

Impact of Image Resolution on Age Estimation with DeepFace and InsightFace

Shiyar Jamo

Master Applied Artificial Intelligence
Amsterdam University of Applied Sciences
Amsterdam, Netherlands
shiyar.jamo@hva.nl

Abstract—Automatic age estimation is widely used for age verification, where input images often vary considerably in resolution. This study evaluates the effect of image resolution on age estimation accuracy using DeepFace and InsightFace. A total of 1000 images from the IMDB-Clean dataset were processed in seven resolutions, resulting in 7000 test samples. Performance was evaluated using Mean Absolute Error (MAE), Standard Deviation (SD), and Median Absolute Error (MedAE). This study shows that image resolution has a clear impact on age estimation accuracy. Both frameworks achieve optimal performance at 224x224 pixels, with an MAE of 10.83 years (DeepFace) and 7.46 years (InsightFace). At low resolutions, MAE increases substantially, while very high resolutions also degrade accuracy. InsightFace is consistently faster than DeepFace across all resolutions.

Index Terms—DeepFace, InsightFace, age estimation, image resolution, resolution, CNN, VGG-Face, ArcFace, age prediction, IMDB-Clean

I. INTRODUCTION

Automatic age estimation using AI is valuable in many real-world scenarios. Examples include self-checkout systems and vending machines that restrict the sale of alcohol or tobacco, online age verification for gaming and social media, and access control at age-restricted events. Pretrained frameworks such as DeepFace [1] and InsightFace [2] can estimate a person's age from facial images. However, these images may vary considerably in resolution, ranging from low-quality to high-resolution inputs. When images are processed by age estimation frameworks, many systems including DeepFace and InsightFace apply internal preprocessing steps such as resizing or cropping to a fixed input size. Common backbone models rely on standard input resolutions, such as 112x112 pixels for ArcFace and 224x224 pixels for VGG-Face [3], [4]. These models were originally developed for face recognition [3], [4], but frameworks such as DeepFace repurpose them as feature extractors for related tasks, including age prediction [1]. Although these preprocessing steps standardize input size, it is unclear whether the resolution of the original image still affects the quality of the final prediction. However, it remains unclear whether input resolution still influences accuracy despite the internal resizing performed by these frameworks. This study uses DeepFace version 0.0.95 and InsightFace version 0.7.3. Previous research has shown that resolution can significantly affect the performance of neural networks in

general [5], but no dedicated study has examined how image resolution impacts age estimation specifically within DeepFace and InsightFace. Therefore, this work investigates whether input image resolution continues to affect age estimation accuracy, even though DeepFace and InsightFace internally resize images to fixed dimensions. If input resolution had no influence, all resolutions would yield identical accuracy. If measurable differences occur, this would indicate that external input resolution still affects the final predictions.

Research question: What is the impact of input image resolution on the accuracy of age estimation in DeepFace and InsightFace (evaluated using MAE, SD, and MedAE), despite their internal preprocessing pipelines?

II. RELATED WORK

A. Background on Age Estimation

Age estimation from facial images is an important problem within computer vision and pattern recognition. The goal is typically to predict the *chronological age* of a person based on a facial photograph. A related task is estimating the *apparent age*, which refers to how old a person appears rather than their biological age [6].

Researchers have approached this problem using deep neural networks. Commonly, large collections of facial images without apparent-age labels are first used to learn a general facial representation, after which the networks are fine-tuned on datasets that do include apparent-age annotations [6]. Nonetheless, the reliability of age estimation remains strongly dependent on the quality and resolution of the input images.

Age estimation can be addressed in two main ways:

- 1) Predicting the *chronological age*, i.e., the actual biological age of the individual;
- 2) Predicting the *apparent age*, i.e., how old the person appears in the image [6], [15].

In practice, *apparent age* is often used, as it aligns more closely with human perception of age and because these labels are easier to obtain through human annotators [20].

B. Regression vs. Classification

State-of-the-art approaches to age estimation can broadly be divided into two categories:

- 1) **Regression-based models:** These models treat age as a continuous variable and attempt to predict the exact

age (e.g., 24 years). Performance is typically measured using metrics such as Mean Absolute Error (MAE), which captures the average absolute deviation between predicted and true ages [7], [8].

