Age Prediction Using Facial and Nose Images with CNN Models

Bram Ekelschot
Master Applied Artificial Intelligence
Amsterdam University of Applied Sciences
Amsterdam, The Netherlands

***Abstract*—This study investigates the effectiveness of using the nose region alone for age classification in convolutional neural networks (CNNs). By comparing CNN models trained on full-face images with those using only nose-region images from the UTKFace dataset, we assess the predictive potential of this feature-specific approach. While full-face models achieved slightly higher accuracy, nose-only models demonstrated promising results, suggesting applications in settings with limited visibility or resources. The findings highlight a pathway toward efficient, adaptable age classification models and underscore the importance of addressing ethnic morphological variations in future research.**

***Index Terms*—Age Prediction, Convolutional Neural Networks (CNN), Feature-Specific Learning, Nose Region, UTKFace Dataset, Facial Recognition, Age Classification, Mean absolute error (MAE)**

I. INTRODUCTION

A. Research Topic

The focus of this research is **feature-specific learning** within convolutional neural networks (CNNs), particularly applied to age classification. CNNs have made significant advancements in facial recognition; however, most models typically analyze full-face images. This approach risks overlooking the predictive potential of individual facial features. This study narrows its focus to the nose, which remains relatively stable compared to other features, such as the eyes or mouth. The nose undergoes distinct age-related changes, which may offer valuable information for accurate age predictions [3], [4].

B. Research Problem

The reliance on full-face images in current facial recognition and age classification models presents practical challenges. In real-world scenarios, such as surveillance or security applications, parts of the face are frequently occluded by masks, hair, or accessories [9]. This full-face dependence can lead to increased computational requirements and may neglect age-related cues that are more prominent in specific facial regions, like the nose. Consequently, a gap exists in understanding whether focusing solely on a feature like the nose can provide reliable age classifications [8]. This study seeks to address this gap by investigating the nose's potential as a predictive feature in age estimation tasks.

C. Research Purpose

The aim of this research is to evaluate the effectiveness of using only the nose region for age classification in CNN models. By comparing the performance of a CNN trained on full-face images with one trained exclusively on nose-cropped images, this study will assess whether the nose can provide sufficient information for accurate age predictions. Filling this research gap will not only enhance understanding of feature-specific learning but also have practical implications for facial recognition systems in environments where full facial visibility is limited [3], [9], [8].

II. RESEARCH QUESTION

The primary research question guiding this study is: ****How does zooming in on the nose region in the UTKFace dataset affect the accuracy of CNN models for age classification?**** This question is pivotal as it addresses the limitations of traditional age classification methods, which predominantly rely on analyzing full-face images.

Focusing on the nose region presents a unique opportunity to explore whether a localized feature can offer reliable age prediction. The nose is a central facial feature that remains relatively stable over time and is less likely to be occluded by external factors, such as hair or accessories. Moreover, age-related changes in the nose can provide valuable cues for classification tasks [3], [9].

By investigating this question, the study aims to fill a significant gap in the existing literature on feature-specific learning, which has largely overlooked the potential of individual facial features for age estimation. Prior research indicates that concentrating on specific areas of the face can yield comparable results to full-face models while potentially reducing computational overhead [4], [8].

Hypothesis: *"This study hypothesizes that the nose-focused model will reduce computational complexity while maintaining accuracy close to that of a full-face model."*

Therefore, this research question is not only relevant but also necessary to advance the understanding of how specific facial features contribute to the performance of CNN models in age classification tasks.

III. METHODS

A. Data Collection and Preparation

1) Dataset: For this study, the UTKFace dataset was chosen due to its diverse age-labeled facial images, which have been widely used in age estimation research and are readily available for academic purposes. The UTKFace dataset includes images spanning various ages, genders, and ethnicities, making it suitable for training CNN models that aim to generalize across demographics. Each image in the dataset is labeled with the age of the person, allowing for supervised learning approaches to age classification [2].

