



OPEN ACCESS

EDITED BY

M. Hassaballah,
South Valley University, Egypt

REVIEWED BY

Khulumani Sibanda,
University of Fort Hare, South Africa
Ayodele Adebisi,
Covenant University, Nigeria

*CORRESPONDENCE

Serestina Viriri
viriris@ukzn.ac.za

SPECIALTY SECTION

This article was submitted to
Machine Learning and Artificial
Intelligence,
a section of the journal
Frontiers in Big Data

RECEIVED 23 August 2022

ACCEPTED 28 September 2022

PUBLISHED 26 October 2022

CITATION

Agbo-Ajala O, Viriri S, Oloko-Oba M,
Ekundayo O and Heymann R (2022)
Apparent age prediction from faces: A
survey of modern approaches.
Front. Big Data 5:1025806.
doi: 10.3389/fdata.2022.1025806

COPYRIGHT

© 2022 Agbo-Ajala, Viriri, Oloko-Oba,
Ekundayo and Heymann. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Apparent age prediction from faces: A survey of modern approaches

Olatunbosun Agbo-Ajala¹, Serestina Viriri^{1,2*},
Mustapha Oloko-Oba¹, Olufisayo Ekundayo¹ and
Reolyn Heymann²

¹Computer Science Discipline, School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Durban, South Africa, ²Electrical and Electronic Engineering Science, University of Johannesburg, Johannesburg, South Africa

Apparent age estimation *via* human face image has attracted increased attention due to its numerous real-world applications. Predicting the apparent age has been quite difficult for machines and humans. However, researchers have focused on machine estimation of “age as perceived” to a high level of accuracy. To further improve the performance of apparent age estimation from the facial image, researchers continue to examine different methods to enhance its results further. This paper presents a critical review of the modern approaches and techniques for the apparent age estimation task. We also present a comparative analysis of the performance of some of those approaches on the apparent facial aging benchmark. The study also highlights the strengths and weaknesses of each approach used for apparent age estimation to guide in choosing the appropriate algorithms for future work in the field. The work focuses on the most popular algorithms and those that appear to have been the most successful for apparent age estimation to improve on the existing state-of-the-art results. We based our evaluations on three facial aging datasets, including looking at people (LAP)-2015, LAP-2016, and APPA-REAL, the most popular and publicly available datasets benchmark for apparent age estimation.

KEYWORDS

apparent age, convolutional neural network, deep learning, facial aging, age prediction

1. Introduction

Age estimation is a very prolific area of research within the computer vision community (Huerta et al., 2015; Drobnýh and Polovinkin, 2017). There has been an increasing interest in age estimation from facial images (Drobnýh and Polovinkin, 2017) due to its increasing demands in various potential applications in security control (Abbas and Kareem, 2018), human-computer interaction (Abbas and Kareem, 2018),

social media (Ruiz-Del-Solar et al., 2009), and forensic studies (Bouchrika et al., 2016). Although this subject has been extensively studied, the ability to estimate human ages reliably and correctly from face images is still far from satisfying human performance level (Onifade, 2015). There exist two kinds of facial age estimation: One is real (biological) age estimation, which determines the precise chronological or biological age of a person from the facial image (Shen et al., 2018); the other is apparent age estimation (Agustsson et al., 2017), which focuses on “how old does a person looks like” rather than predicting the real or biological age. The difference between the traditional real age estimation and apparent age estimation is that the age labels in apparent are annotated by human assessors rather than the real biological age. Some people may appear younger than their real age while others may appear older. As a result, the real age may differ from the apparent age of each subject.

Several methods have been proposed for apparent age estimation. The availability of huge data for training and an increase in computational power has made deep learning with convolutional neural network (CNN) a better method for the estimation task. Many researchers have studied several of these CNN methods, and these methods have improved the results and performances of apparent age estimation tasks. However, due to the challenging nature of apparent age estimation, further attempts to enhance the accuracy of the age estimation are still very much in progress. Researchers continue to examine different CNN and modern methods to enhance the results further. Hence, this paper critically reviews the modern approaches and techniques employed for apparent age estimation. We also present a comparative analysis of the performance of some of those approaches on standard apparent age datasets. The study also highlights the strengths and weaknesses of each approach to apparent age estimation to guide in choosing the appropriate approach to further improve the existing state-of-the-art results in the field. To ensure fairness in evaluating the performance of these approaches, we employed the popular apparent aging datasets and standard evaluation metrics that are widely used in literature in age estimation. Figure 1 displays the overall idea of a typical apparent age estimation system.

The contributions of this paper are highlighted as follows:

1. We outlined different state-of-the-art algorithms and techniques for apparent age estimation.
2. We described the performance evaluation analysis of different state-of-the-art models in apparent age estimation.
3. We presented three facial aging datasets widely employed in the research of apparent age estimation.
4. We also highlighted the standard performance evaluation metrics common in most literature for the apparent age estimation.

2. Application areas for apparent age estimation

Apparent age estimation has many notable real-world applications. Different intelligent application scenarios can benefit from computer-based systems that predict the apparent age of people among such application areas, including the following:

2.1. Medical diagnosis

Computer-based age prediction, as perceived by people, helps determine whether factors like the environment, depression, sickness, fatigue, and stress affect the premature aging of a person. This automatic age prediction will assist in obtaining the required information needed on a decision to improve the person's aging system (Escalera et al., 2015; Agustsson et al., 2017).

2.2. Effect of anti-aging treatment

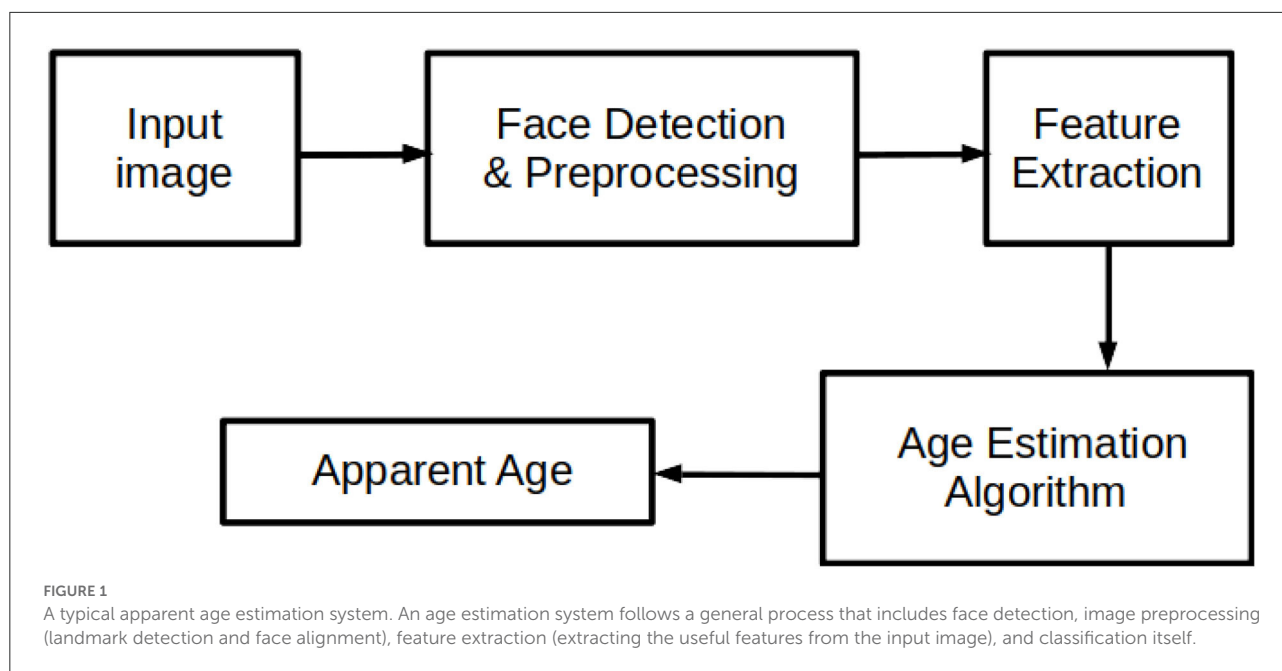
Automatic apparent age estimation is also valuable for knowing the effect of some anti-aging treatments on people. The effectiveness of these anti-aging treatments, like topical treatment and hormone replacement therapy, can only be understood if an apparent age estimator is in place (Escalera et al., 2015; Rothe et al., 2018).

2.3. Facial beauty product development

The effect of some cosmetics products on facial beauty product development can only be discovered with an accurate apparent age predictor. It helps to bring customer insight, marketing story, and aesthetic experience to their product. The estimator assists in determining the best element for the formula's dispensing to deliver in the future for desirables and best products (Padme and Desai, 2015; Rothe et al., 2018).