- 2) **Classification-based models:** These models treat age as a discrete class or interval (e.g., 20–25 years, 26–30 years). The model predicts the age group rather than an exact value [8]. This approach can be more robust in practice, as individuals of different ages may look very similar. For example, a 28-year-old may visually resemble a 30-year-old; placing both in the same interval (26–30 years) helps reduce ambiguity.

C. Age Estimation Processing Pipeline

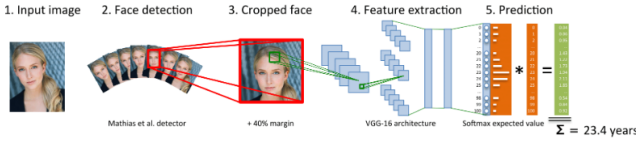


Fig. 1. Processing pipeline for age estimation [15].

Most age estimation models follow a similar processing pipeline (Fig. 1). First, a face is detected in the input image, after which the facial region is cropped with an appropriate margin.

The cropped face is then normalized and passed through a convolutional neural network (such as VGG-16 or ResNet) to extract facial features. Based on the extracted feature representation, the final age is predicted, typically through a regression or classification layer.

A critical step in this pipeline is the **cropping and resizing** of the input images, as this determines which visual details are preserved and made available to the feature extractor. The effectiveness of this step depends strongly on the resolution and quality of the original input image.

D. Datasets for Age Estimation

Large, well-labeled datasets are essential for training and evaluating age estimation models. These datasets vary in size, age distribution, diversity, and label quality, all of which directly affect the performance of deep learning models.

- 1) **IMDB-WIKI** [9] is the largest publicly available dataset for age and gender prediction, containing more than 500,000 images collected from IMDb and Wikipedia. The dataset spans ages from 0 to over 100 years, but suffers from an imbalanced age distribution, with a strong overrepresentation of young adults (15–35 years). Age labels are derived from the estimated date the photo was taken, making them chronological but also prone to noise due to incorrect metadata.
- 2) To reduce this label noise, the cleaned subset **IMDB-Clean** was introduced [10]. This subset removes incorrect or corrupted samples (e.g., blank images, non-face scenes, or incorrect age labels). As a result, IMDB-Clean provides more consistent annotations and is more suitable for experimental research, including this study.



Fig. 2. Comparison of sample images from IMDB-WIKI (top) and the cleaned IMDB-Clean dataset (bottom). Incorrect or irrelevant samples have been removed, improving overall data quality.

- 3) **UTKFace** [11] contains approximately 20,000 images with ages ranging from 0 to 116 years, labeled with age, gender, and ethnicity. Due to its wide age range and relatively balanced distribution, UTKFace is frequently used as a benchmark. However, its image resolution (200×200 px) is relatively low.
- 4) **FairFace** [12] focuses specifically on reducing demographic bias. The dataset includes more than 100,000 images and is balanced across ethnicities and genders. However, age labels are provided in broad categories (e.g., 0–2, 3–9, 10–19, etc.) rather than exact ages, which limits its usefulness for regression-based age estimation.
- 5) **VGGFace2** [13] consists of 3.31 million high-resolution images of 9,131 identities, with considerable variation in pose, age, illumination, and ethnicity. Although the dataset does not provide explicit age labels, it is widely used to train robust face feature extractors (e.g., ResNet and SENet), which are later applied to tasks including age estimation.

TABLE I
OVERVIEW OF COMMONLY USED DATASETS FOR AGE ESTIMATION

Dataset	Number of Images	Resolution / Quality	Age Range
UTKFace	20,000+	~ 200 × 200 px (low)	0–116
FairFace	100,000+	~ 300 × 300 px (medium)	Categories (0–2 to 70+)
IMDB-WIKI	500,000+	often >400 px (mixed)	0–100+
IMDB-Clean	Subset of IMDB-WIKI	up to 1024 px, cleaned (high)	0–100+
VGGFace2	3.31M	high-resolution web images (mixed)	No age labels

In summary: IMDB-WIKI and UTKFace are among the most widely used datasets for age estimation. However, because IMDB-WIKI contains substantial label noise, this study uses the cleaned **IMDB-Clean** subset [10]. A key advantage of IMDB-Clean is that it provides higher-resolution images (up to 1024 px) with more reliable age annotations. Although UTKFace is widely used and offers a balanced age distribution, the dataset consists primarily of low-resolution images (200×200 px). Therefore, UTKFace is less suitable for analyzing the impact of input resolution, as it does not provide sufficiently high-resolution samples for comparison. For this reason, IMDB-Clean is the most appropriate dataset for the objectives of this study.

