electronics

MDPI

Article

Facial Age Estimation Using Multi-Stage Deep Neural Networks

Salah Eddine Bekhouche [1,2], Azeddine Benlamoudi [3,4], Fadi Dornaika [1,5,*], Hichem Telli [2] and Yazid Bounab [6]

1 Department of Computer Science and Artificial Intelligence, Faculty of Informatics, University of the Basque Country—UPV/EHU, 20018 San Sebastian, Spain; sbekhouche001@ikasle.ehu.eus
2 Laboratoire de Vision et des Systémes de Communication, University of Biskra, Biskra 07 000, Algeria; tellihicham@univ-biskra.dz
3 Laboratoire de Génie Électrique, Faculté des Nouvelles Technologies de l'information et de la Communication, Université Kasdi Merbah Ouargla, Ouargla 30 000, Algeria; benlamoudi.azeddine@univ-ouargla.dz
4 Institut d'Electronique de Microélectronique et de Nanotechnologie (IEMN), UMR 8520, Université Polytechnique Hauts de France, Université de Lille, CNRS, 59313 Valenciennes, France
5 IKERBASQUE, Basque Foundation for Science, 48009 Bilbao, Spain
6 Center for Machine Vision and Signal Analysis (CMVS), Faculty of Information Technology and Electrical Engineering (ITEE), University of Oulu, 90570 Oulu, Finland
* Correspondence: fadi.dornaika@ehu.eus

Abstract: Over the last decade, the world has witnessed many breakthroughs in artificial intelligence, largely due to advances in deep learning technology. Notably, computer vision solutions have significantly contributed to these achievements. Human face analysis, a core area of computer vision, has gained considerable attention due to its wide applicability in fields such as law enforcement, social media, and marketing. However, existing methods for facial age estimation often struggle with accuracy due to limited feature extraction capabilities and inefficiencies in learning hierarchical representations. This paper introduces a novel framework to address these issues by proposing a Multi-Stage Deep Neural Network (MSDNN) architecture. The MSDNN architecture divides each CNN backbone into multiple stages, enabling more comprehensive feature extraction, thereby improving the accuracy of age predictions from facial images. Our framework demonstrates a significant performance improvement over traditional solutions, with its effectiveness validated through comparisons with the EfficientNet and MobileNetV3 architectures. The proposed MSDNN architecture achieves a notable decrease in Mean Absolute Error (MAE) across three widely used public datasets (MORPH2, CACD, and AFAD) while maintaining a virtually identical parameter count compared to the initial backbone architectures. These results underscore the effectiveness and feasibility of our methodology in advancing the field of age estimation, showcasing it as a robust solution for enhancing the accuracy of age prediction algorithms.

Keywords: age estimation; deep learning; multilevel deep features; adaptive regression

check for updates

Citation: Bekhouche, S.E.; Benlamoudi, A.; Dornaika, F.; Telli, H.; Bounab, Y. Facial Age Estimation Using Multi-Stage Deep Neural Networks. *Electronics* **2024**, *13*, 3259. https://doi.org/10.3390/electronics13163259

Academic Editors: Prayag Tiwari, Guang Yang, Changshuo Wang, Gang Wang and Ruiping Wang

Received: 14 June 2024
Revised: 12 August 2024
Accepted: 13 August 2024
Published: 16 August 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Biometric technologies, particularly facial analysis, have seen a tremendous surge in advancements, driven largely by the advent of artificial intelligence and deep learning [1]. These remarkable achievements have opened the door to further investigations into the usability of facial features in various applications, including demographics (age, gender, and ethnicity), affective computing (emotion and pain recognition), and security and authentication (face verification, detecting spoofing and facial attacks). These technologies have become increasingly crucial in a wide array of applications, ranging from demographic analysis to security systems. Among these applications, age estimation based on facial features has emerged as a particularly important field due to its diverse applications and inherent challenges [2].

The topic of age estimation from face images continues to receive much attention from computer vision researchers due to its importance and use in various applications [3–5]. Age estimation through facial analysis is paramount in demographic studies, providing vital insights into population age distributions, which in turn are crucial for both sociological research and governmental policy-making [6]. In law enforcement and security, accurate age estimation is indispensable for identifying individuals in investigations and enhancing security protocols through advanced face verification systems [7]. Moreover, in the field of affective computing, age estimation plays a key role in personalizing user experiences and improving human–computer interactions [8].

The use of facial features for age estimation is highly recognized for its reliability and the speed with which assessments can be completed, making it generally preferred over other biometric markers. Advances in machine learning and image processing have further enhanced the viability of facial analysis tools for this purpose. Recent developments such as descriptive human visual cognitive strategies using graph neural networks underscore that innovative approaches in facial expression recognition can contribute to the nuanced detection and interpretation of age-defining facial features [9]. However, despite these technological strides, the field continues to face significant hurdles. One of the core issues is the development of robust age estimation systems that can consistently deliver accurate predictions across diverse demographic groups. This difficulty stems primarily from the varying nature and progression of facial aging among different individuals and populations, a variability that is well-documented in the literature [7,10]. Moreover, factors such as environmental influences, genetics, and lifestyle choices significantly impact the aging process, adding further complexity to accurate age estimation using facial features [6,11]. As a result, ongoing research seeks to refine algorithms and methodologies to better handle these variations and improve the accuracy and applicability of age estimation technologies.

The advent of deep learning, particularly through the development of convolutional neural networks (CNNs), has dramatically transformed the landscape of computer vision, and by extension age estimation models. These deep learning methods have shown remarkable superiority in accuracy and efficiency compared to traditional techniques. Their strength lies particularly in managing large and diverse datasets with enhanced effectiveness [12]. Our research builds upon this innovative groundwork by proposing novel modifications to existing CNN architectures. These enhancements aim to further refine the accuracy of age estimation models while addressing some of the persistent challenges in the field. Utilizing advanced neural network strategies, we seek to push the boundaries of what is currently achievable in age estimation accuracy, making significant contributions to both theory and practical applications [13,14].

The goal of this study is to improve the performance of existing deep neural networks by making simple modifications to CNN architectures that enable more accurate prediction of a person's age and age group based on facial images. The proposed idea divides each CNN backbone into multiple stages, each of which generates a feature map that is then concatenated with the other stages using adaptive average pooling. In addition, an adaptive loss function is used to improve the training of the modified backbones. To empirically validate our proposed modifications, we employ two renowned CNN architectures: EfficientNet and MobileNetV3. These networks were selected for their efficiency and proven effectiveness in various computer vision tasks [15,16]. We conducted extensive experiments using three publicly available datasets: MORPH2, CACD, and AFAD. These datasets were chosen for their diversity and comprehensiveness in order to ensure that our model is robust and able to generalize between different demographic groups.