2.4. Effect of plastic surgery

The essence of plastic surgery procedures is to reshape and restore the appearance of a person's body. The surgery is connected with beautification ideas, which should involve an extensive range of practical operations, including craniofacial surgery, reconstructive surgery, etc. However, to know the impact of plastic surgery procedures, there is a need for an automated system that determines “how old a person is like?” (Fu et al., 2010; Voelkle et al., 2012).



2.5. Movie role casting

An apparent age estimator also plays a role in selecting roles cast in movies, television programs, music videos, stage plays, video documentaries, and television advertisements, among others. In choosing a particular type of an actor, actress, singer, or dancer, for a specific role and character, the need to determine the person's age as perceived by people will be necessary (Padme and Desai, 2015; Rothe et al., 2018).

3. Description of apparent age estimation algorithms

In this section, we present different algorithms and techniques used for apparent age estimation. As shown in Figure 2, most of these techniques fall into five different categories. Apparent age estimation can be modeled as a multi-class classification (MC), metric regression (MR), ranking, deep label distribution learning (DLDL), or a hybrid (combination of two or more techniques). We present a description of these algorithms and suggest the most effective approach in our opinion.

3.1. Multi-class classification

Multi-class classification approach views the ages or age groups category as an independent label and treats age value as a separate category and learns the age classifier to infer the

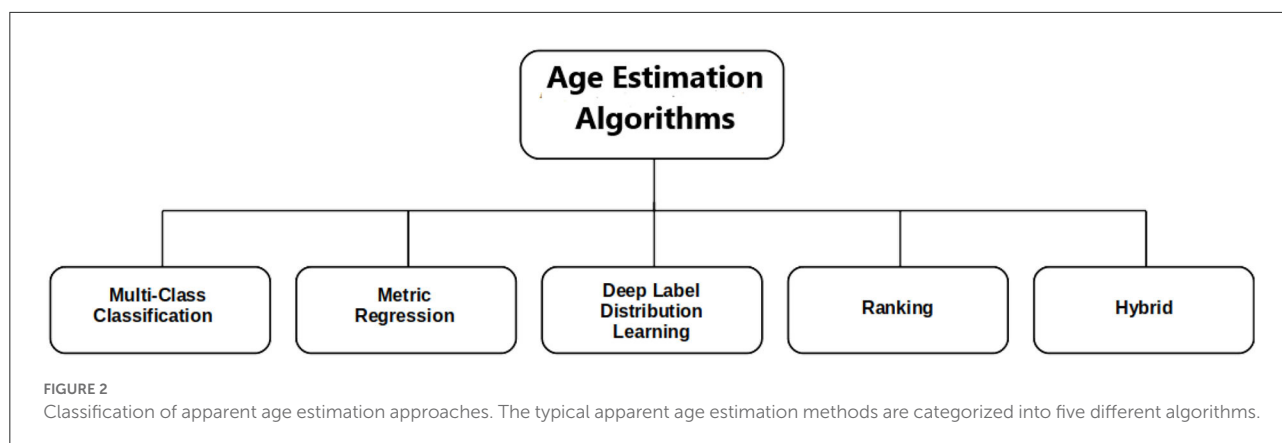
person's age (Zhu et al., 2015; Malli et al., 2016; Feng et al., 2017). MC algorithm maximizes the probability of ground-truth class labels by not considering other classes. Nevertheless, the limited training samples and the class imbalance among most facial aging datasets can result in an overfitting problem (Gao et al., 2018).

3.2. Metric regression

The metric regression-based algorithm views the age class as a linearly progressing relationship and does not display the diversity of the aging method. It learns the trait most appropriate for mapping the age-value space from the feature space using the appropriate regularization method. Although, it is quite normal to address the age estimation task as an MR problem, which does minimize the mean absolute error (MAE) result and improves the performance of the estimation accuracy. However, MR generates an unsteady training mode, causing a significant error term affecting accuracy performance. Some of the typical regression methods include Gaussian Process (Zhang and Yeung, 2010), quadratic regression (Lanitis et al., 2004), and support vector regression (Guo et al., 2008).

3.3. Deep label distribution learning

Deep label distribution learning approach converts real-value age to a discrete-age distribution to fit the entire age distribution. It is an end-to-end learning model that solves the



problem of insufficient training images experienced in most age estimation tasks. It relaxes the demand for a large number of training images and uneven data distribution by converting real age values to discrete age distribution to fit the whole age. The training instances connected with each class label will be increased without an increase in the number of training samples (Gao et al., 2018; Shen et al., 2018). However, it is usually observed that there is a lack of consistency between the employed evaluation metric and the training goals, generating an unsatisfactory result.

3.4. Ranking

The ranking-based algorithm uses age-axis tactics for age-class prediction and utilizes the relative order of the age. It uses relative age ranks instead of real age labels and ranks age class labels in descending order using their relevance to the presented face images to prevent making a decision for each age label that can simplify the problem (Chang et al., 2010; Li et al., 2012; Liu H. et al., 2017). Nonetheless, ranking algorithms can generate suboptimal results, especially when the training objectives and the evaluating metric are inconsistent.

3.5. Hybrid

The hybrid algorithm can be built by combining two or more algorithms in a parallel or hierarchical manner to produce a better performance. The algorithm makes the most of the advantage of the strengths of each algorithm to obtain a more robust system (Guo et al., 2008; Dib and El-saban, 2010; Choi et al., 2011). Unfortunately, combining two or more algorithms can result in a large storage overhead and computational costs, affecting its applicability in resource-constrained machines.

3.6. Summary of apparent age estimation algorithms

In this section, we summarized the main strengths and weaknesses of the different apparent age estimation algorithms in Table 1. Most of the existing state-of-the-art methods used MC and MR algorithms. The hybrid algorithm combines two or more algorithms, and this gives a better and more robust model compensating for the weak points in each algorithm with the strength of others. On the other hand, the ranking algorithm solves the problem peculiar to the classification algorithm by using the ordinal information of various ages and converting it into various binary classification problems. However, with DLDL, we obtained a better model by using the adjacent ages to generate label distribution for each age, even when the label distribution of the dataset is uneven. Table 2 presents the performance of state-of-the-art CNN architectures with a clear distinction of the best results on each dataset for apparent age estimation when evaluated using MAE and ϵ -error.

4. Performance evaluation analysis of state-of-the-art methods in apparent age estimation

Apparent age is “how old a person looks like”? A significant amount of study has been done to extract facial features from faces to determine the apparent age of people. In Rothe et al. (2015) developed a classification-based solution [deep expectation (DEX)] for apparent age. The authors used the VGG-16 architecture that was initially pre-trained on ImageNet before further fine-tuning the newly collected IMDB-WIKI dataset of 500,000 faces of unconstrained images. The CNN-based model was addressed as a deep classification problem. As part of the solution, they employed an open source face detector by Mathias et al. (2014) to locate the face in an image

TABLE 1 Description of state-of-the-art algorithms in apparent age estimation.

Algorithms	Description	Strengths	Weaknesses
Multi-class classification (MC)	A multi-class classification method considers the age value as an independent category and later learns a classifier for the age classification task; It neglects the internal relationships between those age values.	<ul style="list-style-type: none"> • MC maximizes the expectation of the ground-truth set without considering other classes. • It also presents age value as a separate category and later learns a classifier for age estimation. 	<ul style="list-style-type: none"> • MC algorithm can easily lead to over-fitting due to imbalances problem among the age classes and insufficient training images. • The method can also lead to unstable training.
Metric Regression (MR)	MR models the relationship between a set of features and a continuous target variable. It makes some predictions from data by learning the relationship between the features of the data and some continuous-valued answers.	<ul style="list-style-type: none"> • MR minimizes the MAE; The smaller the MAE, the better the performance of an estimator. 	<ul style="list-style-type: none"> • Some outliers in the input data can cause a large error term, leading to an unstable training process and producing an unsatisfactory performance. • MR presents the age category as a linearly growing dependence rather than displaying the diversity of the aging process.
Deep Label Distribution Learning (DLDL)	DLDL is an end-to-end learning model that solves the problem of insufficient training images experienced in most age estimation tasks. It relaxes the demand for a large number of training images and uneven data distribution by converting real age values to discrete age distribution to fit the whole age distribution.	<ul style="list-style-type: none"> • DLDL overcome the uneven age label distribution problem by converting real age value to discrete age distribution to fit the whole age distribution. • DLDL also ease the necessity for a large number of images during training. 	<ul style="list-style-type: none"> • DLDL may be suboptimal. • There might be an inconsistency during the training stage.
Ranking	Ranking uses the age-axis approach for age estimation. It employs the relative order of age to solve the classification bias problem caused by the dataset's unevenness of the dataset's sample images.	<ul style="list-style-type: none"> • The ranking model employs an age-axis approach that uses the relative order of age for the age classification. • The algorithm also transforms age estimation into a series of binary classification problems where the output of the rankers is aggregated directly from those binary outputs for the classification. 	<ul style="list-style-type: none"> • Ranking produces inconsistency in the training objectives and evaluation metric. • Ranking method may be suboptimal at times.
Hybrid	A hybrid algorithm combines two or more modeling methods in a parallel or hierarchical manner to produce a better performance. It makes the most of the advantage of the strengths of each technique employed and is exacted to not only defeat other individual approaches but also make it robust.	<ul style="list-style-type: none"> • A hybrid model makes the most of the advantage of the strengths of each technique used, and it is expected to outperform other individual approaches. • Hybrid also produces a robust classifier. 	<ul style="list-style-type: none"> • Combining two or more models might result in storage overhead and a huge computational cost. • It might be hard to deploy a hybrid model on resource-constrained devices.

before extracting the CNN predictions from an ensemble of 20 networks on the cropped face which was also fine-tuned on the LAP-2015 dataset. The developed model achieved a great result but demanded more computational cost and large

storage overhead to pre-trained the model on huge datasets like ImageNet and IMDB-WIKI.

