Application Vision Transformers On Face Age Regression

Chami, Ahmed Chaouki*; Ajgou, Riadh

El-Oued University

Abstract

Transformers have recently gained signifcant attention in machine learning due to their self-attention mechanisms, which allow models to dynamically assess the importance of different input elements. Although originally designed for Natural Language Processing (NLP), the application of transformers in computer vision tasks, such as image classifcation, has been gaining traction. This work explores the use of Vision Transformers (ViT) in the context of face age regression, focusing on three well-known datasets: MORPH II, AFAD, and CACD. By leveraging ViT in a regression setting, we aim to predict the age of individuals based on facial images. We evaluate the model's performance using the Mean Absolute Error (MAE) on each of these datasets and compare it to traditional models like Convolutional Neural Networks (CNNs). Furthermore, we investigate the computational efciency and performance gains from transfer learning using pre-trained ViT models on the ImageNet dataset. Our experiments demonstrate that Vision Transformers offer a competitive alternative to CNNs for face age regression, with promising results across all three datasets, showing their potential for future applications in age estimation and facial analysis.

Corresponding author email : chaouki1100@gmail.com