E. Choice of Existing Frameworks

For this study, *DeepFace* [1] and *InsightFace* [2] were selected because both frameworks provide pretrained age estimation models that can be used directly without additional training. Each framework integrates detection, alignment, feature extraction, and regression within a standardized pipeline, enabling reproducible comparisons across different input resolutions.

DeepFace is a Python framework for face recognition and related analysis tasks, including age and gender prediction. It supports several state-of-the-art backbone models, such as VGG-Face, ArcFace, FaceNet, Dlib, and DeepID [1]. The standard pipeline consists of face detection, alignment, normalization, and embedding extraction. For age estimation, cropped faces are resized (typically to 224×224 pixels for VGG-Face or 112×112 pixels for ArcFace) and passed to a regression model that outputs a predicted age [3]. According to the official documentation, DeepFace reports an average error of approximately ± 4.65 years (MAE) on public test data [14].



Fig. 3. Example output from DeepFace: age, emotion, gender, and race predictions generated from a facial image [1].

InsightFace is an open-source framework originally developed for face recognition and built around the ArcFace architecture [4]. It provides built-in support for face detection using RetinaFace and SCRFD, both included in the official implementation [2]. Input images are resized by default to 112×112 pixels for processing by the ArcFace backbone. For age estimation, InsightFace includes dedicated models with a `genderage` head, in which face embeddings are first extracted and subsequently passed through a regression or classification layer to estimate age and gender.

Alternative Approaches Considered: Alternative approaches such as training custom CNN models from scratch or using OpenCV- or YOLO-based age estimation systems were considered but not selected. These alternatives generally require extensive training data and long training times, which fall outside the scope of this project.

Framework Selection: DeepFace and InsightFace were selected because they provide well-documented, pretrained age estimation models with integrated pipelines for detection, alignment, and inference. These properties enable consistent and reproducible evaluation across different input resolutions, making both frameworks suitable for the objectives of this study.

TABLE II
COMPARISON OF DEEPFACE AND INSIGHTFACE (2025)

Property	DeepFace [1]	InsightFace [2]
Open source	MIT License	MIT License
Backbones	VGG-Face, FaceNet, OpenFace, DeepID, ArcFace, Dlib	ArcFace, Partial FC, SubCenter ArcFace
Detection models	OpenCV, SSD, Dlib, MTCNN, RetinaFace	RetinaFace, SCRFD
Input format	224×224 (VGG-Face), 112×112 (ArcFace)	112×112 (ArcFace)
Updates	Active (2025, v0.0.95)	Active (2025, v0.7.3)

F. Architectures for Age Estimation

Many age estimation models rely on convolutional neural networks (CNNs). Two commonly used backbone architectures in existing frameworks are **VGG-Face** and **ArcFace**.

The **VGG-Face** model [3] is based on the VGG-16 architecture and uses input images of 224×224 pixels. The network contains approximately 134 million parameters and produces 4096-dimensional embeddings [1]. In frameworks such as DeepFace, VGG-Face is repurposed as a feature extractor, with an additional regression layer used for age prediction [21]. VGG-Face represents a classical, sequential CNN without residual connections and is computationally expensive.

ArcFace [4], on the other hand, is based on a ResNet architecture with residual connections and uses input images of 112×112 pixels. It contains roughly 34 million parameters, generates 512-dimensional embeddings, and introduces the *Additive Angular Margin Loss* to produce more discriminative face representations. Within InsightFace, ArcFace is extended with a `genderage` head, enabling direct prediction of age and gender. In summary, VGG-Face and ArcFace represent two different generations of CNN architectures: a classical, heavy CNN (VGG-Face) versus a more modern and efficient ResNet-based model (ArcFace). This architectural difference partially explains the superior performance and higher speed of InsightFace observed in the experiments of this study. Although more recent architectures exist such as FusionNet [17] or attention-based models [18] these fall outside the scope of this work.

