High-Accurate Age Estimation for Minors

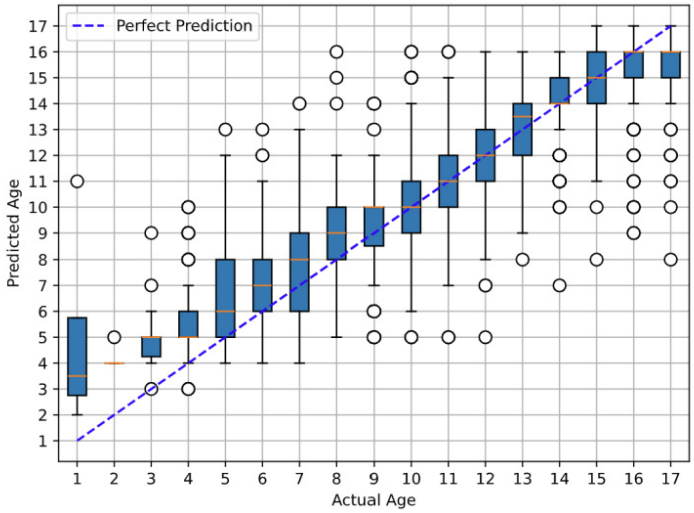

Figure 1: Actual age vs predicted age of the proposed VOLO-D1.The blue box represents the middle 50% of the predicted age distribution for a given age class. The extended lines above and below the box indicate the ranges for each 25% of the distribution (Whisker). 75% of the predicted ages fall below the upper line of the box (Third Quartile - Q3), while 25% fall below the lower line (First Quartile - Q1). Small circles represent outliers in the distribution. The orange line marks the median of the distribution.

Project Manager
Eneldo Loza Mencia

Researchers
Duc Huy Nguyen and Santiago del Rio

Principal Investigator
Dr. Oren Halvani

Project Term
2025 - 2025

Software
PyTorch

Institute
Fraunhofer Institute for Secure Information Technology SIT

University
Technische Universität Darmstadt

Introduction

Age estimation for underage individuals is essential for detecting Child Sexual Abuse Material (CSAM). Detecting CSAM content generally involves identifying whether this content is pornographic, and detecting if it includes minors, which are under 18 years old according to many countries. Furthermore, it is crucial to accurately determine whether the depicted minors are children or adolescents, which are generally separated by the legal age of sexual consent. This task is crucial for legal cases, as a child's offense is generally considered more severe than an adolescent's offense Recently, many existing studies have focused on the age estimation of individuals between underage and adult, separated by the age of 18. There is a lack of research focusing only on the underage group, namely classifying children and adolescents separated by the legal age of sexual consent. This project aimed to fill this gap by introducing an effective age estimation model that can estimate a person's age based on their face, and accurately classify whether this person is a child or an adolescent. The selected age threshold was 14, which is the legal age of sexual consent in Germany. This age estimation model could be integrated with a minor-adult classifier and a nudity detector to form a system capable of detecting CSAM content and evaluating the severity of the cases. Due to the large scale of the dataset and the complexity of architectures such as VOLO-D1 and ResNet-50, the use of high-performance computing (HPC) was essential. HPC enabled efficient model training, hyperparameter optimization, and evaluation across multiple design choices within a feasible time frame.

Methods

The dataset used for this project was Juvenile-80k, the most extensive facial dataset for underage individuals with apparent age and gender labels. This project examined several design choices, including whether to pretrain the model, whether to formulate the task as classification or regression, which classification approach to use and which architecture to use among ResNet-50, VOLO-D1, and MLP. ResNet-50 has been widely used in many age estimation research studies. VOLO-D1 has demonstrated a good performance on the IMDB-Clean dataset. Finally, MLP is a lightweight architecture, but has achieved some promising results on the juvenile age estimation task. Additionally, it evaluates the model on different ages of consent such as 15 and 16, and applies threshold tuning to improve classification.

Results

The results demonstrated that pretraining was effective, with IMDB-Clean pretraining outperforming ImageNet pretraining and no pretraining. Classification outperformed regression, and the median was the most effective classification approach. VOLO-D1 was the most effective architecture for juvenile age estimation. It achieved a mean absolute error (MAE) of 1.27 in age estimation, a balanced accuracy of 0.90 in age group

classification, outperforming other existing baselines such as DEX and MiVOLO-D1. Overall, this thesis can serve as a valuable tool for CSAM investigation.

Discussion

However, some limitations suggest areas for future improvement. Due to the imbalanced dataset, our model predicted certain ages less accurately than others. For example, 1 year olds were commonly predicted as 5 year olds. It did not affect the overall classification performance, but could be addressed with a balanced dataset. Our model also did not generalize well to other datasets, a problem many age estimation models encountered. Therefore, future work should prioritize creating a balanced, representative dataset for underage individuals. Age estimation models would generalize better if the training data is balanced in every aspect such as age, gender, and ethnicity. Furthermore, most underage datasets contain only front-facing images, which do not reflect the reality of CSAM content, often composed of in-the-wild images. Diverse datasets containing images from both scenarios and from various angles would benefit researchers in this field.

Last Update: 2026-07-01 10:34