

Üniversite	: T.C. İstanbul Kültür Üniversitesi
Enstitü	: Lisansüstü Eğitim Enstitüsü
Anabilim Dalı	: Bilgisayar Mühendisliği
Program	: Bilgisayar Mühendisliği
Tez Danışmanı	: Doç. Dr. Fatma PATLAR AKBULUT
Tez Türü ve Tarihi	: Yüksek Lisans – Haziran 2023

ÖZET

AÇIKLANABİLİR VE YORUMLANABİLİR YÜZ DUYGU TANIMA

Günümüzde birçok sektörde otonom sistemler artmaktadır. Bu sistemler ile insanlar arasındaki etkileşimin artmasıyla bu sistemlerin doğru kararlar vermesi oldukça zorlaşmıştır. Otonom sistemlerin karar alırken insanların duygularını dikkate almaması, yanlış kararlar almasına sebep olmaktadır. Oluşturulan sistemlerin kişinin duygu durumuna göre karar vermesi gereken durumlarda, bu sistemlerin kişilerin duygularını yüz ifadelerinden hızlı ve doğru bir şekilde tahmin etmelerini gerekli kılmıştır. Bu çalışma yüz ifadesi tanımak için yapılan çalışmalarda oluşturulan yapay zekâ modellerinin aldığı kararların açıklanması ve eksikliklerinin belirlenmesi konularında yol gösterici olacaktır. İlk olarak Literatürde bu alanda yer alan çalışmalar incelenerek bu çalışmada kullanılacak model, veriseti ve teknikler belirlenmiştir. Bu çalışmada daha önceden eğitilmiş modeller olan VGG, ResNet ve Inception modelleri üzerinde iyileştirmeler yapılarak yeni üç model oluşturulmuştur. Oluşturulan modeller Fer2013 veriseti ile eğitilmiş ve modeller ilk olarak eğitim ve test verileri üzerinden karşılaştırılmıştır. Daha sonra ise modeller yedi farklı duygu durumu için seçilen yedi görüntüyü tahmin etmeleri sağlanarak ikinci defa karşılaştırılmıştır. Her iki karşılaştırmada da en iyi model VGG olmuştur. Modellerin yaptıkları tahminlerin doğruysa neden doğru ve yanlışsa neden yanlış olduğunu açıklanması için beş farklı Açıklanabilir Yapay Zekâ (XAI) tekniği kullanılmıştır. Kullanılan XAI teknikleri GradCAM, GradCAM++, Saliency haritası, SHAP ve LIME teknikleridir. Her modelin yedi farklı duygu durumunu tahmin etmesi sağlandıktan sonra bu tahminin beş farklı teknikle açıklaması yapılmıştır. En açıklayıcı bilgiler GradCAM++ ve Saliency haritası teknikleriyle elde edilmiştir.

Anahtar Kelimeler: GradCAM, Saliency Haritası, SHAP, LIME, Açıklanabilir Yapay Zekâ

University : T.C. İstanbul Kültür University
Institute : Institute of Graduate Studies
Department : Computer Engineering
Program : Computer Engineering
Thesis Advisor : Assoc. Prof. Dr. Fatma PATLAR AKBULUT
Degree Awarded And Date : MA – June 2023

ABSTRACT

Explainable and Interpretable Facial Emotion Recognition

In today's world, autonomous systems are increasingly prevalent in various industries. As the interaction between these systems and humans increases, it becomes more challenging for these systems to make accurate decisions. One of the reasons for the systems making incorrect decisions is that they do not take into account human emotions when making decisions. In situations where these systems need to make decisions based on a person's emotional state, it has become necessary for these systems to quickly and accurately predict people's emotions from their facial expressions. This study aims to provide guidance on explaining the decisions made by artificial intelligence models developed for facial expression recognition and identifying their shortcomings. Firstly, existing studies in this field were examined to determine the model, dataset, and techniques to be used in this study. In this study, three new models were created by improving the pre-trained models VGG, ResNet, and Inception. These models were trained using the Fer2013 dataset, and initially compared using training and test data. Subsequently, the models were compared again by having them predict seven images selected for seven different emotional states. In both comparisons, the VGG model performed the best. To explain why the models made correct or incorrect predictions, five different eXplainable Artificial Intelligence (XAI) techniques were used. The XAI techniques employed were GradCAM, GradCAM++, Saliency Map, SHAP, and LIME. After each model predicted the seven different emotional states, the predictions were explained using the five techniques. The most informative results were obtained using GradCAM++ and Saliency Map techniques.

Keywords: GradCAM, Saliency Map, SHAP, LIME, Explainable Artificial Intelligence