The main contributions of our work are summarized as follows:

- We introduce a straightforward yet effective architecture that leverages channels at various stages of a deep neural network within the regression module.
- Our approach is implemented with two distinct CNN architectures: EfficientNet and MobileNetV3. Notably, this scalable method can be adapted to work with any CNN architecture.

- We conduct experiments using three public datasets: MORPH2, CACD, and AFAD. We provide an extensive comparison with several leading-edge methods, demonstrating that age estimation accuracy can be enhanced using the same backbone network without incurring additional costs.

The remainder of this paper is structured as follows. In Section 2, we dive into the different methods of determining age from facial images. This section comprehensively reviews the evolution of techniques, starting from classical methods that rely on texture and shape feature extraction to the more sophisticated deep learning approaches that have revolutionized the field. We discuss the strengths and limitations of each method, providing a holistic understanding of the landscape of age estimation techniques.

Section 3 introduces our proposed framework, which marks a significant shift from traditional methods. We detail our innovative approach, which involves modifications to CNN architectures to enhance their predictive accuracy in age estimation. This section explains the rationale behind dividing the CNN backbone into multiple stages and the benefits of concatenating these feature maps through adaptive average pooling. Moreover, we elaborate on the adaptive loss function, a crucial component designed to fine-tune the training process of these modified backbones.

In Section 4, we present our experimental setup and the results obtained from testing our framework on three public datasets: MORPH2, CACD, and AFAD. This section is crucial as it not only validates our proposed method but also provides a comprehensive comparison with existing state-of-the-art methods. We discuss the implications of our findings, emphasizing how our modifications lead to improvements in age estimation without incurring additional computational costs.

Finally, Section 5 summarizes the key points of our research, highlighting the advances made in age estimation through our approach. We also discuss potential areas for future research, including suggestions that can further enhance the accuracy and efficiency of age estimation techniques in biometric systems.

This paper aims to not only contribute to the academic discourse in computer vision and biometrics but also to offer practical insights that can be applied in various domains where age estimation is crucial, such as security, marketing, and healthcare.

2. Related Work

Age estimation from facial images has attracted significant attention in the field of computer vision [17], primarily due to its wide-ranging applications in areas such as security, marketing and human–computer interaction. This research domain has evolved through two major methodologies: classical methods and deep learning approaches.

Classical approaches to age estimation largely focus on the extraction and analysis of facial features. Key techniques include Local Binary Patterns (LBP); introduced by Bekhouche et al. [18], LBP has been extensively used for texture analysis. Its effectiveness in capturing fine-grained textural changes makes it particularly useful for distinguishing age-related features in facial images. Scale-Invariant Feature Transform (SIFT), introduced by Ren et al. [19], leverages SIFT to capture keypoints that are resistant to changes in image scale, rotation, and illumination. This method is instrumental in identifying age-related variations in facial structure. Active Appearance Model (AAM), as discussed by Tian et al. [20], provides a framework for both shape and texture analysis, allowing for comprehensive modeling of facial dynamics with aging. Binarized Statistical Image Features (BSIF) and Local Phase Quantization (LPQ), highlighted in works by Dornaika [21] and Bekhouche [22], respectively, focus on encoding texture information in a way that is more resistant to variations in lighting and other extrinsic factors.

In conjunction with these feature extraction techniques, various machine learning algorithms such as Support Vector Machine (SVM) [23], Partial Least Squares (PLS) [24], and Coupled Similarity Reference Coding Model (CSRC) [25] have been employed to perform the classification or regression tasks necessary for age estimation.

The advent of deep learning has revolutionized age estimation, offering more robust and accurate models. As mentioned by Kong et al. [26], Convolutional Neural Networks (CNNs) have become a fundamental tool in age estimation thanks to their ability to automatically and hierarchically extract features from raw images. Shen et al.'s [27] proposal of Deep Regression Forests (DRFs) showcases an innovative integration of CNNs with decision forests, tailoring the model to handle the diverse and complex data distributions found in age-related features. The Identity-Preserving Generative Adversarial Network (IPCGAN) framework by Wang et al. [28] signifies a significant leap in the generation of advanced facial images while preserving the identity of the subjects. This approach has implications not only for age estimation but also for digital entertainment and forensics. The introduction of robust loss functions by Dornaika et al. [29] and innovative label distribution learning techniques by Akbari et al. [30] exemplify the ongoing efforts to refine the training process of deep neural networks for more accurate and stable age estimation.

Transfer learning has become a cornerstone in this domain [31], allowing researchers to apply the knowledge gained from one task to another. In age estimation, this means utilizing pretrained models on large datasets, which is especially beneficial in light of the high variability and subtlety of age-related features in facial images. This approach not only saves significant computational resources but also improves the generalization capabilities of age estimation models.

Gil Levi and Tal Hassner [13] utilized convolutional neural networks; their paper represents an early transition phase in which deep learning began to gain prominence for age and gender classification. Their focus was on the effectiveness of CNNs in handling real-world uncontrolled images. Rothe et al. [12] introduced the DEX algorithm, which estimates apparent age from a single image using deep learning techniques without relying on facial landmarks. Antipov et al. [32] focused on the challenge of age estimation in children using a combination of general and child-specialized deep learning models. Yang et al. [33] presented SSR-Net, a new compact and efficient architecture for age estimation. Hossein Hassani and Amirhossein Hosseini [34] introduced AgeNet, combining regression and classification approaches in deep learning for age estimation. These papers illustrate the evolution from classic methods to sophisticated deep learning techniques in the field of face age estimation.

The field of age estimation from facial images has seen remarkable advancements, evolving from initial reliance on classical feature-based methods to the incorporation of sophisticated deep learning models [35]. This evolution is a testament to the dynamic nature of research in biometrics and computer vision. Initially, classical feature-based methods laid the groundwork. These methods, focusing on the characteristics of texture, shape, and appearance extracted from facial images, were crucial in establishing a basic understanding of how facial characteristics correlate with age. Techniques such as Local Binary Patterns (LBP), Gabor filters, and Active Appearance Models (AAM) were among those commonly employed. Although these methods provided significant information, they were often limited by their dependency on manual feature selection and susceptible to variations in image quality, lighting, and facial expressions.

The advent of deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized the field [36]. These models automatically learn hierarchical feature representations, making them highly effective in capturing complex age-related patterns in facial images. Techniques such as transfer learning and data augmentation further enhance the robustness and accuracy of these models. Deep learning approaches excel in handling large datasets and diverse age groups, offering a level of precision and generalization that the classical methods struggled to achieve.

