu ayrı bir sınıf (etiket) olarak ele alınmıştır. Görseller, farklı açılardan, ışık koşullarında ve yüz ifadelerinde çekilmiş görüntülerden oluşmaktadır.

Veri seti eğitim ve test olmak üzere iki ana gruba ayrılmıştır. Eğitim veri seti modelin öğrenme sürecinde kullanılırken, test veri seti modelin daha önce görmediği veriler üzerindeki başarısını ölçmek amacıyla kullanılmıştır. Her futbolcunun görüntülerini kendi ismini taşıyan klasörler içerisinde tutulmuş ve etiketleme işlemi otomatik olarak gerçekleştirilmiştir.

Veri setindeki tüm görüntüler RGB formatında olup, model eğitiminden önce çeşitli ön işleme adımlarından geçirilmiştir. Bu yapı, denetimli öğrenme yaklaşımı için uygun bir veri seti sunmaktadır.

III. YÖNTEM VE SÜREÇ

A. Veri Hazırlığı ve Ön İşleme

Ham görüntüler üzerinde doğrudan eğitim yapılması, arka plan gürültüsü ve yüz dışı detaylar nedeniyle model performansını olumsuz etkileyebilmektedir. Bu nedenle çalışmada öncelikle yüz tespiti işlemi gerçekleştirilmiştir.

Yüz tespiti için OpenCV kütüphanesinde yer alan Haar Cascade tabanlı yüz tespit algoritması kullanılmıştır. Yanlış pozitif yüz tespitlerini azaltmak amacıyla ek olarak göz tespiti yapılmış ve en az bir göz algılanan yüzler geçerli kabul edilmiştir.

Tespit edilen yüz bölgeleri, yüzün kenar kısımlarının kayboldurulması için yaklaşık %15 oranında genişletilerek kırpılmıştır. Elde edilen yüz görüntüleri, her futbolcu için ayrı klasörler halinde kaydedilmiştir. Bu işlem hem eğitim hem de test veri seti için ayrı ayrı uygulanmıştır.

B. Özellik Çıkarımı (Embedding)

Yüz tanıma işlemi için her yüz görüntüsünden ayırt edici özelliklerin çıkarılması gerekmektedir. Bu çalışmada, face-recognition kütüphanesi kullanılarak her yüz için sayısal özellik vektörleri (embedding) elde edilmiştir.

Yüz tespiti aşamasında Histogram of Oriented Gradients (HOG) tabanlı yöntem tercih edilmiştir. Tespit edilen yüzlerden 128 boyutlu embedding vektörleri çıkarılmıştır. Bu vektörler, yüzlerin karakteristik özelliklerini temsil eden sayısal veriler olup, sınıflandırma aşamasında girdi olarak kullanılmıştır.

C. Model Mimarisi (SVM)

Elde edilen yüz embedding'lerinin sınıflandırılması amacıyla Support Vector Machine (SVM) algoritması kullanılmıştır. SVM, özellikle yüksek boyutlu veri uzaylarında başarılı

sonuçlar vermesi ve küçük–orta ölçekli veri setleri için uygun olması nedeniyle tercih edilmiştir.

Modelde lineer çekirdek (linear kernel) kullanılmıştır. Ayrıca, sınıflandırma sonuçlarının güven skorlarının hesaplanması için probability parametresi aktif hale getirilmiştir. Bu sayede model, yaptığı tahminlerin ne kadar güvenilir olduğunu da ifade edebilmektedir.

IV. BULGULAR VE ANALİZ

Eğitilen model, test veri seti üzerinde değerlendirilmiştir. Modelin performansı doğruluk (accuracy), precision, recall ve F1-score gibi metrikler kullanılarak ölçülmüştür.

Elde edilen sonuçlar, modelin futbolculara ait yüzleri yüksek doğruluk oranı ile tanıyalabildiğini göstermektedir. Sınıflandırma raporu incelendiğinde, her bir futbolcu sınıfı için dengeli ve tutarlı başarı oranları elde edilmiştir.

Ayrıca model, tek bir fotoğraf üzerinden yüz tespiti ve tanıma işlemini başarıyla gerçekleştirmiştir. Güven skoru belirli bir eşik değerinin altında olan tahminler “Unknown” olarak etiketlenmiş ve bu sayede hatalı tanımların önüne geçilmiştir.

V. DEĞERLENDİRME VE TARTIŞMA

Modelin genel başarısı yüksek olmakla birlikte, bazı durumlarda yanlış sınıflandırmalar gözlemlenmiştir. Bu hataların temel nedenleri arasında düşük çözünürlükli görüntüler, aşırı ışık yansımaları, yüzün yan profilden görünmesi ve yüz ifadelerindeki büyük değişiklikler yer almaktadır.

Güven skoru mekanizmasının kullanılması, modelin emin olmadığı durumlarda yanlış tahmin yapmasını engelleyerek sistemin güvenilirliğini artırmıştır.

VI. SONUÇ

Bu projede, futbol yıldızlarına ait yüz görüntülerinin otomatik olarak tespit edilmesi ve tanınması amacıyla makine öğrenmesi tabanlı bir yüz tanıma sistemi geliştirilmiştir. Haar Cascade tabanlı yüz tespiti, yüz embedding çıkarımı ve SVM sınıflandırıcısı kullanılarak etkili bir sistem oluşturulmuştur.