Liu et al. (2015) later presented a hybrid model (AgeNet) that fuses a regression (real-value) and classification (Gaussian

TABLE 2 Description of state-of-the-art convolutional neural network (CNN) architectures in apparent age estimation.

References	Approach	Pre-trained models	Algorithm	Dataset	External data	ϵ -Score	MAE
Rothe et al. (2015)	20 CNN networks model	VGG	MC	LAP-2015	ImageNet; IMDb; WIKI;	0.278	3.221
Liu et al. (2015)	Real-value + Gaussian label distribution	GoogleNet	Hybrid	LAP-2015	CASIA-WebFace; CACD; WebFaceAge; MORPH-II;	0.2872	3.3345
Zhu et al. (2015)	Multiple models + RF + SVR	GoogLeNet	MR	LAP-2015	CASIA-WebFace; Adience; MORPH-II; FGNET; Lifespan; CACD; Private;	0.295	-
Ranjan et al. (2015)	DCNN-H-3NNR(Gaussian loss)	DCNN	MR	LAP-2015	CASIA-WebFace; Adience; MORPH-II;	0.373	-
Huo et al. (2016)	KL divergence + Softmax	VGG-16 + new CNN	DLDL	LAP-2015	MORPH-II; FG-Net; Adience; Web;	0.3057	-
Malli et al. (2016)	Ensemble of 3 CNNs	VGG-16	MC	LAP-2016	IMDb-WIKI;	0.3668	-
Antipov et al. (2016)	LDL + Classification	VGG-16	Hybrid	LAP-2016	IMDb-WIKI; Private;	0.241	-
Gurpinar et al. (2016)	VGG-Face + Kernel ELM	VGG-16	Hybrid	LAP-2016	-	0.3740	3.85
Liu W. et al. (2017)	GA-DFL + Multi-path CNN	VGG-16	MC	LAP-2015	-	0.369	4.21
Ranjan et al. (2017)	Euclidean + Gaussian loss functions	Novel CNN	MR	LAP-2015	MORPH-II; IMDb-WIKI; Adience;	0.293	-
Gao et al. (2017)	VGG-Face + DLDL(KL loss function)	ZF-Net + VGG-Net	DLDL	LAP-2015	-	0.31	3.51
Agustsson et al. (2017)	Residual DEX	VGG-16	MC	APPA-REAL	-	-	4.082
Gao et al. (2018)	LDL + Expectation Regression	ThinAgeNet	Hybrid	LAP-2015; LAP-2016;	-	0.272; 0.267	3.135; 3.452
Duan et al. (2018a)	CNN + ELM	AgeNet	Hybrid	LAP-2016	ImageNet; IMDb-WIKI; MORPH-II;	0.3679	-
Rothe et al. (2018)	DEX	VGG-16	MC	LAP-2015	ImageNet; IMDb-WIKI;	0.282	3.252
Li et al. (2019)	CNN + BridgeNet	VGGNet	MR	LAP-2015	-	0.26	-
Liu et al. (2019)	ODL (cross-entropy)	VGGNet	Ranking	LAP-2016	-	0.312	-
Deng et al. (2021)	CNN	VGG-16 + ResNet + GoogLeNet + AlexNet	Hybrid	LAP-2016	FGNET; MORPH	-	2.94; 2.97
Zhao et al. (2022)	adaptive mean residue loss	VGG-16 + ResNet50	DLDL	LAP-2016	FGNET	0.3882	3.61
Kjærran et al. (2021)	CNN	AgeNet	MC	APPA-REAL	UTK; IMDb	0.3882	3.61
Kjærran et al. (2021)	Gabor feature fusion + PCA + SVM + KPCA + CIAO-SA	Novel CNN	Hybrid	LAP-2016	Adience	-	2.10
Deng et al. (2021)	CNN	GenderNet; AgeNet; RaceNet	MR;Ranking	LAP-2016	MORPH2; FGNET	-	2.47; 2.59; 2.67

label distribution) to solve the apparent age estimation task. The two models employed GoogleNet CNN to learn informative age representations after preprocessing the images using face detection, facial landmark localization, and face normalization. The models were initially pre-trained on a large-scale facial aging dataset with identity labels and then fine-tuned on another large-scale age dataset with unconstrained age label before it was fine-tuned on the training images of the original LAP-2015 with apparent age labels. The hybridized model achieved a second place position in the 2015 edition of the Chalearn Looking At People (LAP) competition. However, the employed GoogleNet 22-layers deep convolution neural network was too deep to be implemented on resource-constrained devices.

Zhu et al. (2015) studied a method that utilized the deep representations trained in a cascaded way. The approach also employed GoogleNet design, initially pre-trained with face images without age labels, then on data with chronological age labels to fine-tune the network parameters before finally fine-tuning the apparent age model on the apparent age dataset itself. The proposed approach consists of four different processes: an image pre-processing stage (face detection and landmark localization), a CNN design architecture in a cascade way, a coarse-to-fine design consisting of age grouping and local age estimators, and fused predictors. Although the model achieved an ϵ -error of 0.2948 on the Chalearn LAP-2015 dataset with a third position at the 2015 edition of ChaLearn LAP challenge, it demands more computational cost and overhead to pre-train very large datasets on equally large CNN architecture.

In Ranjan et al. (2015), developed a regression-based automatic age estimation model. The approach estimates apparent age from unconstrained images using deep CNN (DCNN). The architecture consists of four steps: face detection, face alignment, feature extraction, and a 3-layer neural network regression. They employed a deep pyramid deformable part designed for the face detection phase and an ensemble of regression trees method (dlib C++ library) for the face alignment phase. However, for the feature extraction phase, they proposed a method that obtained traits from the pool5 layer of a pre-trained DCNN model without needing to re-tune the pre-trained DCNN network for face description tasks on age estimation data. For the age estimation phase, they used a 3-layer neural network regression model with the Gaussian loss function and a hierarchical learning approach to further boost the result of their work. Consequently, their approach achieved comparable results when evaluated on the LAP-2015 dataset. However, some outliers in the input data can cause a large error term, leading to an unstable training process.

Huo et al. (2016) introduced a deep CNN with distribution-based loss functions. The distributions utilized the ambiguity induced *via* manual labeling by learning a better model rather than using ages as the target. The method employed two different types of deep CNN models with different architectures: the first is based on the popular pre-trained VGG-16 CNN; the

second is based on a different CNN architecture. The VGG-16 architecture is fine-tuned on three different datasets before finally fine-tuning the two models on the competition dataset. The fusion of the two outputs generated the final predicted ages. The method achieved an ϵ -error of 0.3057 when evaluated on the LAP-2015 dataset. However, a distribution-based loss function might yield an inconsistent result during the training stage.

Malli et al. (2016) investigated an approach that used VGG-16 deep CNN models pre-trained on the IMDb-WIKI dataset and fine-tuned on the same dataset. The approach, an ensemble of deep learning methods, extracted deep features from the 7th layer (FC) of the VGG-FACE model and trained a 3-layer neural network with two hidden layers using deep features. They treated age estimation as a classification problem; as such, they assigned age labels within standard deviation boundaries as true. The approach significantly improved after fine-tuning the VGG-16 models for testing the age-shifted grouping technique. Although the designed fusion schemes with an ensemble of deep learning methods achieved a ϵ -error of 0.3668 on the test set of the LAP-2016 dataset when evaluated, it can lead to overfitting if an appropriate regularization algorithm is not utilized.

In Antipov et al. (2016), the authors proposed a pre-trained VGG-16 CNN model that combined two separate models: the general and the children. The general model was initially trained on the huge IMDb-WIKI dataset for biological age estimation and then fine-tuned for the apparent age estimation task. The children model used a pre-trained VGG-16 network and trained the children (private dataset of children between 0 and 12 years old) dataset before it was also fine-tuned on the original apparent age estimation dataset. The children's network was fine-tuned from the general network. The method involves using separate age encoding strategies for training the general and children networks; a label distribution encoding for the general network and 0/1 classification encoding for the children network. The hybridized approach combines the strength of the two algorithms hierarchically but employs a more lighter CNN architecture that allows its applicability on a resource-constrained device.