However, deep learning models come with their own set of challenges. They require substantial computational resources and large annotated datasets for training. Moreover, the "black box" nature of these models often leads to difficulties in interpretability, making it challenging to understand the specific features the model is focusing on for age estimation.

As we move into the next section focusing on our proposed approach, it is important to consider both the advantages and disadvantages of these methodologies. Our approach aims to integrate the interpretability and simplicity of classical methods with the advanced feature extraction capabilities of deep learning models. This hybrid strategy is designed to take advantage of the strengths of both paradigms, offering enhancing accuracy and robustness while maintaining computational efficiency and ensuring a degree of interpretability. Through this approach, we aim to push the boundaries of automatic age estimation while balancing the trade-offs between complexity, accuracy, and practical applicability.

3. Methodology

Age estimation, like any image-related task, essentially involves two main steps: preprocessing, which includes face detection, normalization, and augmentation; and modeling, which encompasses feature extraction, modeling, and prediction. In this pipeline, the quality and appropriateness of preprocessing and feature extraction play a crucial role in enhancing the performance of age estimation models. Figure 1 shows the overall architecture that we adopt for estimating facial age. These steps collectively form a comprehensive pipeline that ensures the robustness and generalization ability of age estimation models across diverse facial characteristics and conditions.

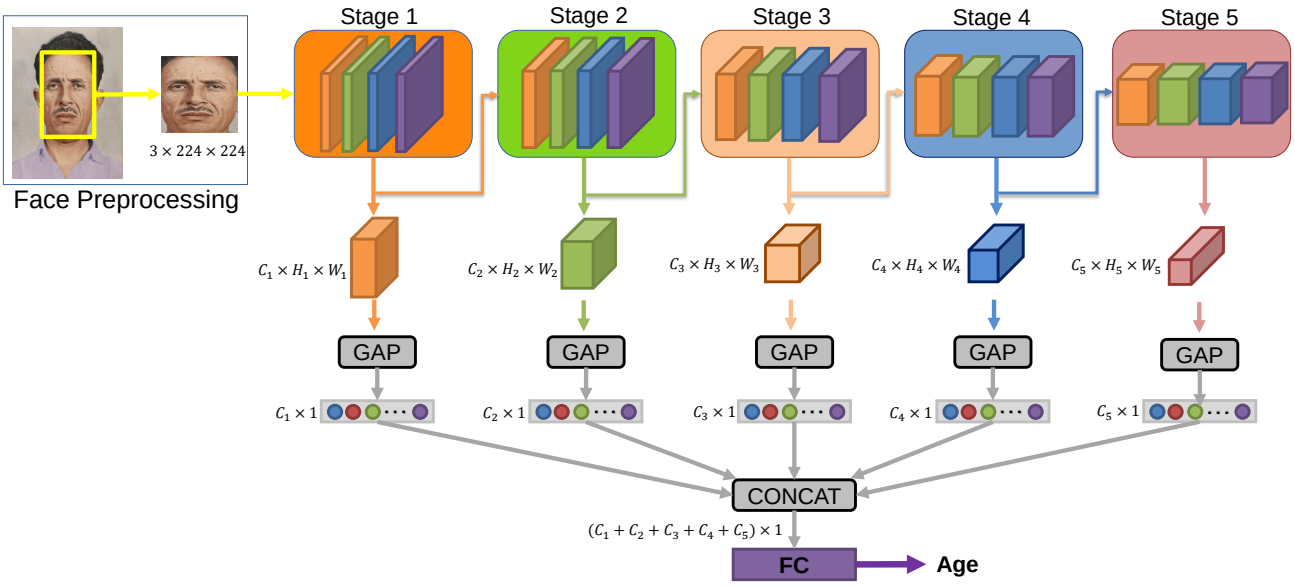


Figure 1. General structure of a multistage deep neural network based on a backbone with five stages.

3.1. Preprocessing

The initial step of age estimation from facial images involves performing face detection, including locating and identifying facial regions within an image. Following detection, the faces are normalized to standardize their size, orientation, and illumination. Normalization ensures that subsequent analyses are not influenced by variations in facial appearance. Augmentation techniques are then applied to diversify the dataset, enhancing the model's ability to generalize to different facial variations, expressions, and environmental conditions [37]. This step is crucial for training robust age estimation models capable of handling real-world scenarios.

To ensure the quality of the preprocessing steps, the highly efficient MTCNN face detection and alignment model [38] is employed. MTCNN excels in handling challenging scenarios, including partial occlusion and shadows. It is utilized to acquire a cropped and aligned face, which is then adjusted to meet the 224 × 224 standards of the age model backbone. Throughout the training phase of the proposed approach, various data enhancement techniques are applied to the aligned face images. These techniques encompass alterations

to brightness, contrast, and saturation, random conversion of a color image to grayscale, and horizontal flipping of the image.

3.2. Proposed Architecture: Multi-Stage Deep Neural Nets (MSDNN)

In contrast to traditional machine learning approaches, deep learning fundamentally alters the landscape of age estimation by sidestepping the labor-intensive process of manual feature extraction. Instead, these advanced neural networks automatically learn relevant features from the data, a capability that is especially potent in the context of facial analysis for age estimation. Classical machine learning techniques are heavily dependent on domain expertise for feature selection, which may not effectively capture all the subtle complexities of aging facial features; however, deep learning excels at discerning and leveraging intricate nonlinear relationships within data through its multiple layers and transformations. This ability enables the model to identify and use subtle visual cues associated with age, such as fine lines, wrinkles, and texture changes, which can be challenging to quantify manually but are critical for accurate age predictions. Moreover, as these networks delve deeper they refine these features, enhancing their ability to generalize from training data to new unseen images, thereby improving reliability and accuracy in real-world applications. This shift not only streamlines the analytical pipeline but also potentially increases the accuracy of age estimation models, making them more robust across diverse populations and varying image conditions.

Furthermore, CNN architectures are designed to capture various levels of abstraction at different layers within the network, which is pivotal for progressively refining the understanding of image content relevant to tasks such as age estimation. In the initial layers of a CNN, the extracted features are typically low-level, such as edges, textures, and colors. These foundational elements are essential for the preliminary interpretation of visual data. As the data progress through the network, subsequent layers focus on higher-level features that incorporate more complex aspects of the image, such as specific facial attributes associated with aging, for example the shape and sag of facial contours or the presence of age-related spots [39,40]. Each successive layer abstracts and compounds information from the previous layers, enabling the network to make more sophisticated inferences about age from facial characteristics. This hierarchical processing not only improves the precision of age estimation but also enhances the network's ability to adapt to different facial idiosyncrasies and demographic variations, ultimately leading to more accurate and robust age prediction models.