Elde edilen sonuçlar, geliştirilen sistemin daha önce görümediği görüntüler üzerinde de başarılı tahminler yapabileğini göstermektedir. Modelin hem doğruluk oranı hem de güven skoru üretme yeteneği, sistemin pratikte kullanılabilir olduğunu ortaya koymaktadır.

VII. KAYNAKÇA

KAYNAKLAR

- [1] OpenCV Library
- [2] face-recognition Python Library
- [3] Scikit-learn Documentation
- [4] Kaggle – Football Players Dataset
- [5] Transfermarkt

Abstract—In this study, a machine learning-based face recognition system is developed for the automatic detection and recognition of football players' faces. Face recognition systems are widely used in many fields such as biometric authentication, security systems, social media applications, and the entertainment industry. The proposed system consists of three main stages: face detection, feature extraction, and classification. Haar Cascade is employed for face detection, facial embedding techniques are used for feature extraction, and Support Vector Machines (SVM) are applied for classification.

Index Terms—Face Recognition, Machine Learning, SVM, Football Players, Image Processing

VIII. INTRODUCTION

Face recognition systems are widely used in many fields such as biometric authentication, security systems, social media applications, and the entertainment industry. With the advances in image processing and machine learning, automatic detection and identification of human faces have become feasible.

In the sports domain, automatic recognition of famous football players' faces plays an important role in media archiving, match analysis, statistical studies, and entertainment-oriented applications. Traditional approaches rely on manual processes, which are time-consuming and prone to errors.

In this study, a machine learning-based face recognition system is developed for the automatic detection and recognition of football players' faces. Instead of classical deep learning methods, a hybrid approach is adopted in which discriminative facial features are classified using Support Vector Machines (SVM).

The proposed system consists of three main stages: face detection, feature extraction, and classification. The main objective of this study is to develop a practical face recognition model that can recognize different football players with high accuracy, produce confidence scores, and be applicable in real-world scenarios.

IX. DATASET

The dataset used in this study consists of face images belonging to different football stars. The dataset is image-based, and each football player is considered as a separate class (label). The images include various viewing angles, lighting conditions, and facial expressions.

The dataset is divided into two main subsets: training and testing. While the training dataset is used during the learning phase of the model, the test dataset is used to evaluate the performance of the model on previously unseen data. Images of each football player are stored in folders named after the player, enabling automatic labeling.

All images are in RGB format and are subjected to several preprocessing steps before model training. This structure provides a suitable dataset for supervised learning approaches.

X. METHODOLOGY

A. Data Preparation and Preprocessing

Training a model directly on raw images may negatively affect performance due to background noise and non-facial

details. Therefore, face detection is performed as an initial preprocessing step.

Haar Cascade-based face detection provided by the OpenCV library is used. To reduce false positive detections, eye detection is additionally applied, and only faces with at least one detected eye are considered valid.

Detected face regions are cropped with approximately 15% padding to prevent the loss of facial boundaries. The cropped face images are stored in separate folders for each football player. This process is applied to both training and test datasets.

B. Feature Extraction (Embedding)

For face recognition, discriminative features must be extracted from each face image. In this study, numerical feature vectors (embeddings) are generated for each face using the face-recognition library.

During face detection, the Histogram of Oriented Gradients (HOG) based method is preferred. From each detected face, a 128-dimensional embedding vector is extracted. These vectors numerically represent the distinctive characteristics of faces and are used as input for the classification stage.

C. Model Architecture (SVM)

Support Vector Machines (SVM) are employed to classify the extracted facial embeddings. SVM is chosen due to its effectiveness in high-dimensional feature spaces and its suitability for small to medium-sized datasets.

A linear kernel is used in the model. Additionally, probability estimation is enabled to compute confidence scores for classification results. This allows the model to indicate how reliable each prediction is.

XI. RESULTS AND ANALYSIS

The trained model is evaluated on the test dataset. Performance is measured using metrics such as accuracy, precision, recall, and F1-score.

The results indicate that the model can recognize football players' faces with high accuracy. The classification report shows balanced and consistent performance across different football player classes.

Furthermore, the model successfully performs face detection and recognition from a single image. Predictions with confidence scores below a predefined threshold are labeled as "Unknown", thereby preventing incorrect identifications.

XII. EVALUATION AND DISCUSSION

Although the overall performance of the model is high, some misclassifications are observed. The main reasons include low-resolution images, excessive lighting variations, side-profile faces, and significant changes in facial expressions.

The use of a confidence score mechanism improves system reliability by preventing uncertain predictions from being incorrectly classified.

XIII. CONCLUSION

In this study, a machine learning-based face recognition system is developed for the automatic detection and recognition of football players' faces. An effective system is constructed using Haar Cascade-based face detection, facial embedding extraction, and SVM classification.

The results demonstrate that the proposed system can successfully recognize previously unseen images. The ability to achieve high accuracy and generate confidence scores indicates that the system is practical and applicable in real-world scenarios.

XIV. REFERENCES

REFERENCES

- [1] OpenCV Library
- [2] face-recognition Python Library
- [3] Scikit-learn Documentation
- [4] Kaggle Football Players Dataset
- [5] Transfermarkt