Gurpinar et al. (2016) proposed a two-level method for apparent age estimation of facial images. They classified samples into overlapping age groups, and within each age group, apparent age is estimated with local regressors before fusing the output for the final age estimation task. The method involved three phases: face alignment, feature extraction, and model learning. They used a deformable parts model (DPM) for the face detector and a pre-trained CNN for feature extraction from already aligned images before employing Kernel extreme learning machines for classification. The method's effectiveness was evaluated on the ChaLearn LAP 2016 dataset, and they reported an ϵ -error of 0.374 on the test set and MAE of 3.85.

Liu W. et al. (2017) proposed a group-aware deep feature learning (GA-DFL) technique for apparent age estimation. The CNN-based method acquired the needed feature descriptor

directly from raw pixels of the face images. The ordinal ages were split into sets of discrete collections to learn deep feature transformations. They also designed a multi-path CNN approach to combine the corresponding information. The experimental results proved that the approach produced an excellent performance when compared with state-of-the-art methods. It was evaluated on three known face aging datasets and obtained an MAE of 3.93 (FG-NET), 3.25 (MORPH-II), and 4.21 (LAP-2015) with an error of 0.369.

Ranjan et al. (2017) presented a novel multi-purpose CNN model that concurrently solves the problem of apparent age estimation, genders recognition, face detection, pose estimation, landmark localization, smile detection, face verification, and recognition from any unconstrained face image. The approach, an improvement of the work in Ranjan et al. (2015), was trained in an MTL framework that develops a synergy among several face-related tasks improving the performance of each of those tasks through learning robust features for the distinct tasks. They employed multiple tasks enabling the network to learn the correlations between data from many distributions in an efficient way. They used Chalearn LAP-2015 and FG-NET datasets to model the age classifier, and it achieved a comparable result when it was evaluated on the same datasets.

Gao et al. (2017) also proposed a DLDL approach, an end-to-end deep learning design that employed label ambiguity in both feature and classifier learning. The approach prevented the network from overfitting even when an inadequate training dataset was used. They converted the label of each image into a discrete label distribution and learned the label distribution by minimizing a Kullback-Leibler divergence between the ground-truth and predicted label distributions using deep ConvNets. This is necessary to resolve the problem of ambiguous information among labels. Extensive experimental results revealed that the proposed approach for apparent age estimation performed significantly better than state-of-the-art methods on the same task when evaluated on MORPH-II and LAP-2015 datasets.

Agustsson et al. (2017) proposed a model called Residual DEX, an enhancement of DEX (Rothe et al., 2015). The apparent age model was addressed as a classification problem considering the age value as an independent category. The idea of residual is that the error between the ground truth labels and the rough DEX estimation can be tackled with a unique model. The model learns a new regressor with the same CNN architecture in Rothe et al. (2015) to predict DEX residuals, and this significantly contributed to the performance of the new apparent age estimator.

The authors in Gao et al. (2018) designed a lightweight CNN architecture that collectively learned age distribution and regressed it. The CNN-based approach, ThinAgeNet, employed the compression rate of 0.5. The network was further trained on a quite small model with a compression rate of 0.25 and called it TinyAgeNet. The method combined LDL and

expectation regression into a unified structure to ease the disparity between the training and evaluation stages. The proposed approach efficiently enhanced the performance of the earlier DLDL on both prediction error and inference speed for age classification. The approach's effectiveness was validated for real and apparent age estimation tasks using MORPH-II, LAP-2015, and LAP-2016 datasets.

Duan et al. (2018a) developed a robust age estimator which employed an ensemble structure: CNN2ELM, which includes CNN and extreme learning machine (ELM). The model updated the work presented in Duan et al. (2018b) and has a three-level approach, including feature extraction, age grouping using an ELM classifier, and age estimation with an ELM regressor. The model initially pre-trained on the ImageNet dataset was fine-tuned on the IMDb-WIKI and MORPH-II datasets. The experimental analysis performed on the LAP-2016 dataset outperformed the existing state-of-the-art age estimation models and achieved an ϵ -error of 0.3679.

The deep expectation model based on VGG-16 architecture was proposed (Rothe et al., 2018). The approach is a solution to real and apparent age estimation from a single face image without the use of facial landmarks. The DEX model was pre-trained on both ImageNet and IMDb-WIKI datasets to achieve better performance. They evaluated their method on three standard datasets, MORPH-II, FG-NET, and LAP-2015, and obtained state-of-the-art results for both real and apparent age estimations. The approach achieved an MAE of 3.09 (FG-NET), 2.68 (MORPH-II), and 6.521 (CACD), an ϵ -error of 0.2650 (LAP-2015), and 64.0% (Exact), and 96.6% (1-off) accuracy on the Adience dataset outperformed the existing state-of-the-art age estimation methods. It recorded an ϵ -error of 0.3679 on the LAP-2016 dataset.

Li et al. (2019) proposed a CNN-based technique called BridgeNet for real and apparent age estimation. The proposed model comprises two components: local regressors and gating networks that can jointly be learned end-to-end. The first component (local regressors) addressed heterogeneous data by partitioning the data space. In contrast, the second one (gating networks) employed a bridge-tree structure that learns the continuity-aware weights used by the local regressors. Experimental results on the MORPH II, FG-NET, and Chalearn LAP-2015 datasets show the effectiveness of the BridgeNet CNN in outperforming the state-of-the-art methods.

Liu et al. (2019) developed a method that is an extension of their work in Liu et al. (2018). The work is an end-to-end ordinal deep learning (ODL) framework, including two ordinal regression loss functions; square loss and cross-entropy loss. The proposed ranking-based ordinal deep feature learning (ODFL) method learns features needed for face representation directly from raw image pixels and then independently learns the procedures of feature extraction and apparent age estimation. The work was evaluated on state-of-the-art face aging datasets

and achieved superior performance compared to state-of-the-art methods in apparent age estimation.

The work presented by [Dagher and Barbara \(2021\)](#) employed transfer learning from four pre-trained CNNs models to develop a facial age estimation model of human faces. The authors aimed to find the optimum age gap and to achieve high age estimation accuracy. The model was trained and evaluated on the FG-NET + MORPH datasets. These datasets were deemed suitable because it has varieties of age range from 0 to 77 years. The proposed model achieved an MAE of 2.94 and 2.97 on FGNET and MORPH datasets, respectively.

Recently, [Zhao et al. \(2022\)](#) proposed an adaptive mean-residue loss effective for facial age estimation. The proposed mean loss can penalize the age probabilities between the estimated age distribution's mean and the apparent age. Experiments for the model were performed on popular facial datasets (FGNET and CLAP2016) and evaluated using the MAE and ϵ -error to achieve 3.61 and 0.3882, respectively. The experimental results show superior performance on both datasets when applied to state-of-the-art models (VGG-16, ResNet-50) compared to the existing mean-variance loss methods. The authors conclude with some recommendations that can improve the performance of their model.

A deep learning model that comprises five convolutional layers and three fully-connected layers trained from scratch was developed and presented in [Kjærran et al. \(2021\)](#) for individual age estimation from facial images. The model was trained on three benchmarks (APPA, UTK, and IMDB) datasets and evaluated on separate held-out data and the Adience benchmark datasets. The experimental results show an inferior performance on the Adience dataset compared to existing models on the same dataset. In contrast, improved performance was obtained when the model was evaluated on the held-out dataset to achieve exact accuracy of 0.304% and one-off accuracy of 0.463%.

Deep learning models applied to facial age estimation tasks and the different data modalities employed in studying aging were presented in [Ashiqur Rahman et al. \(2021\)](#). The present study presents four broad classes of measures for quantifying algorithms' performances concerning biological age estimation. Based on the findings, the direction for the future endeavor in the apparent age estimation research was identified with significant potential for improvement in understanding the individual's health status with respect to body shape, blood samples, and physical activities.

A hybrid facial age estimation method using the Gabor feature fusion with an atomic search algorithm for feature selection was proposed by [Lu et al. \(2022\)](#). Gabor filter with five scales and eight directions was first used to extract facial age features and then employed a histogram to carry out coding and fusion for the indices in each direction of the Gabor. An algorithm called the chaotic improved atom search optimization with simulated annealing (CIASO-SA) was then presented to improve the accuracy and the number of feature selection that

is more adaptive to solving optimization problems in high dimensions. Experimental results found that Gabor achieved the best results on the 48 x 48 image resolutions to obtain one-off accuracy of 85.9%.