TABLE III
ARCHITECTURES OF COMMONLY USED FACE RECOGNITION MODELS [1]

Model	Input Shape	Embedding Dim.	Parameters	Number of Layers
FaceNet-128d	$160 \times 160 \times 3$	128	$\sim 22\text{M}$	447
VGG-Face	$224 \times 224 \times 3$	4096	$\sim 134\text{M}$	36
ArcFace	$112 \times 112 \times 3$	512	$\sim 34\text{M}$	162
Dlib	$150 \times 150 \times 3$	128	$\sim 63\text{M}$	34
OpenFace	$96 \times 96 \times 3$	128	$\sim 3\text{M}$	166

Impact of Image Resolution: Factors such as resolution and image sharpness play a crucial role in age estimation, as subtle details (e.g., wrinkles and skin texture) can be lost at lower resolutions. The study “*Effects of Varying Resolution on Performance of CNN based Image Classification*” [5] shows that image resolution has a significant influence on CNN performance. Accuracy decreases notably when a model is trained on high-resolution images but tested on lower-resolution data.

TABLE IV
EFFECT OF VARYING IMAGE RESOLUTION ON CNN PERFORMANCE ON CIFAR-10. ACCURACY, PRECISION, AND F1-SCORE DECREASE SIGNIFICANTLY AT LOWER RESOLUTIONS [5].

CIFAR-10 Dataset Resolution	TOTV			TVT		
	Accuracy	Precision	F1 Score	Accuracy	Precision	F1 Score
28×28	0.9927	0.99269	0.99263	0.9927	0.99269	0.99263
21×21	0.9905	0.99062	0.99045	0.9924	0.99247	0.99234
14×14	0.9773	0.97828	0.97749	0.9854	0.98586	0.98544
7×7	0.6791	0.79071	0.68767	0.7770	0.84433	0.78416

Table IV confirms that image resolution is a critical factor when using pretrained frameworks for age estimation such as *DeepFace* [1] and *InsightFace* [2]. Similar effects have been

observed in other domains; for example, in medical image analysis (e.g., endoscopy), lower image resolution also leads to a clear drop in CNN performance [19]. For this reason, the present study explicitly investigates how different input image resolutions affect the accuracy of age predictions in DeepFace and InsightFace.

G. Dataset and Preprocessing

For this study, a customized subset of the IMDB-Clean dataset¹ was used. The selection consists of 1000 images that meet the following criteria:

- captured in 2010 or later (based on file metadata),
- at least 1024 pixels in width or height,
- containing exactly one face (detected using the default InsightFace model:

```
FaceAnalysis(allowed_modules=["detection"])).
```

Filtering was performed using a Python script that automatically removed images with no detectable face or with multiple faces. The final dataset is publicly available on Hugging Face.²

Each original image was resized into seven versions with the following resolutions: 64×64 , 112×112 , 224×224 , 256×256 , 512×512 , 720×720 , and 1080×1080 pixels. The Pillow library was used for resizing. Since all images are square, no letterboxing (padding) was required.

All versions of each image were stored in a single directory, with the resolution appended to the filename (e.g., `_64.jpg`, `_512.jpg`).



Fig. 4. Example of a single image in seven different resolutions.

H. Experimental Environment

All experiments were conducted in Google Colab using an NVIDIA T4 GPU and Python 3 as the runtime environment (see Fig. 5). The frameworks DeepFace (v0.0.95) and InsightFace (v0.7.3) were installed via pip:

```
!pip install deepface==0.0.95 insightface==0.7.3
!pip install onnxruntime-gpu-1.23.0
```

For DeepFace, age prediction was performed using:

```
DeepFace.analyze(img_path=img_path, actions=["age"],
                 enforce_detection=False)
```

This uses the default DeepFace backbone for age estimation, namely VGG-Face, as implemented in version 0.0.95.

For InsightFace, the following configuration was used:

```
FaceAnalysis(allowed_modules=["detection", "genderage"])
```

In this setup, face detection is performed using SCRFD, and age estimation is handled by the built-in genderage head operating on the ArcFace backbone. The full code and processing pipeline are publicly available [16].