The motivation behind this research is to harness the benefits offered by varying levels of abstraction and seamlessly integrate them into the age regression module. In contrast to age estimation methods solely based on deep learning, our approach stands out by directly amalgamating all types of features through the regression module. This strategy ensures a comprehensive utilization of diverse information encapsulated in different levels of abstraction, paving the way for a more holistic and nuanced understanding of age-related characteristics in the data. By integrating multiple layers of abstraction, our methodology aims to enhance the accuracy and robustness of age regression, ultimately contributing to the refinement and optimization of age estimation models (Equation (2)).

Consider a CNN architecture consisting of $\boldsymbol{L}$ layers, each performing similar or varied transformations. The initial layer takes an input image and generates the primary feature map. Subsequently, each following layer $\boldsymbol{i}$ processes the feature map FM_{i-1} generated by its predecessor $\boldsymbol{i-1}$. As we progress from the first layer ($\boldsymbol{i = 0}$) to the final layer ($\boldsymbol{i = L}$), the level of abstraction of the features systematically increases. This progressive increase in abstraction allows the network to transition from simple discernible patterns to more complex abstract representations. This enhances its ability to analyze and interpret the intricate details pertinent to tasks such as age estimation from facial features.

$$FM_i = FM_1 \oplus FM_2 \oplus \cdots \oplus FM_i \oplus \cdots \oplus FM_L \tag{1}$$

$$Features = FM_1 \oplus FM_2 \oplus \cdots \oplus FM_i \oplus \cdots \oplus FM_L \quad (2)$$

The flowchart of the proposed architecture is shown in Figure 1. The proposed architecture was tested on two types of CNN backbones, namely, EfficientNet and MobileNetV3.

3.2.1. Efficientnet

Since AlexNet won the ImageNet competition in 2012, CNNs have been the de facto method for numerous deep learning tasks, especially computer vision. Since 2012, researchers have experimented in attempting to develop succeedingly better architectures to increase the accuracy of models on various tasks [41]. Tan et al. proposed the EfficientNet concept in "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks" [15]. They studied scaling and discovered that a careful balance between the depth, width, and resolution of the network can improve performance. EfficientNet is a design for a convolutional neural network that uses a compound to equally scale all dimensions of depth/width/resolution to l. Unlike the common practice of setting these elements freely, the EfficientNet scaling approach scales width, depth, and resolution uniformly using a set of predetermined coefficients. The EfficientNet B0 base model provides a good balance between accuracy and FLOPs using neural architecture research. The base model has been scaled using scaling approaches to create the B1–B7 family of models. Despite recent claims of improvements in learning speed or inference, these are often simply EfficientNet in terms of their parameters and FLOP efficiency. EfficientNet Edge is a variant of EfficientNet with a different input size and mesh depth and width. In this work, we use as a backbone architecture EfficientNet Edge Large B, which assumes an input size of 300×300 and extends the depth and width by 1.4 and 1.2, respectively.

3.2.2. Mobilenetv3

MobileNet [42] is a lightweight deep neural network designed for mobile computer vision applications. MobileNetV3 [43] is a new generation of the MobileNet family. It is a platform-independent network that uses NAS [44] and NetAdapt [45] to search and define two models: MobileNetV3-Small for low-resource use cases, and MobileNetV3-Large for high-resource use cases. MobileNetV3 uses different activation functions with nonlinearity depending on the target layer, and uses squeeze-and-excite in its bottleneck blocks. In this work, we use MobileNetV3-Large pretrained on ImageNet-1K with a scaling factor of 1.25 for the width of the layers.

3.3. Loss Function Used

The adaptive loss function used in [29] exploits both the ℓ_2- and ℓ_1-norms. It is provided by

$$L_{Ada} = \frac{1}{N} \sum_{i=1}^{N} \frac{(1+\sigma)\,(Y_i - T_i)^2}{|Y_i - T_i| + \sigma}, \quad (3)$$

where σ is a positive parameter that affects the shape of the loss function, while Y_i and T_i denote the predicted and ground-truth ages of the i-th image. The adaptive loss function becomes comparable to the MAE loss function when σ approaches zero, and becomes identical to the MSE loss function as σ approaches infinity.

4. Performance Evaluation

We begin this section by introducing the three public datasets used in our experiments. Subsequently, we outline the evaluation metrics we employed. We present the performance of our proposed scheme and compare it with other competing methods. The section includes an analysis and discussion of the results obtained from each dataset. Following that, we provide a concise ablation study. Finally, we conclude with a cross-dataset evaluation that assesses the generalization capability of the proposed solution.

4.1. Datasets

To assess the effectiveness of the proposed approach, a comprehensive evaluation using three publicly available benchmark datasets (AFAD, CACD, MORPH2) is conducted. These datasets were chosen due to their extensive use in the field of face age estimation, allowing for meaningful comparisons with other state-of-the-art-approaches. In addition, a cross-dataset protocol is adopted across the three datasets, further enhancing the rigor and the generalizability of the overall assessment. This protocol, which spans across multiple datasets, not only increases the comprehensiveness of the experiments but also ensures that the findings are transferable and applicable in diverse real-world scenarios. The datasets are divided into training, validation, and test sets following the same split proposed by [46].

MORPH2 (Craniofacial Longitudinal Morphological Face Database) [10] is a large longitudinal face database containing 55,134 images of 13 K subjects labeled with age, gender, and race. Thousands of facial images of subjects taken in the real world make up the MORPH database corpus (uncontrolled conditions). Ages used in this dataset ranged from 16 to 70 years.

CACD (Cross-Age Celebrity Dataset) [47] contains 159,449 images of 2000 celebrities collected from the internet ranging in age from 14 to 62 years. The images were collected from search engines using celebrity names and years as keywords.

AFAD (Asian Face Age Dataset) [48] contains 165,501 faces ranging in age from 15 to 40 years. The images were collected from the RenRen Social Network, which is used by Asian students. The dataset includes middle school, high school, undergraduate, and graduate students.

4.2. Evaluation Metrics

The performance measures used to evaluate the proposed facial age estimation model are the Mean Absolute Error (MAE) and the Cumulative Score (CS). The MAE is the average of the absolute errors between the ground-truth ages and predicted ages. The MAE is calculated as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|p_i - g_i| \tag{4}$$

where N, p_i, and g_i are the total number of samples, predicted age, and ground-truth age, respectively. The CS indicates the percentage of tested cases for which the age estimation error is smaller than a threshold. The CS is provided by

$$CS(T) = \frac{N_{e\leq T}}{N}\%, \tag{5}$$

where T, N, and $N_{e\leq T}$ are the error threshold (years), total number of samples, and number of samples for which the age estimate has an absolute error no greater than the threshold T, respectively; thus, CS indicates the percentage of tested samples that were correctly predicted within the tolerance T.