[Deng et al. \(2021\)](#) proposed a multifeatured learning and fusion method for age estimation. Three subnetworks were employed to learn age, gender, and race information. The race and gender information and the age features were concatenated to form a robust feature extraction for age estimation. These features were then converted into exact age using the regression ranking age feature estimator. Three popular benchmark datasets (MORPH2, FGNET, and LAP) were used to experiment and validate the performance efficiency of the model. The proposed model achieved an MAE of 2.47%, 2.59%, and 2.67%, respectively, for the three datasets, compared to existing methods. The model is also suitable for mobile device deployment for age estimation due to memory compactness of 20 MB only.

A gender-specific facial age estimation system was proposed by [Raman et al. \(2022\)](#) to classify gender images and estimate age. The system is composed of two separate models. One of the components is meant to classify the facial images into males and females separately. The other component is made up of models that were trained independently on the male and female images. The models were trained on and evaluated on the UTK-Face dataset, while cross-validation was performed with the FGNET dataset. The experimental results report 90.86% of accuracy on the female-specific model and 89.21% of accuracy for the male-specific model.

5. Apparent age estimation datasets

We briefly highlighted three popular facial aging datasets that were widely employed in the research of apparent age estimation; Chalearn LAP-2015, Chalearn LAP-2016, and APPA-REAL datasets.

ChaLearn LAP-2015 dataset. [Escalera et al. \(2015\)](#) was collected for the purpose of ChaLearn LAP challenge competition 2015 edition. It is the first dataset on apparent age estimation. The dataset comprised 4,699 images, with each image labeled by at least ten different users. LAP-2015 dataset is divided into three sets: 2,476 (training), 1,136 (validation), and 1,087 (testing).

ChaLearn LAP-2016 dataset. [Escalera et al. \(2016\)](#), an extension of the LAP-2015 dataset, consists of 7,591 face images collectively labeled by different human annotators. The dataset is divided into three sets: 4,113 training, 1,500 validation, and 1,978 testing images. The testing set labels are separated from the other sets but with similar age distribution.

APPA-REAL dataset. [Agustsson et al. \(2017\)](#) is the first state-of-the-art database with both real and apparent age annotations. The images are collected using a labeling

application, crowd-sourcing data collection, data from the AgeGuess platform, and with the assistance of Amazon Mechanical Turk (AMT) workers. APPA-REAL database contains a total of 7,591 images with an age range between 0 and 95 of images of subjects taken under different conditions.

The sample images from these datasets are presented in Figure 3, while sample size, subjects size, and the age range information of those datasets are presented in Table 3. Figure 4 shows facial aging datasets by the number of publications, while Figure 5 shows the performance in terms of ϵ -error achieved by authors on the Chalearn LAP 2015 dataset.

6. Performance evaluation metrics for the state-of-the-art methods in apparent age estimation

The standard evaluating metrics used for facial age estimation are MAE (Onifade and Akinyemi, 2014), cumulative score (CS) (Lin et al., 2012), accuracy (exact and 1-off) (Levi and Hassner, 2015), and normal score (ϵ -error) (Duan et al., 2018a). MAE, CS, and accuracy are commonly employed for real age estimation, while MAE and ϵ -error are the common evaluation metrics for apparent age estimation. MAE is defined as the average of the absolute errors between the estimated ages and the ground truth, and CS measures the performance of an estimator when the training data has images at nearly every age; exact accuracy calculates the percentage of face images that were classified into correct age and gender; the ratio of the accurate predictions to the total number of the ground-truth label while one-off accuracy measures whether the ground-truth class label matches the predicted class label or if the ground-truth label exists in the two adjacent bins.

Mean Absolute Error is defined mathematically by:

$$MAE = \sum_{k=1}^N \frac{|I_k - I_k^*|}{N} \tag{1}$$

where

I_k : the estimated age.

I_k^* : the ground truth age for the test image k .

N : the total number of test images.

Normal (ϵ)-error metric was employed for the purpose of the Chalearn LAP competition in 2015. It calculates the error between an estimated value and the average labeled age. ϵ -error is calculated as follows:

$$\epsilon = 1 - e^{-\frac{(x-\sigma)^2}{2\mu^2}} \tag{2}$$

where

x : the estimated age.

σ : apparent age label provided for a given face image.

μ : standard deviation of all estimated ages for the given face image.

ϵ -error not only calculates the error between the estimated value x and the averaging labeled age σ but also considers the standard deviation of μ . The final ϵ -error result is the average overall prediction, and obviously, the lower the ϵ -error, the better the estimator's performance.

Figures 6, 7 present the MAE and the ϵ -error performance reported by authors on the Chalearn LAP 2015 and Chalearn LAP 2016 datasets.

7. Discussion

In recent times, a study in apparent age estimation has progressed steadily in terms of performance. From the literature presented above, it was clear that no concise conclusion can be made on the best type of algorithm for the estimation task. In Table 1, we present a description of each of these

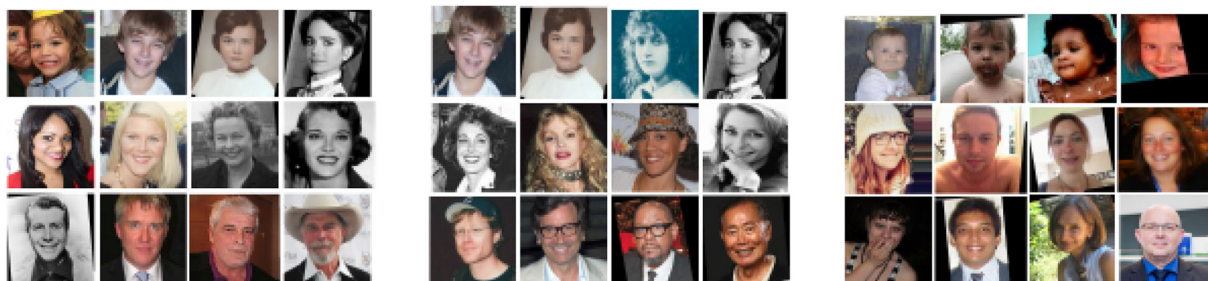
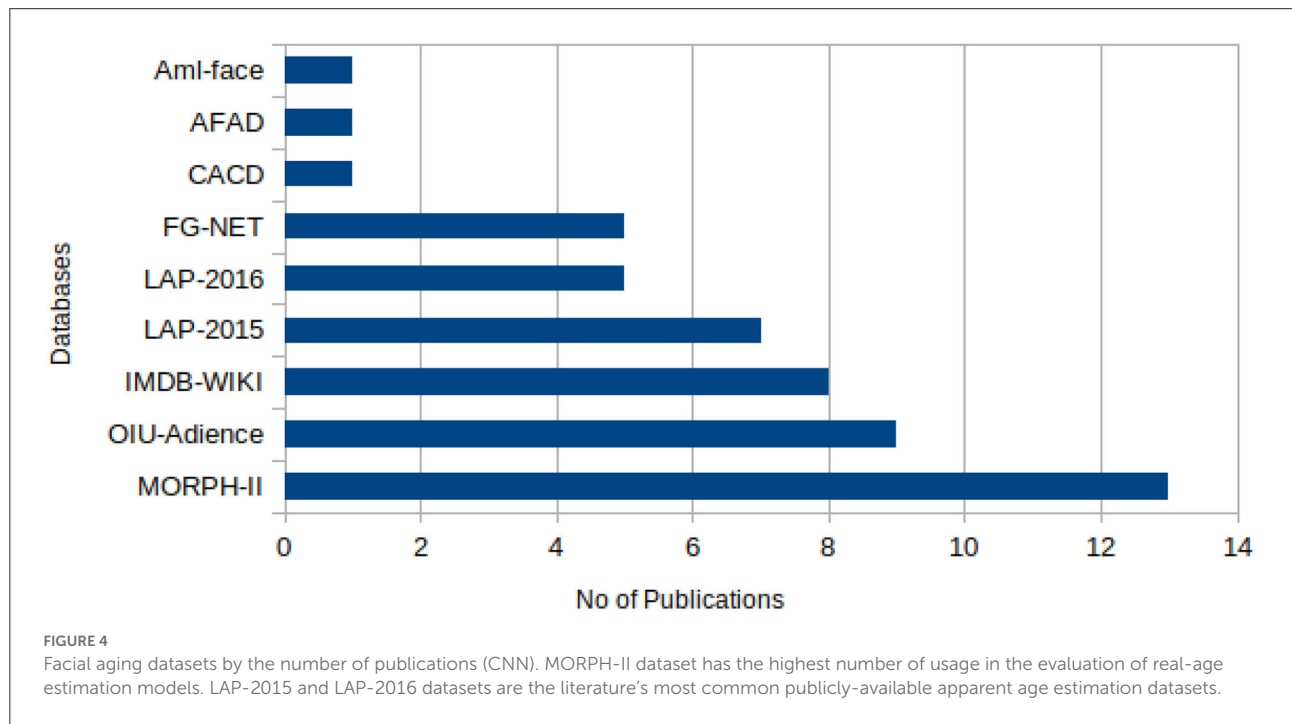


FIGURE 3

Sample images from LAP-2015, LAP-2016, and APPA-REAL datasets. LAP-2015 and LAP-2016 datasets were collected purposely for the Chalearn looking at people (LAP) competition in 2015 and 2016, respectively. It contains facial images with apparent ages. APPA-REAL dataset, on the other hand, was introduced by Agustsson et al. (2017) and contained images with both real and apparent age annotations.