¹<https://www.kaggle.com/datasets/yuulind/imdb-clean>

²<https://huggingface.co/datasets/codershiyar/imdb-clean-2010plus-singleface>

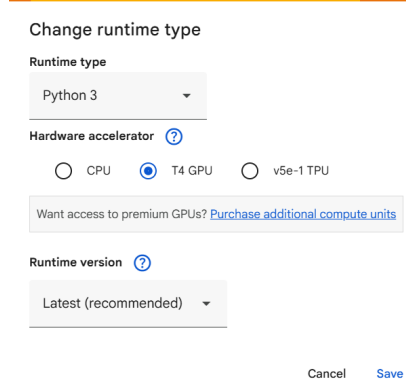


Fig. 5. Screenshot of the Colab runtime (T4 GPU, Python 3).

I. Experimental Design

All experiments used the same 1000 images (the selected subset of IMDB-Clean). For each image, seven different resolutions were evaluated (64×64 up to 1080×1080). Each resolution was tested using both frameworks:

- **DeepFace (v0.0.95)** – default detection and pretrained regression-based age analysis.
- **InsightFace (v0.7.3)** – default detection (SCRFD) and age estimation via the built-in genderage head.

Applying both frameworks systematically across all resolutions ensures a fair comparison. Performance was evaluated using MAE, SD, and MedAE (see Section II-J).

J. Evaluation Metrics

To assess the performance of DeepFace and InsightFace across different resolutions, the **Mean Absolute Error (MAE)** was computed for each model–resolution combination. MAE represents the average difference, in years, between the predicted and true age.

In addition, two supplementary statistics were included:

- **Standard Deviation (SD)** – measures the variability of the absolute errors. A high SD indicates irregular or unstable predictions, while a low SD indicates consistent performance.
- **Median Absolute Error (MedAE)** – the median absolute error. This metric is more robust than the mean and highlights whether MAE is influenced by outliers.

Together, MAE, SD, and MedAE provide a more complete picture of both the accuracy and the reliability of the age predictions at each resolution.

All calculations were performed on the exact same set of 1000 faces per resolution, ensuring a fair comparison across models and resolutions. The final results are summarized in Table V.

K. Formulas Used

For each resolution and each model, the following formulas were applied, where y_i denotes the predicted age, \hat{y}_i the ground-truth age, and n the number of images:

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

$$SD = \sqrt{\frac{1}{n} \sum_{i=1}^n (|y_i - \hat{y}_i| - MAE)^2}$$

$$MedAE = \text{median}(|y_1 - \hat{y}_1|, \dots, |y_n - \hat{y}_n|)$$

L. Processing Time Measurement

For both frameworks (DeepFace and InsightFace), 100 images were analyzed per resolution (from 64×64 to 1080×1080 pixels), resulting in a total of 700 images per framework. Each image was processed ten times to minimize variation caused by background processes or system activity. For each image, the mean processing time and standard deviation were computed; subsequently, the average of these values was calculated per resolution in seconds (see Table VII). The processing time includes all internal steps performed by the framework, such as face detection and inference, which were executed automatically by the framework.

III. RESULTS AND DISCUSSION

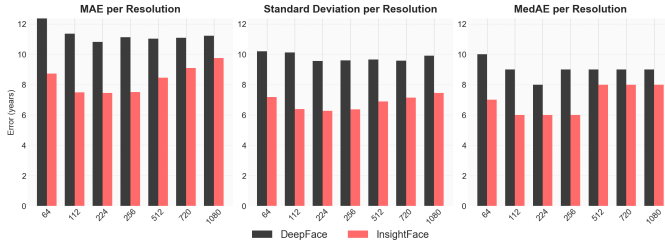


Fig. 6. Comparison of DeepFace and InsightFace across different image resolutions based on MAE, SD, and MedAE.

TABLE V

MEAN ABSOLUTE ERROR (MAE), STANDARD DEVIATION (SD), AND MEDIAN ABSOLUTE ERROR (MEDAE) PER RESOLUTION.