4.3. Implementation

MSDNN is implemented in Python 3.8 and uses the Pytorch 2.0 and Pytorch lightning frameworks. The model is trained with distributed data parallelism, which can be run on single/multiple machines, and half-precision tensors are used to reduce training time. The implementation was run on a laptop with a Core I7 CPU and a GeForce RTX 2070 GPU. By default, we trained the model using the AdamW optimizer with ReduceLRonPlateau scheduling and a batch size of 32 for 100 epochs; the initial learning rate of the optimizer was set to 1×10^{-3} and the learning rate limit to 1×10^{-7}.

4.4. Comparison with State-of-the-Art Methods

The results show that the proposed MSDNN architecture based on the Efficient-Net backbone for facial age estimation using three public datasets (AFAD, CACD, and

MORPH2) outperforms previous state-of-the-art approaches in terms of MAE, as shown in Table 1.

The proposed MSDNN architecture achieves a significant improvement, with an MAE of 4.90 compared to the best SOTA approach, (CORAL-CNN), which had an MAE of 5.25 using the CACD dataset. In another comparison the performance was notably better, with an MAE of 2.59 compared to the best SOTA approach (ADPF), which had an MAE of 2.71 using the MORPH2 dataset. Similarly, the difference in performance was minimal, with an MAE of 3.25 compared to the best SOTA approach (CDCNN), which had an MAE of 3.30 using the AFAD dataset. In addition, the MSDNN model has 8.5 k parameters, whereas the best models for ADPF, CDCNN, and CORAL-CNN have 14 M, roughly 21 M (based on ResNet34), and 138 M (based on VGG16), respectively.

Figure 2 shows the cumulative scores of the proposed model based on the EfficientNet backbone on the three public datasets. Comparing the results from the three datasets (see the curves in Figure 2), it can be seen that the results for the MORPH2 dataset are more stable and accurate compared to the AFAD and CACD datasets, which is due to the age imbalance in the datasets. From the figure, the percentages of correct age estimates with respect to 5 years of absolute error are 86.66%, 65.31%, and 80.17% for the MORPH2, CACD, and AFAD datasets, respectively. The latter shows that the CACD dataset is more challenging compared to the other two datasets, as the percentage of correct age estimates with respect to 10 years of absolute error in the CACD dataset is only 88.86%.

Overall, the use of EfficientNet as a backbone combined with features from different stages shows significant performance improvements across the three datasets. The highest performance was reported on the MORPH2 dataset, with an MAE of 2.59, followed by the CACD and AFAD datasets, with MAEs of 3.25 and 4.90, respectively.

The significant improvements in MAE suggest that the combination of EfficientNet and MSDNN effectively captures age-related features across diverse datasets. The experimental effects demonstrate that the proposed architecture can generalize well even on datasets with varying characteristics. Despite these promising results, this study has limitations. For instance, the performance may vary with different preprocessing techniques or when applied to datasets not included in this study. Potential biases could arise from the specific demographics of the datasets used here, which may not fully represent the global population. Future research should explore these aspects and test the model's robustness across a wider range of data in order to validate its applicability and address any inherent biases.

In our experiments, we observed that the Multi-Stage Deep Neural Network (MSDNN) architecture significantly improved the accuracy of age estimation compared to the baseline models. Specifically, MSDNN demonstrated a lower Mean Absolute Error (MAE) across all tested datasets. This improvement can be attributed to the effective utilization of feature maps at different stages of the CNN backbone, which provides a more comprehensive representation of age-related features. Despite these positive results, our study has several limitations. First, the datasets we used (MORPH2, CACD, and AFAD) may not fully represent the diversity of real-world facial images, potentially limiting the generalizability of our findings. Additionally, our model's performance may be influenced by the quality and resolution of the input images as well as variations in lighting and facial expressions. Future studies should aim to validate our approach on more diverse and comprehensive datasets. Our study may be subject to bias related to the demographic compositions present in the datasets; for example, the MORPH2 dataset mainly includes images of individuals from a specific age range and ethnic background, which might not reflect the global population. To mitigate this, we recommend that future research include more diverse datasets encompassing various age groups, ethnicities, and environmental conditions in order to help ensure that the model's performance is robust and unbiased across different demographic groups.

Table 1. Mean absolute error (MAE) (years) obtained with the three public datasets: comparison with state-of-the-art.

Approach	**MORPH2**
SGD [49]	5.69
OR-CNN [48]	3.27
Google-Net [50]	2.94
CORN [51]	2.98
Soft-Ranking [52]	2.83
IMDB-WIKI (ResNet-50) [53]	2.81
deep-JREAE [54]	2.77
CDCNN [55]	2.76
Adaptive-CNN [29]	2.75
SADAL [56]	2.75
MSFCL [57]	2.73
Wasserstein [58]	2.71
CORAL-CNN [46]	2.64
ADPF [59]	2.71
Ours	**2.59**
Approach	**CACD**
CAAE [60]	44.20
Lifespan [61]	11.70
IPCGAN [28]	9.10
S^2 GAN [62]	8.40
PADA [63]	7.71
UAN [64]	7.57
OPDA-BP [65]	7.48
ORUDA [66]	7.26
DANCE [67]	7.13
Self-Estimate [68]	6.70
ADPF [59]	5.39
CORAL-CNN [46]	5.25
Ours	**4.90**
Approach	**AFAD**
PADA [63]	7.11
OPDA-BP [65]	6.84
UAN [64]	6.73
DANCE [67]	6.25
ORUDA [66]	6.19
CNN + LSVR [14]	5.56
CORAL-CNN [46]	3.47
OR-CNN [48]	3.34
CDCNN [55]	3.30
Ours	**3.25**

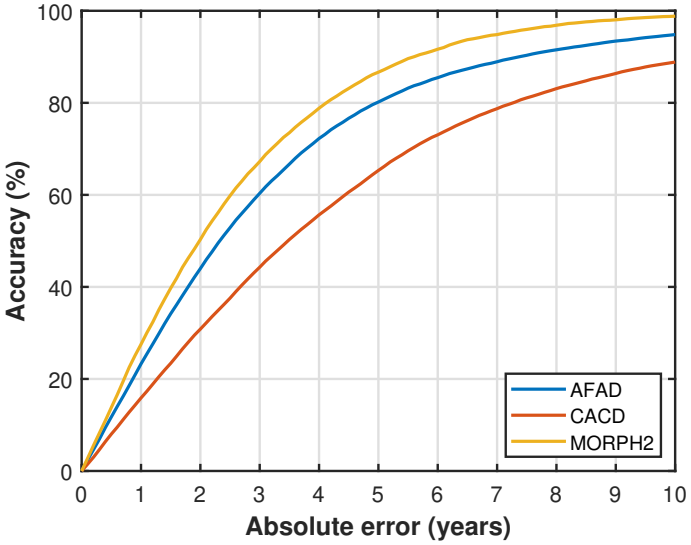


Figure 2. Cumulative scores obtained by the proposed approach on the three public datasets.