TABLE 3 A summary of apparent age estimation databases.

Database	#Faces	#Subj.	Range	Age type	Year	In-the-wild?
ChaLearn (Escalera et al., 2015)	4,699	-	0–100	Apparent	2015	yes
ChaLearn (Escalera et al., 2016)	7,591	-	-	Apparent	2016	yes
APPA-REAL (Agustsson et al., 2017)	7,591	7,000+	0–95	Real and Apparent	2017	yes



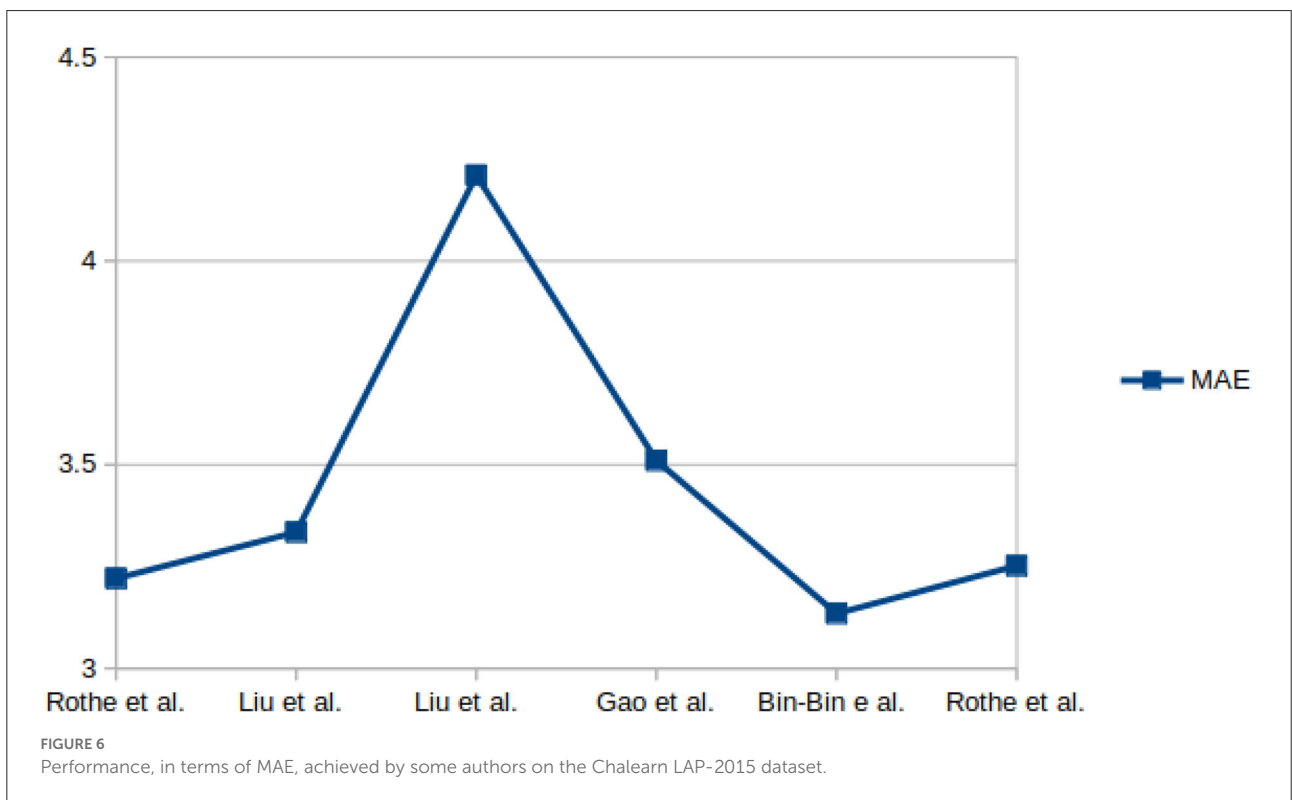
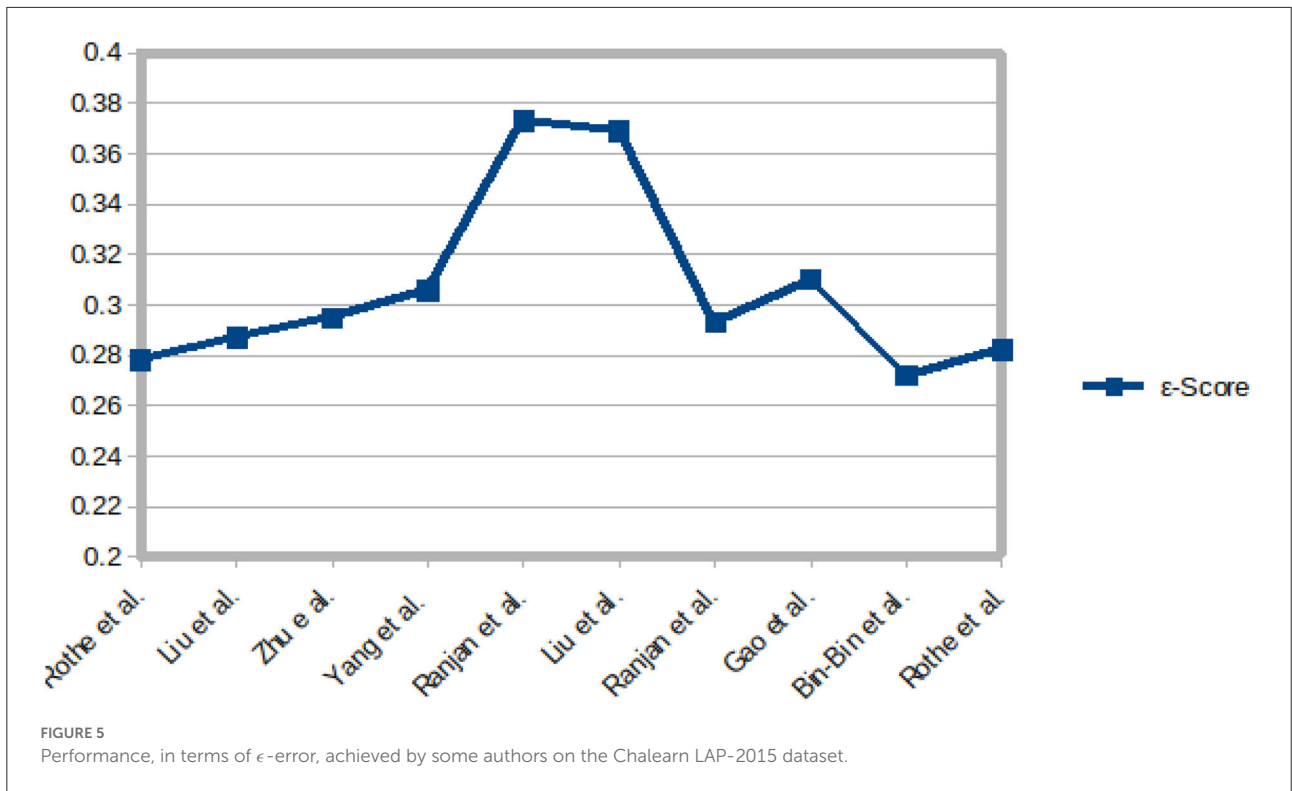
algorithms employed by researchers, highlighting their strengths and weaknesses; this will help to choose the right algorithm for future work.