The choice of UTKFace is advantageous because it includes images with age labels from 0 to 116 years, covering a wide range of age-related features that are essential for training accurate age classification models. Additionally, the dataset's balanced representation of age groups makes it possible to mitigate bias toward specific age ranges, which is crucial for developing robust models in age prediction tasks.

B. Data Preprocessing

In this study, a convolutional neural network (CNN) was designed for age prediction using facial images, specifically focusing on the nose region. The architecture of the model was constructed to efficiently extract features from the input images while minimizing overfitting and maximizing predictive accuracy.

Prior to training, the images were processed as follows:

- **Image Loading**: OpenCV was used to read images from a specified directory and convert them to grayscale to facilitate nose detection with a pre-trained Haar Cascade classifier ('haarcascade_mcs_nose.xml').
- **Nose Detection**: The Haar Cascade classifier was employed to detect nose regions, ensuring that only images with detectable noses were included in the dataset. The error handling here is that if no nose is found the data is not added to the dataset to ensure both models have the exact same dataset with the only difference that one has cropped images.
- **Image Resizing**: Each detected nose image was resized to 32x32 pixels to standardize the input size for the CNN, balancing computational efficiency and feature representation.
- **Normalization**: Pixel values were scaled to a range between 0 and 1 by dividing by 255.0, accelerating the training process and improving model convergence. This step is critical, especially when comparing different regions like the nose versus the entire face, as it ensures consistency in feature scaling across models. Normalization reduces the influence of varying pixel intensity ranges that could affect model training and improve convergence. Since face and nose regions differ in size and level of detail, normalizing these regions to the same scale helps in achieving comparability between models trained on different input types.

The resulting dataset consisted of processed nose images and their corresponding age labels, which were split into training and testing sets using an 80-20 ratio.

C. CNN Architecture

The CNN model was structured as follows for a simple base CNN [1]:

- **Input Layer**: Accepts images of size 32x32 pixels with three color channels (RGB).
- **Convolutional Layers**:
  - The first convolutional layer consists of 32 filters of size 5x5, applying the ReLU activation function. This layer extracts low-level features from the input images.
  - The second convolutional layer includes 64 filters of size 5x5, also using the ReLU activation function, capturing more complex features.
- **Pooling Layers**: After each convolutional layer, a max pooling layer with a pool size of 2x2 is applied to reduce the spatial dimensions of the feature maps, retaining the most significant features while reducing computational load.
- **Flatten Layer**: Converts the 2D feature maps into a 1D feature vector, preparing the data for the fully connected layers.
- **Fully Connected Layers**:
  - A dense layer with 128 neurons and ReLU activation follows the flattening step, allowing for non-linear combinations of features.
  - A dropout layer with a rate of 0.5 is included to reduce overfitting by randomly setting half of the neurons to zero during training.
  - The output layer consists of a single neuron with no activation function, as the model is set up for age regression.

D. Model Compilation and Training

The model was compiled using the Adam optimizer and Mean Squared Error (MSE) loss function, which is suitable for regression tasks. The performance of the model was measured using the Mean Absolute Error (MAE) metric.

Training consisted of 200 epochs with a batch size of 32, using the training dataset for fitting and the validation dataset for monitoring the model's performance.

E. Limitations

One limitation encountered was the variability in the predictive power of the nose region alone. While the nose offers stable features, its age-related cues may not always be sufficient for accurate age classification. In cases where age prediction relies heavily on changes in other facial features, the nose-only model may underperform compared to the full-face model. This limitation suggests that the nose-only approach may be best suited for constrained scenarios, and integrating other facial regions could enhance model accuracy.