4.5. Ablation Study

To prove the effectiveness of the MSDNN concept, two different experiments were conducted by training the two different backbones with and without MSDNN. The results are shown in Table 2. For the MORPH2 dataset and the EfficientNet backbone, the use of multistage features in the regression module improved MAE by 0.15 years, while the number of parameters increased by 304 over 8.6 million. For the MORPH2 dataset and the MobileNetV3 backbone, the improvement was 0.19 years and the number of parameters increased by 248 over 4.4 million.

In both cases, the increase in model parameters is negligible in terms of computational cost compared to the improvement in estimation capability, which is considered a good tradeoff.

Table 2. Performance on the MORPH2 dataset.

Backbone	MSDNN	Params	MAE (Years)	Time
EfficientNet	×	8,693,521	2.74	16.1 ms
	✓	8,693,825	2.59	16.2 ms
MobileNetV3	×	4,399,945	2.93	13.0 ms
	✓	4,400,193	2.72	13.0 ms

4.6. Cross-Dataset Evaluation

An interesting property in age estimation is the generalization ability of a given model. This generalization can be quantified by performing a cross-dataset evaluation using one whole dataset for training and another dataset for testing. We conducted extensive cross-dataset testing using the same three public datasets. The experiments were performed using test subsets, with Table 3 showing the results of all three datasets. The best generalization was obtained with MORPH2 as the training set and the AFAD dataset as the testing set. The results on the AFAD test dataset seem to be close to each other, which can be attributed to the age interval of this dataset, which ranges from 15 to 40 years. In general, the performance decreased significantly over the course of the experiments, which can be explained by the demographic diversity of the faces and the photo conditions of each dataset.

Table 3. Performance in cross-dataset experiments on the testing subsets.

Train	Test	MAE	CS
MORPH2	CACD	10.62	26.23%
	AFAD	7.15	40.22%
CACD	MORPH2	8.38	33.41%
	AFAD	9.54	33.89%
AFAD	MORPH2	8.10	41.89%
	CACD	9.59	29.42%

5. Conclusions

The accurate estimation of age is not merely a scientific pursuit but a pivotal tool with far-reaching implications across diverse domains, from healthcare and forensic science to biometric security and demographic analysis. In this paper, we introduce a novel Multi-Stage Deep Neural Network (MSDNN) architecture that uses feature maps at different levels of a CNN backbone to yield more discriminative features. This approach shows its usefulness in the inference phase. We tested MSDNN on three public datasets: AFAD, CACD, and MORPH2, where we also adopted a cross-dataset testing protocol. As backbones, we used two famous edge device architectures: EfficientNet and MobileNetV3. Due to the portability of these architectures, we used them to validate the effectiveness of the proposed architecture against state-of-the-art approaches in terms of time and accuracy. In addition, we conducted an ablation study to investigate the efficiency of the proposed MSDNN concept.

The results show that when the MSDNN was trained on MORPH2, the MEA for AFAD and CACD was 8.38 and 8.10, respectively. On the other hand, when the MSDNN was trained on CACD, the MAEs were 10.62 and 9.59 for MORPH2 and AFAD, respectively. Finally, when the MSDNN was trained on AFAD, the MAEs were 7.15 and 9.54 for MORPH2 and CACD, respectively. These results can be explained by the fact that the datasets each contain different information, including variations in lighting conditions, image sizes, and image quality. Overall, despite these different dynamics, MSDNN shows the ability to mostly exceed the performance of SOTA approaches on the different datasets, indicating its effectiveness and robustness.

In the future, we plan to use modules for spatial and channel attention between the different stages; we also envision the use of visual transformers to improve the performance of age estimation.

Author Contributions: Conceptualization, S.E.B. and F.D.; methodology, S.E.B., F.D. and A.B.; software, S.E.B. and H.T.; validation, S.E.B., F.D., A.B., H.T. and Y.B.; writing—original draft preparation, S.E.B., F.D. and A.B.; writing—review and editing, S.E.B., F.D., A.B. and Y.B.; supervision, F.D. and S.E.B.; funding acquisition, F.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work was partially supported by grant PID2021-126701OB-I00 funded by MCIN/AEI/10.13039/ 501100011033 and by 'ERDF: A way of making Europe'.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare that this research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Zhao, W.; Chellappa, R.; Phillips, P.J.; Rosenfeld, A. Face recognition: A literature survey. *ACM Comput. Surv. (CSUR)* **2003**, *35*, 399–458. [CrossRef]
2. Jain, A.K.; Li, S.Z. *Handbook of Face Recognition*; Springer: Berlin/Heidelberg, Germany, 2011; Volume 1.