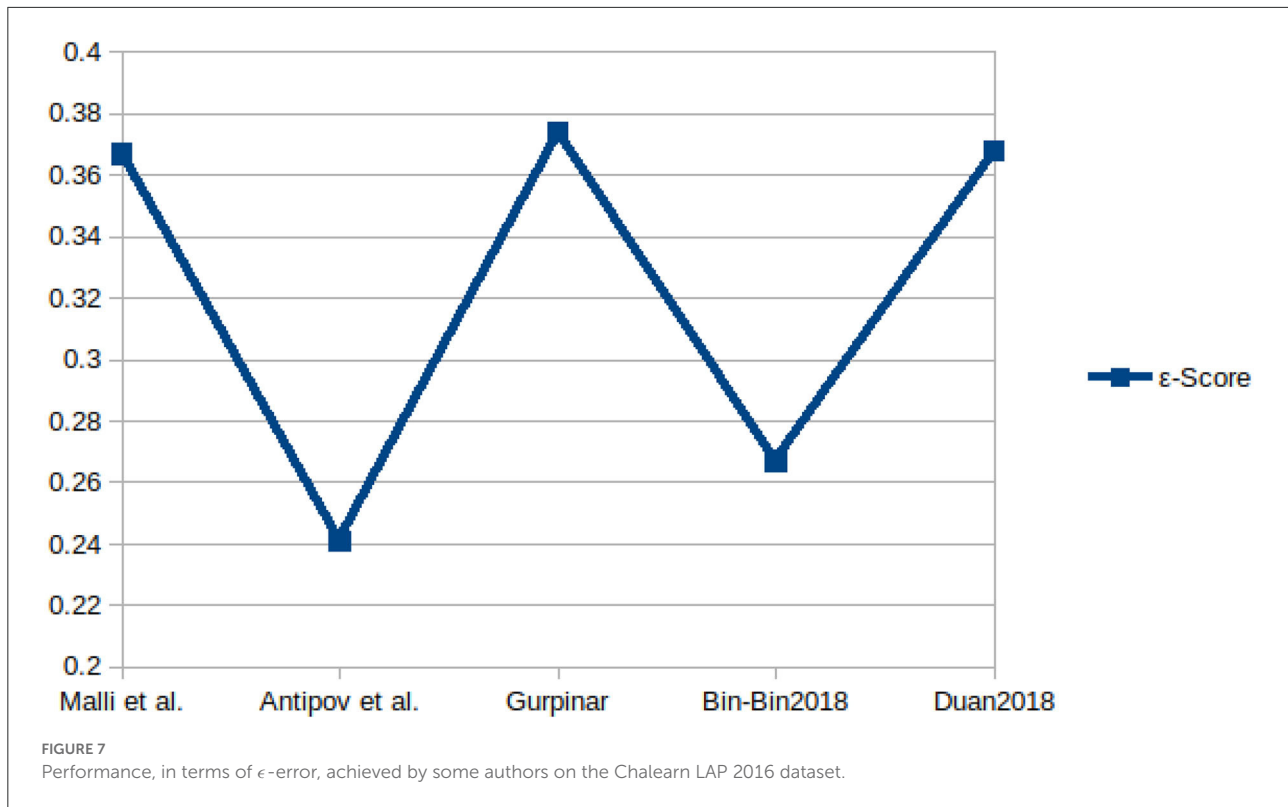
As presented in Figure 4, it was observed that the MORPH-II dataset had been far more used for real age estimation experiments than any other datasets, and this is probably due to the more significant number of facial images, which helps algorithms to learn the many individual aging traits better. However, LAP-2015, LAP-2016, and APPA-REAL are the three publicly-available facial aging datasets with apparent age labels for apparent age estimation. In order to avoid biased, considering the peculiarities of each dataset, our analysis will be based on these three datasets. As such, we summarized in Table 2, the performance of state-of-the-art algorithms showing a clear distinction between which result is best so far on each of these datasets for apparent age estimation when evaluated using MAE and ϵ -error metrics.

Consequently, significant observations from earlier research work in the estimation help make some

justifiable conclusions. Hence, we highlighted those observations stating some of the points which we deem very important:

- It was observed that the image processing method employed for face detection, facial landmark, and face alignment has an impact on the performance of an apparent age estimator.
- It was also important to assert that the performance of learning algorithms was determined by many factors, among which are the size and label distribution of the employed dataset, the degree of image variability, etc. The deep learning algorithms perform differently on different datasets, most likely due to the peculiarities of each dataset.
- Data augmentation improves the performance of age estimation models, especially on unevenly distributed not-too-large datasets.
- We also observed from the literature that models pre-trained on large-scale datasets before fine-tuning on the





original dataset performed better than training the model on just the original dataset.

- From this review, we observed that MC algorithms had been the most popularly used individual algorithm for apparent age estimation on the mentioned datasets.
- We also observed that ranking and DLDL are the most suitable algorithms for the estimation when the label distribution of the dataset is uneven.

For apparent age estimation on LAP-2015, a hybrid algorithm (LDL and expectation regression) showed the best performance on MAE and ϵ -error metrics (Figures 5, 6). However, on the LAP-2016 dataset (in Figure 7), a hybridized algorithm combining label distribution and classification algorithm also presented the best performance. The choice of any of these algorithms continues to be debated in the facial aging research community, which is evident in Table 2. This demonstrates that no approach may be individually suitable for apparent age estimation; instead, an MC, DLDL, or hybrid (the combination of DLDL with another algorithm) seem more effective and consistent.

8. Conclusion and future directions

This work presents a comprehensive review and the suitability of modern algorithms for apparent age estimation focusing on those algorithms that are most popular and those that appear to have been the most successful. In this review, we based our evaluations on LAP- 2015, LAP-2016, and APPAREAL, the most popularly-used publicly available facial aging datasets for apparent age estimation. Apparent age estimation, estimate “how old the person looks like”? A hybridized algorithm showed the best performance on LAP-2015 and LAP-2016 datasets. From this study, we could deduce that MC was the most popularly used individual algorithm. We also asserted that the performance of these age estimation algorithms is mostly influenced, not only by choice of the approach but also by some other factors, among which are the method of image pre-processing applied, size and the distributions of the employed datasets, etc.

However, many encouraging future directions in apparent age estimations may improve performance. Huge datasets with apparent age label annotation rather than real age will help improve the research accuracy of apparent age estimation. Also, it is necessary to investigate further study that focuses more on predicting apparent age (how old does the person look?)

rather than a person's biological age to enhance the practical and real-world applications of this research.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships