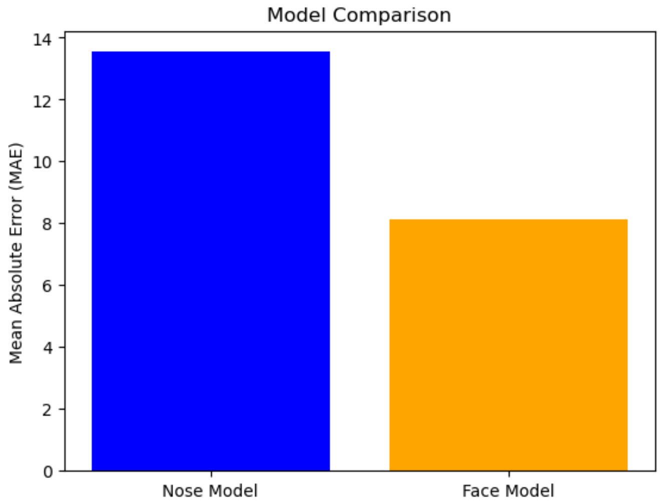

Fig. 1. MAE comparison between the Full-Face Model and Nose-Only Model

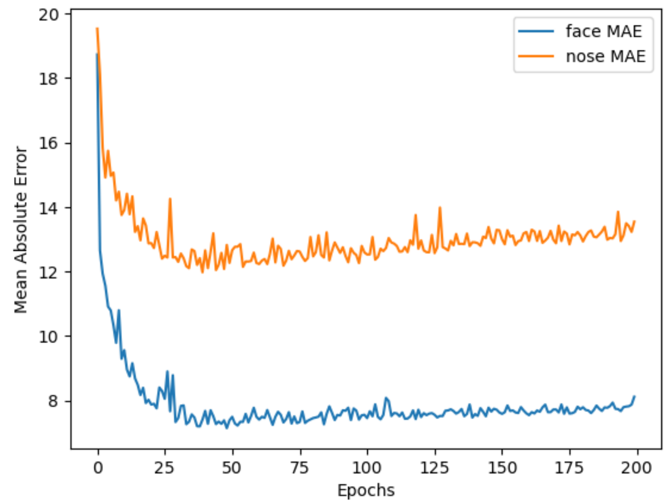

Fig. 2. MAE progression over epochs for both models

IV. Evaluation and Results

The primary metric used to evaluate the performance of the CNN models was the Mean Absolute Error (MAE), as it provides an intuitive measure of the average error in age prediction. Lower MAE values indicate better performance, as the model's predictions are closer to the true age values [5].

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

Two CNN models were compared:

- **Full-Face Model**: Trained on full facial images.
- **Nose-Only Model**: Trained solely on images cropped to the nose region.

Figure 1 shows the MAE comparison between the two models. The full-face model achieved a lower MAE than the nose-only model, indicating that analyzing the entire face provided more accurate age predictions. This suggests that features from other facial regions, in addition to the nose, contribute valuable information for age estimation.

However, the nose-only model demonstrated promising results, highlighting the predictive potential of the nose region alone. While it did not achieve the same accuracy as the full-face model, the nose-only model offers a practical alternative for scenarios where full-face visibility may be compromised, such as in masked or partially occluded images.

Figure 2 illustrates the MAE progression over epochs for both models. As observed, both models improved with each epoch, with the full-face model converging to a lower MAE. These results underscore the benefit of including more facial features while also supporting the feasibility of feature-specific learning for age estimation tasks.

V. Discussion

The findings from this study hold practical implications for advancing age classification in facial recognition technology, particularly by exploring feature-specific learning within convolutional neural networks (CNNs). The comparative analysis between models trained on full-face images and those trained solely on nose-region images revealed that, while the full-face model achieved a lower Mean Absolute Error (MAE), the nose-only model also performed notably well, underscoring the potential of localized feature learning for age estimation tasks.

A significant implication of this research is the feasibility of developing lightweight models that focus on specific facial features, such as the nose, to achieve reliable age predictions. Such models could be computationally more efficient and better suited for deployment in resource-constrained environments, where power and processing capabilities are limited [6]. For instance, in crowded spaces, surveillance applications, or settings with low-light conditions, where full-face visibility is often obstructed, using only the nose as a predictive feature could offer a practical alternative.