3. Bekhouche, S.E. Facial Soft Biometrics: Extracting Demographic Traits. Ph.D. Thesis, Faculté des Sciences et Technologies, Vandœuvre-lès-Nancy, France, 2017.
4. Guehairia, O.; Dornaika, F.; Ouamane, A.; Taleb-Ahmed, A. Facial Age Estimation Using Tensor Based Subspace Learning and Deep Random Forests. *Inf. Sci.* **2022**, *609*, 1309–1317. [CrossRef]
5. Kim, S.; Kim, H.; Lee, E.S.; Lim, C.; Lee, J. Risk score-embedded deep learning for biological age estimation: Development and validation. *Inf. Sci.* **2022**, *586*, 628–643. [CrossRef]
6. Lanitis, A.; Draganova, C.; Christodoulou, C. Comparing different classifiers for automatic age estimation. *IEEE Trans. Syst. Man Cybern. Part B (Cybern.)* **2004**, *34*, 621–628. [CrossRef] [PubMed]
7. Fu, Y.; Guo, G.; Huang, T.S. Age synthesis and estimation via faces: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 1955–1976. [PubMed]
8. Asif, M.K.; Nambiar, P.; Ibrahim, N.; Al-Amery, S.M.; Khan, I.M. Three-dimensional image analysis of developing mandibular third molars apices for age estimation: A study using CBCT data enhanced with Mimics & 3-Matics software. *Leg. Med.* **2019**, *39*, 9–14. [CrossRef]
9. Liu, S.; Huang, S.; Fu, W.; Lin, J.C.W. A descriptive human visual cognitive strategy using graph neural network for facial expression recognition. *Int. J. Mach. Learn. Cybern.* **2024**, *15*, 19–35. [CrossRef]
10. Ricanek, K.; Tesafaye, T. Morph: A Longitudinal Image Database of Normal Adult Age-Progression. In Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition (FGR06), Southampton, UK, 10–12 April 2006; pp. 341–345.
11. Guo, G.; Mu, G. A framework for joint estimation of age, gender and ethnicity on a large database. *Image Vis. Comput.* **2014**, *32*, 761–770. [CrossRef]
12. Rothe, R.; Timofte, R.; Van Gool, L. Deep expectation of real and apparent age from a single image without facial landmarks. *Int. J. Comput. Vis.* **2018**, *126*, 144–157. [CrossRef]
13. Levi, G.; Hassner, T. Age and Gender Classification using Convolutional Neural Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Boston, MA, USA, 7–12 June 2015; pp. 34–42.
14. Wang, X.; Guo, R.; Kambhamettu, C. Deeply-Learned Feature for Age Estimation. In Proceedings of the 2015 IEEE Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 5–9 January 2015; pp. 534–541.
15. Tan, M.; Le, Q. Efficientnet: Rethinking Model Scaling for Convolutional Neural Networks. In Proceedings of the International Conference on Machine Learning, Long Beach, CA, USA, 9–15 June 2019; pp. 6105–6114.
16. Howard, T.C. *Why Race and Culture Matter in Schools: Closing the Achievement Gap in America's Classrooms*; Teachers College Press: New York, NY, USA, 2019.
17. Aonty, S.S.; Deb, K.; Sarker, I.H. Attention-Based Human Age Estimation from Face Images to Enhance Public Security. *Data* **2023**, *8*, 145. [CrossRef]
18. Bekhouche, S.E.; Ouafi, A.; Taleb-Ahmed, A.; Hadid, A.; Benlamoudi, A. Facial Age Estimation using BSIF and LBP. In Proceedings of the First International Conference on Electrical Engineering ICEEB'14, Biskra, Algeria, 7–8 December 2014.
19. Ren, H.; Li, Z.N. Age Estimation Based on Complexity-Aware Features. In Proceedings of the Asian Conference on Computer Vision, Singapore, 1–5 November 2014; pp. 115–128.
20. Tian, Q.; Cao, M.; Chen, S.; Yin, H. Relationships Self-Learning Based Gender-Aware Age Estimation. *Neural Process. Lett.* **2019**, *50*, 2141–2160. [CrossRef]
21. Dornaika, F.; Arganda-Carreras, I.; Belver, C. Age estimation in facial images through transfer learning. *Mach. Vis. Appl.* **2019**, *30*, 177–187. [CrossRef]
22. Bekhouche, S.E.; Ouafi, A.; Benlamoudi, A.; Taleb-Ahmed, A.; Hadid, A. Facial Age Estimation and Gender Classification using Multi Level Local Phase Quantization. In Proceedings of the 2015 3rd International Conference on Control, Engineering & Information Technology (CEIT), Tlemcen, Algeria, 25–27 May 2015; pp. 1–4.
23. Guo, G.; Mu, G.; Fu, Y.; Huang, T.S. Human Age Estimation using Bio-Inspired Features. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 112–119.
24. Guo, G.; Mu, G. Simultaneous Dimensionality Reduction and Human Age Estimation via Kernel Partial Least Squares Regression. In Proceedings of the CVPR, Colorado Springs, CO, USA, 20–25 June 2011; pp. 657–664.
25. Wu, Y.; Hu, H.; Li, H. Age-invariant face recognition using coupled similarity reference coding. *Neural Process. Lett.* **2019**, *50*, 397–411. [CrossRef]
26. Kong, C.; Wang, H.; Luo, Q.; Mao, R.; Chen, G. Deep Multi-Input Multi-Stream Ordinal Model for age estimation: Based on spatial attention learning. *Future Gener. Comput. Syst.* **2023**, *140*, 173–184. [CrossRef]
27. Shen, W.; Guo, Y.; Wang, Y.; Zhao, K.; Wang, B.; Yuille, A.L. Deep Regression Forests for Age Estimation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 2304–2313.
28. Wang, Z.; Tang, X.; Luo, W.; Gao, S. Face Aging with Identity-Preserved Conditional Generative Adversarial Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7939–7947.
29. Dornaika, F.; Bekhouche, S.E.; Arganda-Carreras, I. Robust regression with deep CNNs for facial age estimation: An empirical study. *Expert Syst. Appl.* **2020**, *141*, 112942. [CrossRef]
30. Akbari, A.; Awais, M.; Bashar, M.; Kittler, J. A Theoretical Insight Into the Effect of Loss Function for Deep Semantic-Preserving Learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *34*, 1–15. [CrossRef] [PubMed]