that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Abbas, A. R., and Kareem, A. R. (2018). Intelligent age estimation from facial images using machine learning techniques. *Iraqi J. Sci.* 59, 724–732. doi: 10.24996/ij.s.2018.59.2A.10
- Agustsson, E., Timofte, R., Escalera, S., Baro, X., Guyon, I., and Rothe, R. (2017). "Apparent and real age estimation in still images with deep residual regressors on appa-real database," in *Proceedings - 12th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2017 - 1st International Workshop on Adaptive Shot Learning for Gesture Understanding and Production, ASLAGUP 2017* (Washington, DC: IEEE), 87–94.
- Antipov, G., Baccouche, M., Berrani, S. A., and Dugelay, J. L. (2016). "Apparent age estimation from face images combining general and children-specialized deep learning models," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* (Las Vegas, NV: IEEE), 801–809.
- Ashiqur Rahman, S., Giacobbi, P., Pyles, L., Mullett, C., Doretto, G., and Adjeroh, D. A. (2021). Deep learning for biological age estimation. *Brief. Bioinform.* 22, 1767–1781. doi: 10.1093/bib/bbaa021
- Bouchrika, I., Harrati, N., Ladjailia, A., and Khedairia, S. (2016). "Age estimation from facial images based on hierarchical feature selection," in *16th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering, STA 2015* (Monastir), 393–397.
- Chang, K. Y., Chen, C. S., and Hung, Y. P. (2010). "A ranking approach for human age estimation based on face images," in *Proceedings-International Conference on Pattern Recognition* (Istanbul), 3396–3399.
- Choi, S. E., Lee, Y. J., Lee, S. J., Park, K. R., and Kim, J. (2011). Age estimation using a hierarchical classifier based on global and local facial features. *Pattern Recognit.* 44, 1262–1281. doi: 10.1016/j.patcog.2010.12.005
- Dagher, I., and Barbara, D. (2021). Facial age estimation using pre-trained cnn and transfer learning. *Multimed. Tools Appl.* 80, 20369–20380. doi: 10.1007/s11042-021-10739-w
- Deng, Y., Teng, S., Fei, L., Zhang, W., and Rida, I. (2021). A multifeature learning and fusion network for facial age estimation. *Sensors* 21, 4597. doi: 10.3390/s21134597
- Dib, M. Y. E., and El-saban, M. (2010). *Human Age Estimation Using Enhanced Bio-Inspired Features (EBIF)*. Cairo: Faculty of Computers and Information, Cairo University.
- Drobnyh, K. A., and Polovinkin, A. N. (2017). Using supervised deep learning for human age estimation problem. *Int. Arch. Photogram. Remote Sens. Spatial Inf. Sci.* XLII-2/W4, 97–100. doi: 10.5194/isprs-archives-XLII-2-W4-97-2017
- Duan, M., Li, K., and Li, K. (2018a). An ensemble CNN2ELM for age estimation. *IEEE Trans. Inf. Forensics Security* 13, 758–772. doi: 10.1109/TIFS.2017.2766583
- Duan, M., Li, K., Yang, C., and Li, K. (2018b). A hybrid deep learning CNN-ELM for age and gender classification. *Neurocomputing* 275, 448–461. doi: 10.1016/j.neucom.2017.08.062
- Escalera, S., Fabian, J., Pardo, P., Baro, X., Gonzalez, J., Escalante, H. J., et al. (2015). "ChaLearn looking at people 2015: apparent age and cultural event recognition datasets and results," in *Proceedings of the IEEE International Conference on Computer Vision, Vol. 2015* (Santiago: IEEE), 243–251.
- Escalera, S., Torres, M. T., Martinez, B., Baro, X., Escalante, H. J., Guyon, I., et al. (2016). "ChaLearn looking at people and faces of the world: face analysis workshop and challenge 2016," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* (Las Vegas, NV: IEEE), 706–713.
- Feng, S., Lang, C., Feng, J., Wang, T., and Luo, J. (2017). Human facial age estimation by cost-sensitive label ranking and trace norm regularization. *IEEE Trans. Multimedia* 19, 136–148. doi: 10.1109/TMM.2016.2608786
- Fu, Y., Guo, G., and Huang, T. S. (2010). Age synthesis and estimation via faces: a survey. *IEEE Trans. Pattern Anal. Mach. Intell.* 32, 1955–1976. doi: 10.1109/TPAMI.2010.36
- Gao, B. B., Xing, C., Xie, C. W., Wu, J., and Geng, X. (2017). Deep label distribution learning with label ambiguity. *IEEE Trans. Image Process.* 26, 2825–2838. doi: 10.1109/TIP.2017.2689998
- Gao, B. B., Zhou, H. Y., Wu, J., and Geng, X. (2018). "Age estimation using expectation of label distribution learning," in *IJCAI International Joint Conference on Artificial Intelligence* (Stockholm), 712–718.
- Guo, G., Fu, Y., Dyer, C. R., and Huang, T. S. (2008). Image-based human age estimation by manifold learning and locally adjusted robust regression. *IEEE Trans. Image Process.* 17, 1178–1188. doi: 10.1109/TIP.2008.924280
- Gurpinar, F., Kaya, H., Dibeklioglu, H., and Salah, A. A. (2016). "Kernel ELM and CNN based facial age estimation," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* (Las Vegas, NV: IEEE), 785–791.
- Huerta, I., Fernández, C., Segura, C., Hernando, J., and Prati, A. (2015). A deep analysis on age estimation. *Pattern Recognit. Lett.* 68, 239–249. doi: 10.1016/j.patrec.2015.06.006
- Huo, Z., Yang, X., Xing, C., Zhou, Y., Hou, P., Lv, J., et al. (2016). "Deep age distribution learning for apparent age estimation," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* (Las Vegas, NV: IEEE), 722–729.
- Kjærø, A., Vennerød, C. B., and Bugge, E. S. (2021). Facial age estimation using convolutional neural networks. *arXiv preprint arXiv:2105.06746*. doi: 10.48550/arXiv.2105.06746
- Lanitis, A., Draganova, C., and Christodoulou, C. (2004). Comparing different classifiers for automatic age estimation. *IEEE Trans. Syst. Man Cybern. B* 34, 621–628. doi: 10.1109/TSMCB.2003.817091
- Levi, G., and Hassner, T. (2015). "Age and gender classification using convolutional neural networks," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* (Boston, MA: IEEE), 34–42.
- Li, C., Liu, Q., Liu, J., and Lu, H. (2012). "Learning ordinal discriminative features for age estimation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Providence, RI: IEEE), 2570–2577.

- Li, W., Lu, J., Feng, J., Xu, C., Zhou, J., and Tian, Q. (2019). "Bridgenet: a continuity-aware probabilistic network for age estimation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Long Beach, CA: IEEE), 1145–1154.
- Lin, C. T., Li, D. L., Lai, J. H., Han, M. F., and Chang, J. Y. (2012). Automatic age estimation system for face images. *Int. J. Adv. Robot. Syst.* 9, 1–9. doi: 10.5772/52862
- Liu, H., Lu, J., Feng, J., and Zhou, J. (2017). "Ordinal deep feature learning for facial age estimation," in *Proceedings-12th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2017* (Washington, DC: IEEE), 157–164.
- Liu, H., Lu, J., Feng, J., and Zhou, J. (2019). Ordinal deep learning for facial age estimation. *IEEE Trans. Circ. Syst. Video Technol.* 29, 486–501. doi: 10.1109/TCSVT.2017.2782709
- Liu, W., Chen, L., and Chen, Y. (2018). Age classification using convolutional neural networks with the multi-class focal loss. *IOP Conf. Ser. Mater. Sci. Eng.* 428, 12043. doi: 10.1088/1757-899X/428/1/012043
- Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., and Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing* 234, 11–26. doi: 10.1016/j.neucom.2016.12.038
- Liu, X., Li, S., Kan, M., Zhang, J., Wu, S., Liu, W., et al. (2015). "AgeNet: deeply learned regressor and classifier for robust apparent age estimation," in *Proceedings of the IEEE International Conference on Computer Vision* (Santiago: IEEE), 258–266.
- Lu, D., Wang, D., Zhang, K., and Zeng, X. (2022). "Age estimation from facial images based on gabor feature fusion and the ciao-sa algorithm," in *CAAI Transactions on Intelligence Technology*. doi: 10.1049/cit2.12084
- Malli, R. C., Aygun, M., and Ekenel, H. K. (2016). "Apparent age estimation using ensemble of deep learning models," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* (Las Vegas, NV: IEEE), 714–721.
- Mathias, M., Benenson, R., Pedersoli, M., and Van Gool, L. (2014). Face detection without bells and whistles. *Lecture Notes Comput. Sci.* 8692, 720–735. doi: 10.1007/978-3-319-10593-2_47
- Onifade, O. F. (2015). A review on the suitability of machine learning approaches to facial age estimation. *Int. J. Modern Educ. Comput. Sci.* 7, 17–28. doi: 10.5815/ijmecs.2015.12.03
- Onifade, O. F. W., and Akinyemi, J. D. (2014). A GW ranking approach for facial age estimation. *Egypt. Comput. Sci. J.* 38, 63–74.
- Padme, S. E., and Desai, P. S. (2015). Estimation of age from face images. *Int. J. Sci. Res.* 4, 1927–1931. doi: 10.21275/v4i12.NOV152411
- Raman, V., ELKarazle, K., and Then, P. (2022). Gender-specific facial age group classification using deep learning. *Intell. Autom. Soft Comput.* 34, 105–118. doi: 10.32604/iasc.2022.025608
- Ranjan, R., Sankaranarayanan, S., Castillo, C. D., and Chellappa, R. (2017). "An all-in-one convolutional neural network for face analysis," in *Proceedings-12th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2017-1st International Workshop on Adaptive Shot Learning for Gesture Understanding and Production, ASLAGUP 2017* (Washington, DC: IEEE), 17–24.
- Ranjan, R., Zhou, S., Chen, J. C., Kumar, A., Alavi, A., Patel, V. M., et al. (2015). "Unconstrained age estimation with deep convolutional neural networks," in *Proceedings of the IEEE International Conference on Computer Vision* (Santiago: IEEE), 351–359.
- Rothe, R., Timofte, R., and Van Gool, L. (2015). "DEX: deep expectation of apparent age from a single image," in *Proceedings of the IEEE International Conference on Computer Vision, Vol. 2015* (Santiago: IEEE), 252–257.
- Rothe, R., Timofte, R., and Van Gool, L. (2018). Deep expectation of real and apparent age from a single image without facial landmarks. *Int. J. Comput. Vis.* 126, 144–157. doi: 10.1007/s11263-016-0940-3
- Ruiz-Del-Solar, J., Verschae, R., and Correa, M. (2009). Recognition of faces in unconstrained environments: a comparative study. *EURASIP J. Adv. Signal Process.* 2009, 1–5. doi: 10.1155/2009/184617
- Shen, W., Guo, Y., Wang, Y., Zhao, K., Wang, B., and Yuille, A. (2018). "Deep regression forests for age estimation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Salt Lake City, UT: IEEE), 2304–2313.
- Voelkle, M. C., Ebner, N. C., Lindenberger, U., and Riediger, M. (2012). Let me guess how old you are: effects of age, gender, and facial expression on perceptions of age. *Psychol. Aging* 27, 265–277. doi: 10.1037/a0025065
- Zhang, Y., and Yeung, D. Y. (2010). "Multi-task warped Gaussian process for personalized age estimation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (San Francisco, CA: IEEE), 2622–2629.
- Zhao, Z., Qian, P., Hou, Y., and Zeng, Z. (2022). Adaptive mean-residue loss for robust facial age estimation. *arXiv preprint arXiv:2203.17156*. doi: 10.1109/ICME52920.2022.9859703
- Zhu, Y., Li, Y., Mu, G., and Guo, G. (2015). "A study on apparent age estimation," in *Proceedings of the IEEE International Conference on Computer Vision* (Santiago: IEEE), 267–273.