Resolution	DeepFace			InsightFace		
	MAE ↓	SD ↓	MedAE ↓	MAE ↓	SD ↓	MedAE ↓
64×64	12.36	10.19	10.0	8.72	7.19	7.0
112×112	11.36	10.12	9.0	7.48	6.38	6.0
224×224	10.83	9.57	8.0	7.46	6.28	6.0
256×256	11.13	9.60	9.0	7.51	6.36	6.0
512×512	11.02	9.65	9.0	8.45	6.89	8.0
720×720	11.08	9.58	9.0	9.09	7.14	8.0
1080×1080	11.22	9.90	9.0	9.76	7.45	8.0

Table V and Fig. 6 show the error metrics (MAE, SD, and MedAE) for each resolution using DeepFace and InsightFace. Both frameworks achieve their lowest MAE at 224×224 pixels: 10.83 years for DeepFace and 7.46 years for InsightFace.

At 64×64 , the error increases to 12.36 (DeepFace, +14.1%) and 8.72 (InsightFace, +16.9%) compared to 224×224 . Previous work suggests that this may be related to the loss of age-relevant facial details such as wrinkles and skin texture when the resolution is too low [5].

At very high resolutions (1080×1080), DeepFace remains relatively stable (11.22 years; +3.6%), while InsightFace degrades more strongly, reaching 9.76 years (+30.8%). DeepFace remains consistently around ~ 11 years MAE, whereas InsightFace exhibits a gradual decline in performance when

moving above 224×224 . This may indicate that InsightFace’s internal resizing step handles extreme downscaling from high-resolution images to 112×112 pixels less effectively, causing loss of useful details.

A notable observation is that InsightFace performs slightly better at an external input resolution of 224×224 (MAE = 7.46 years) than at its native 112×112 input (MAE = 7.48 years), even though the internal resize still converts all inputs to 112×112 . This suggests that higher input resolutions may improve face detection (SCRFD) and alignment before the internal resizing occurs, thereby improving overall accuracy.

TABLE VI
AVERAGE ERROR METRICS (MAE, SD, MedAE) ACROSS ALL RESOLUTIONS.

Framework	MAE ↓	SD ↓	MedAE ↓
DeepFace	11.28	9.80	9.00
InsightFace	8.35	6.81	7.00

Across all resolutions, InsightFace outperforms DeepFace on average (Table VI). The mean MAE is 8.35 years for InsightFace and 11.28 years for DeepFace. In addition to MAE, the Standard Deviation (SD) and Median Absolute Error (MedAE) provide further insight into model behavior. For InsightFace at 224×224 pixels, the SD is 6.28 years and the MedAE is 6.0 years, indicating reasonably consistent errors and showing that half of the predictions deviate by at most 6 years from the true age. DeepFace exhibits higher SD and MedAE values, suggesting greater variability and lower typical accuracy.

These differences can be linked to the underlying architectures: InsightFace uses ArcFace with a ResNet backbone and Additive Angular Margin Loss, whereas DeepFace relies on VGG-Face with softmax loss [3], [4].

Processing time: Although the primary research question focuses on accuracy, processing time was also examined to assess the computational efficiency of both frameworks at different resolutions. This is relevant for real-time applications and systems with limited computational resources. Table VII reports the average per-image processing time for both frameworks across all resolutions.

InsightFace is substantially faster and remains nearly constant between 0.015 s and 0.021 s per image across all resolutions. DeepFace, however, shows a clear increase in processing time as resolution increases from an average of 0.052 ± 0.070 s at 64×64 to 0.733 ± 0.106 s at 1080×1080 . Notably, DeepFace is slightly slower at 64×64 than at 112×112 , which may be due to internal upscaling requirements or less stable detection at very low resolutions.

Considering the combination of lower error rates (Table V) and significantly higher processing speed, InsightFace can be considered both **more accurate** and **more efficient** than DeepFace in these experiments.

IV. CONCLUSION

This study demonstrates that input image resolution has a clear and measurable impact on the accuracy of age estimation

TABLE VII

AVERAGE PROCESSING TIME PER RESOLUTION (IN SECONDS) FOR DEEPFACE AND INSIGHTFACE, INCLUDING STANDARD DEVIATION.