Furthermore, the morphology of the nose varies significantly across different ethnic groups due to both genetic and environmental factors. These variations could influence the accuracy of age prediction models if they are not properly accounted for. Prior research has shown that the structure of the nose adapts to diverse climatic conditions, with wider nostrils more commonly found in populations from warmer climates and narrower nostrils in populations from colder climates [7]. Such physiological differences suggest that age classification models based on the nose alone may need further fine-tuning to ensure equitable performance across ethnicities and avoid unintended biases. Future research could expand on this study by examining whether these morphological differences in the nose impact the age prediction accuracy across diverse demographics.

The insights from this research also open avenues for further exploration into other facial features as standalone predictors for age classification. While this study focused on the nose, other stable features, such as the eyes or mouth, may contain age-related cues that could contribute to reliable predictions in isolation. Exploring these localized features could refine the scope of feature-specific learning, potentially enabling modular and efficient age classification systems that are adaptable to varying real-world conditions.

In conclusion, this study contributes to the growing body of knowledge on feature-specific learning in CNNs by demonstrating the practical utility of the nose as a predictive feature for age classification. This approach has the potential to reduce computational requirements, improve accessibility in resource-limited environments, and encourage a more inclusive approach in facial recognition systems by accounting for ethnic variations. These findings pave the way for future research into the integration of additional facial features, ultimately enhancing the adaptability and fairness of age classification models.

VI. Conclusion

This study evaluated the effectiveness of using only the nose region versus full facial images for age classification in convolutional neural networks (CNNs). By comparing the performance of two CNN models trained on full-face images and nose-only images, we found that while the full-face model achieved higher accuracy, the nose-only model also provided meaningful results. This supports the potential of feature-specific learning for age prediction, especially in scenarios where full-face visibility may be compromised, such as in surveillance or security settings.

The findings underscore the practicality of developing lightweight, feature-specific models that can operate efficiently in resource-constrained environments. Furthermore, the varying morphology of the nose across ethnic groups points to the need for further research to ensure that feature-specific age prediction models perform equitably across diverse populations. Future work could investigate additional facial features as standalone inputs, further enhancing the adaptability of CNNs in age estimation tasks.

In conclusion, this research contributes to advancing facial recognition technology by highlighting the utility of individual facial features in age classification. With continued exploration, feature-specific learning could lead to more efficient, accessible, and inclusive applications of age prediction models in real-world scenarios.

References

[1] Abien Fred Agarap. An architecture combining convolutional neural network (cnn) and support vector machine (svm) for image classification. *arXiv preprint arXiv:1712.03541*, 2017.

[2] Moritz Mistol. Age estimation on the utkface dataset. https://wandb.ai/moritzm00/UTKFace-Age-Regression/reports/Age-Estimation-on-the-UTKFace-Dataset--Vmlldzo0MzE3MjE3, 2021.

[3] D. V. Pakulich, S. A. Yakimov, and S. A. Alyamkin. Age recognition from facial images using convolutional neural networks. *Optoelectronics, Instrumentation and Data Processing*, 55(3):255–262, May 2019.

[4] V. Sheoran, S. Joshi, and T. R. Bhayani. Age and gender prediction using deep cnns and transfer learning. arXiv preprint arXiv:2110.12633, 2021.

[5] Abhishek V Tatachar. Comparative assessment of regression models based on model evaluation metrics. *International Research Journal of Engineering and Technology*, 2021.

[6] Sheorey T. Ojha A Thakur, P.S. Vgg-icnn: A lightweight cnn model for crop disease identification. *Multimed Tools Appl*, 2023.

[7] Y. F. Wong and et al. Inter-ethnic/racial facial variations: A systematic review and bayesian meta-analysis of photogrammetric studies. *PLOS ONE*, 14(8):e0221556, 2019.

[8] K. Zhang, N. Liu, and X. Yuan. Fine-grained age estimation in the wild with attention lstm networks. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 1–9, 2018.

[9] G. Özbulak, Y. Aytar, and H. K. Ekenel. How transferable are cnn-based features for age and gender classification? In *Proc. International Conference of the Biometrics Special Interest Group (BIOSIG)*, pages 1–6, 2016.