31. Van, E.; Hospital, R. Innovative Approaches to Clinical Diagnosis: Transfer Learning in Facial Image Classification for Celiac Disease Identification. *Appl. Sci.* **2024**, *14*, 6207. [CrossRef]
32. Antipov, G.; Baccouche, M.; Berrani, S.A.; Dugelay, J.L. Apparent Age Estimation from Face Images Combining General and Children-Specialized Deep Learning Models. In Proceedings of the Conference on Computer Vision and Pattern Recognition Workshops, Las Vegas, NV, USA, 27–30 June 2016; pp. 96–104.
33. Yang, T.Y.; Huang, Y.H.; Lin, Y.Y.; Hsiu, P.C.; Chuang, Y.Y. Ssr-net: A compact soft stagewise regression network for age estimation. In Proceedings of the IJCAI, Stockholm, Sweden, 13–19 July 2018; Volume 5, p. 7.
34. Liu, X.; Li, S.; Kan, M.; Zhang, J.; Wu, S.; Liu, W.; Han, H.; Shan, S.; Chen, X. Agenet: Deeply Learned Regressor and Classifier for Robust Apparent Age Estimation. In Proceedings of the IEEE International Conference on Computer Vision Workshops, Santiago, Chile, 7–13 December 2015; pp. 16–24.
35. ELKarazle, K.; Raman, V.; Then, P. Facial Age Estimation Using Machine Learning Techniques: An Overview. *Big Data Cogn. Comput.* **2022**, *6*, 128. [CrossRef]
36. Kang, J.S.; Kim, C.S.; Lee, Y.W.; Cho, S.W.; Park, K.R. Age estimation robust to optical and motion blurring by deep residual CNN. *Symmetry* **2018**, *10*, 108. [CrossRef]
37. Liu, X.; Zou, Y.; Kuang, H.; Ma, X. Face image age estimation based on data augmentation and lightweight convolutional neural network. *Symmetry* **2020**, *12*, 146. [CrossRef]
38. Zhang, K.; Zhang, Z.; Li, Z.; Qiao, Y. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Process. Lett.* **2016**, *23*, 1499–1503. [CrossRef]
39. Yang, J.; Qian, H.; Zou, H.; Xie, L. Learning decomposed hierarchical feature for better transferability of deep models. *Inf. Sci.* **2021**, *580*, 385–397. [CrossRef]
40. Zhou, J.; Gan, J.; Gao, W.; Liang, A. Image retrieval based on aggregated deep features weighted by regional significance and channel sensitivity. *Inf. Sci.* **2021**, *577*, 69–80. [CrossRef]
41. Authors, V. Dental Age Estimation Using Deep Learning: A Comparative Survey. *Computation* **2023**, *8*, 145. [CrossRef]
42. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.
43. Howard, A.; Sandler, M.; Chu, G.; Chen, L.C.; Chen, B.; Tan, M.; Wang, W.; Zhu, Y.; Pang, R.; Vasudevan, V.; et al. Searching for mobilenetv3. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 1314–1324.
44. Tan, M.; Chen, B.; Pang, R.; Vasudevan, V.; Sandler, M.; Howard, A.; Le, Q.V. Mnasnet: Platform-Aware Neural Architecture Search for Mobile. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 2820–2828.
45. Yang, T.J.; Howard, A.; Chen, B.; Zhang, X.; Go, A.; Sandler, M.; Sze, V.; Adam, H. Netadapt: Platform-Aware Neural Network Adaptation for Mobile Applications. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 285–300.
46. Cao, W.; Mirjalili, V.; Raschka, S. Rank consistent ordinal regression for neural networks with application to age estimation. *Pattern Recognit. Lett.* **2020**, *140*, 325–331. [CrossRef]
47. Chen, B.C.; Chen, C.S.; Hsu, W.H. Cross-Age Reference Coding for Age-Invariant Face Recognition and Retrieval. In Proceedings of the Computer Vision—ECCV 2014, Zurich, Switzerland, 6–12 September 2014; Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T., Eds.; Springer: Cham, Switzerland, 2014; pp. 768–783.
48. Niu, Z.; Zhou, M.; Wang, L.; Gao, X.; Hua, G. Ordinal Regression with Multiple Output cnn for Age Estimation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 4920–4928.
49. Akbari, A.; Awais, M.; Feng, Z.; Farooq, A.; Kittler, J. A Flatter Loss for Bias Mitigation in Cross-Dataset Facial Age Estimation. In Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10–15 January 2021; pp. 10629–10635.
50. Dagher, I.; Barbara, D. Facial age estimation using pre-trained CNN and transfer learning. *Multimed. Tools Appl.* **2021**, *80*, 20369–20380. [CrossRef]
51. Shi, X.; Cao, W.; Raschka, S. Deep neural networks for rank-consistent ordinal regression based on conditional probabilities. *Pattern Anal. Appl.* **2023**, *26*, 941–955. [CrossRef]
52. Zeng, X.; Huang, J.; Ding, C. Soft-Ranking Label Encoding for Robust Facial Age Estimation. *IEEE Access* **2020**, *8*, 134209–134218. [CrossRef]
53. Paplham, J.; Franc, V. A Call to Reflect on Evaluation Practices for Age Estimation: Comparative Analysis of the State-of-the-Art and a Unified Benchmark. *arXiv* **2023**, arXiv:2307.04570.
54. Tian, Q.; Cao, M.; Sun, H.; Qi, L.; Mao, J.; Cao, Y.; Tang, J. Facial age estimation with bilateral relationships exploitation. *Neurocomputing* **2021**, *444*, 158–169. [CrossRef]
55. Zhang, B.; Bao, Y. Cross-dataset learning for age estimation. *IEEE Access* **2022**, *10*, 24048–24055. [CrossRef]
56. Liu, H.; Sun, P.; Zhang, J.; Wu, S.; Yu, Z.; Sun, X. Similarity-aware and variational deep adversarial learning for robust facial age estimation. *IEEE Trans. Multimed.* **2020**, *22*, 1808–1822. [CrossRef]
57. Xia, M.; Zhang, X.; Liu, W.; Weng, L.; Xu, Y. Multi-Stage Feature Constraints Learning for Age Estimation. *IEEE Trans. Inf. Forensics Secur.* **2020**, *15*, 2417–2428. [CrossRef]

58. Akbari, A.; Awais, M.; Fatemifar, S.; Khalid, S.S.; Kittler, J. A Novel Ground Metric for Optimal Transport-Based Chronological Age Estimation. *IEEE Trans. Cybern.* **2021**, *52*, 9986–9999. [CrossRef] [PubMed]
59. Wang, H.; Sanchez, V.; Li, C.T. Improving Face-Based Age Estimation with Attention-Based Dynamic Patch Fusion. *IEEE Trans. Image Process.* **2022**, *31*, 1084–1096. [CrossRef] [PubMed]
60. Zhang, Z.; Song, Y.; Qi, H. Age Progression/Regression by Conditional Adversarial Autoencoder. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 5810–5818.
61. Or-El, R.; Sengupta, S.; Fried, O.; Shechtman, E.; Kemelmacher-Shlizerman, I. Lifespan Age Transformation Synthesis. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020; Springer: Berlin/Heidelberg, Germany, 2020; pp. 739–755.
62. He, Z.; Kan, M.; Shan, S.; Chen, X. S2gan: Share Aging Factors across Ages and Share Aging Trends among Individuals. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 9440–9449.
63. Cao, Z.; Ma, L.; Long, M.; Wang, J. Partial Adversarial Domain Adaptation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 135–150.
64. You, K.; Long, M.; Cao, Z.; Wang, J.; Jordan, M.I. Universal Domain Adaptation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 2720–2729.
65. Saito, K.; Yamamoto, S.; Ushiku, Y.; Harada, T. Open Set Domain Adaptation by Backpropagation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 153–168.
66. Boris, C.; Sadek, A.; Wolf, C. Universal Domain Adaptation in Ordinal Regression. *arXiv* **2021**, arXiv:2106.11576.
67. Saito, K.; Kim, D.; Sclaroff, S.; Saenko, K. Universal domain adaptation through self supervision. *arXiv* **2020**, arXiv:2002.07953.
68. Li, Z.; Jiang, R.; Aarabi, P. Continuous Face Aging via Self-Estimated Residual Age Embedding. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 15008–15017.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.