Resolution	DeepFace (Mean \pm SD)	InsightFace (Mean \pm SD)
64x64	0.052 \pm 0.070 s	0.016 \pm 0.006 s
112x112	0.039 \pm 0.002 s	0.015 \pm 0.001 s
224x224	0.074 \pm 0.005 s	0.016 \pm 0.001 s
256x256	0.086 \pm 0.007 s	0.016 \pm 0.001 s
512x512	0.222 \pm 0.022 s	0.017 \pm 0.001 s
720x720	0.381 \pm 0.044 s	0.019 \pm 0.001 s
1080x1080	0.733 \pm 0.106 s	0.021 \pm 0.001 s

using DeepFace and InsightFace, despite their internal preprocessing steps. Both frameworks achieve optimal performance at an input resolution of 224×224 pixels, yielding the lowest Mean Absolute Error (MAE). Deviations from this resolution, both lower and higher, lead to reduced accuracy. InsightFace is more sensitive to extremely high resolutions (+31% at 1080×1080) than DeepFace (+3.6%).

A notable finding is that InsightFace performs slightly better with 224×224 pixel inputs (MAE: 7.46 years) compared to its native 112×112 input (MAE: 7.48 years), even though all images are internally resized to 112×112 . This indicates that the external input resolution still influences the final prediction quality.

In addition to achieving higher accuracy, InsightFace also demonstrates substantially faster processing times across all resolutions, reflecting a more efficient underlying architecture. Because the IMDB-Clean dataset consists of images of celebrities, the demographic distribution (e.g., age, gender, ethnicity) may not fully represent the general population, which could limit the generalizability of the findings.

Future research could therefore focus on evaluating DeepFace and InsightFace across different demographic groups to identify potential bias in age estimation. Considering the balance between speed and accuracy, InsightFace with an input resolution of 224×224 pixels is recommended for real-time applications and systems with limited computational resources.

REFERENCES

- [1] S. Serengil and A. Ozpinar, "A benchmark of facial recognition pipelines and co-usability performances of modules," *Journal of Information Technologies*, vol. 17, no. 2, pp. 95–107, 2024. doi: 10.17671/gaz-ibtd.1399077. Available: <https://dergipark.org.tr/en/download/article-file/3573195>. Accessed: Oct. 14, 2025.
- [2] J. Guo, J. Deng, and S. Zafeiriou, "InsightFace: 2D and 3D Face Analysis Project," GitHub repository, 2025. Available: <https://github.com/deepinsight/insightface>. [Accessed: Oct. 5, 2025]
- [3] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *Proc. British Machine Vision Conf. (BMVC)*, 2015. Available: <https://www.robots.ox.ac.uk/~vgg/publications/2015/Parkhi15/parkhi15.pdf>. Accessed: Oct. 14, 2025.
- [4] J. Deng, J. Guo, J. Yang, N. Xue, I. Kotsia, and S. Zafeiriou, "ArcFace: Additive angular margin loss for deep face recognition," *arXiv preprint arXiv:1801.07698*, Jan. 2018. Available: <https://arxiv.org/abs/1801.07698>. Accessed: Oct. 14, 2025.
- [5] S. P. Kannoja and G. Jaiswal, "Effects of varying resolution on performance of CNN based image classification: An experimental study," *Int. J. Comput. Sci. Eng.*, vol. 6, no. 9, pp. 451–456, Sept. 2018. doi: 10.26438/ijcse/v6i9.451456. Available: https://www.researchgate.net/publication/328960034_Effects_of_Varying_Resolution_on_Performance_of_CNN_based_Image_Classification_An_Experimental_Study. Accessed: Oct. 14, 2025.

