

## GENEL BİLGİLER

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

## KISA ÖZET

### YÜZ İFADESİ TANIMADA ÖNYARGISIZ VE AYRIŞTIRILMIŞ TEMSİLLERİN ÖĞRENİLMESİ

Sözsüz iletişimin bir parçası olan yüz ifadelerinin tanınmasına yönelik geliştirilen sistemler incelendiğinde yüzde bulunan kişisel özelliklerin ve aksesuarların duygu tanıma sistemlerini olumsuz etkilediği tespit edilmektedir. Bu problemin ortadan kaldırılmasına yönelik literatürde kullanılan yöntemler incelendiğinde ayrıştırılmış temsil öğrenimi öne çıkmaktadır. Hassas değişkenlerin, hedef temsilinden ayrıştırılması değişimsel oto kodlayıcılar ile mümkün olmaktadır. Çalışmada bir temel sınıflandırıcı, bir beta-vae kullanılarak eğitilen bir sınıflandırıcı, son olarak da temsillerin doğru bir şekilde ayrıştığını kontrol edebilmek için ortogonalite kaybı kullanan bir mimari ile yüz ifadesi sınıflandırma modelleri eğitilmiştir. Hassas olarak seçilen özellikler yüz ve boyun bölgesinde bulunan kişisel özellikler ve aksesuarlar olarak belirlenmiştir. Kullanılan veri setine göre seçilen duygu gülümseme üzerinden değerlendirilmektedir. Elde edilen model sonuçları temel sınıflandırıcı modeli ile karşılaştırılmıştır. Belirlenen hassas özellikler için ayrıştırılmış temsil yönteminin doğruluğu olumlu yönde etkilediği gözlemlenmiştir. Bunun yanı sıra hassas özelliklerin hedef temsilden çıkarılması ile daha adil bir sınıflandırıcı elde edilmiştir.

**Anahtar Kelimeler:** Ayrıştırılmış Temsil Öğrenimi, B-VAE, Duygu Tanıma Sistemleri, Duygu Tanıma Sistemlerinde Önyargı

## GENERAL INFORMATION

University : İstanbul Kültür University

Institute : Institute of Graduate Studies

Department : Computer Engineering

Programme : Computer Engineering

Supervisor : Assoc. Prof. Dr. Fatma PATLAR AKBULUT

Degree Awarded and Date : Master's degree – June 2024

## ABSTRACT

### LEARNING DEBIAS AND DISENTANGLED REPRESENTATION IN FACIAL EXPRESSION RECOGNITION

When the systems developed for the recognition of facial expressions, which are a part of non-verbal communication, are examined, it is determined that characteristic features and accessories on the face negatively affect emotion recognition systems. When the methods used in the literature to eliminate this problem are examined, disentangled representation learning stands out. Separating sensitive variables from the target representation is possible with variational autoencoders. In the study, facial expression classification models were trained with a base classifier, a classifier trained using beta-VAE, and finally an architecture that uses orthogonality loss to check that the representations are correctly disentangled. The features selected as sensitive features were determined as characteristic features and accessories located in the face and neck area. Depending on the dataset used, the selected emotion is evaluated through smile. The obtained model results were compared with the base classifier model. It has been observed that the disentangled representation method has a positive impact on accuracy for the identified sensitive features. In addition, a fairer classifier was obtained by removing sensitive features from the target representation.

**Keywords:** Disentangled Representation Learning, B-VAE, Facial Expression Recognition, Bias in Facial Expression Recognition System