- [6] Y. Zhu, Y. Li, G. Mu, and G. Guo, "A study on apparent age estimation," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCVW)*, Santiago, Chile, Dec. 2015, pp. 25–31. doi: 10.1109/ICCVW.2015.43. Available: <https://ieeexplore.ieee.org/document/7406392>. Accessed: Oct. 14, 2025.
- [7] A. Garain, B. Ray, P. K. Singh, and A. Ahmadian, "GRANet: A deep learning model for classification of age and gender from facial images," *IEEE Access*, vol. 9, pp. 85672–85689, June 2021. doi: 10.1109/ACCESS.2021.3085971.
- [8] J. Chen, L. Cheng, X. Yang, J. Liang, B. Quan, and S. Li, "Joint learning with both classification and regression models for age prediction," *J. Phys.: Conf. Ser.*, vol. 1168, 032016, 2019. doi: 10.1088/1742-6596/1168/3/032016.
- [9] R. Rothe, R. Timofte, and L. Van Gool, "DEX: Deep EXpectation of apparent age from a single image," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCVW)*, Dec. 2015. Available: <https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki/>. Accessed: Oct. 14, 2025.
- [10] Y. Lin, L. Shen, Y. Wang, and M. Pantic, "FP-Age: Leveraging face parsing attention for facial age estimation in the wild," *arXiv preprint arXiv:2106.11145*, Jun. 2021. doi: 10.48550/arXiv.2106.11145. Available: <https://arxiv.org/abs/2106.11145>. Accessed: Oct. 14, 2025.
- [11] Z. Zhang, Y. Song, and H. Qi, "Age progression/regression by conditional adversarial autoencoder," *arXiv preprint arXiv:1702.08423*, Mar. 2017. doi: 10.48550/arXiv.1702.08423. Available: <https://arxiv.org/abs/1702.08423>. Accessed: Oct. 14, 2025.
- [12] K. Karkkainen and J. Joo, "FairFace: Face attribute dataset for balanced race, gender, and age for bias measurement and mitigation," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2021, pp. 1548–1558. doi: 10.1109/WACV48630.2021.00159. Available: <https://ieeexplore.ieee.org/document/9423296>. Accessed: Oct. 14, 2025.
- [13] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman, "VGGFace2: A dataset for recognising faces across pose and age," in *Proc. 2018 13th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG 2018)*, Xi'an, China, May 2018, pp. 67–74. doi: 10.1109/FG.2018.00020. Available: <https://ieeexplore.ieee.org/document/8373813>. Accessed: Oct. 14, 2025.
- [14] S. Serengil, "DeepFace: A lightweight face recognition and facial attribute analysis framework," GitHub repository, 2025. Available: <https://github.com/serengil/deepface>. Accessed: Oct. 14, 2025.
- [15] R. Rothe, R. Timofte, and L. Van Gool, "Deep expectation of real and apparent age from a single image without facial landmarks," *Int. J. Comput. Vis.*, vol. 126, pp. 144–157, 2018. doi: 10.1007/s11263-016-0940-3.
- [16] S. Jamo, "Experimentcode: Effect van beeldresolutie op leeftijdschatting," Colab. Available: <https://colab.research.google.com/drive/1roSZ2Y4Ne9Yy71DM2mRmY0VetfuydbEi>. Accessed: Oct. 27, 2025.
- [17] H. Wang, X. Wei, V. Sanchez, and C.-T. Li, "Fusion network for face-based age estimation," *arXiv preprint arXiv:1807.10421*, Jul. 2018. doi: 10.48550/arXiv.1807.10421. Available: <https://arxiv.org/abs/1807.10421>. Accessed: Oct. 14, 2025.
- [18] H. Wang, V. Sanchez, and C.-T. Li, "Improving face-based age estimation with attention-based dynamic patch fusion," *arXiv preprint arXiv:2112.10167*, Dec. 2021. doi: 10.48550/arXiv.2112.10167. Available: <https://arxiv.org/abs/2112.10167>. Accessed: Oct. 14, 2025.
- [19] V. Thambawita, S. A. Hicks, I. Strümke, P. Halvorsen, S. Parasa, and M. A. Riegler, "Impact of image resolution on deep learning performance in endoscopy image classification," *Diagnostics*, vol. 11, no. 12, Art. no. 2183, Dec. 2021. doi: 10.3390/diagnostics11122183. Available: <https://www.mdpi.com/2075-4418/11/12/2183>. Accessed: Oct. 5, 2025.
- [20] O. Agbo-Ajala, S. Viriri, M. Oloko-Oba, and O. Ekundayo, "Apparent age prediction from faces: A survey of modern approaches," *Frontiers in Big Data*, vol. 5, Art. no. 1025806, Oct. 2022. doi: 10.3389/fdata.2022.1025806. Available: <https://doi.org/10.3389/fdata.2022.1025806>. Accessed: Oct. 7, 2025.
- [21] Z. Qawaqneh, A. A. Mallouh, and B. D. Barkana, "Deep convolutional neural network for age estimation based on VGG-Face model," *arXiv preprint arXiv:1709.01664*, Sep. 2017. doi: 10.48550/arXiv.1709.01664. Available: <https://arxiv.org/abs/1709.01664>. Accessed: Oct. 14, 2025.