

Face Age Estimation using Shortcut Identity Connection of Convolutional Neural Network

Shohel Pramanik, Dr. Hadi Affendy Bin Dahlan
Information Science and Technology
University Kebangsaan Malaysia
Selangor, Malaysia

Abstract—Depletion of skin and muscle tone has a considerable impact on the appearance of the face, which is constantly evolving. Algorithms necessitate a large number of aging faces for this purpose. Another popular deep learning technique is convolutional neural networks. In a recent study, many computer vision and pattern recognition problems have been successfully tackled using it. But these methods have architectural issues (e.g., the training process) that have a negative impact on their age estimation performance. As a result, a whole new approach is proposed in this research to address the issue. Using a convolutional neural network framework and resnet50 architecture, researchers were able to detect the age of a human face. This proposed shortcut identity connection strategy, which enables age estimation from the face image, has improved the success of the resnet50 architecture. To be able to tell a person's age from a picture of their face, it was important to know the characteristics of aging. As a result, the rhetorical classification method, which employs the resnet50 structure, is used to shift the face aging levels to a probability level. All of the 50 layers in the proposed residual network have a residual block that connects them. As far as face-aging databases go, ImageNet and FG-NET are both good choices for the proposed age estimation process. In the training session, the experiment results are 2.27 and 2.38, based mostly on the mean absolute error. The test accuracy results for the ImageNet dataset are 81.75% with the FG-NET dataset and 57% with the ImageNet dataset.

Keywords—Shortcut identity connection; ImageNet; residual connection; face aging

I. INTRODUCTION

There are numerous real-world applications for age estimation from facial photographs, including security monitoring, missing-person inquiries, entertainment, crime investigation, forensic art, cosmetology, biometrics, and facial recognition. Because of the rising interest in this topic among scientists, the automatic age estimate has attracted a large number of them. It is possible to define an automatic age estimate as either a predefined age range or an actual age.

Multiple ways have been investigated to depict age using facial photos to estimate it. Unfortunately, most of the techniques require a significant number of facial aging datasets and an excessive level of architectural fitness to be implemented. Due to their inaccuracies, these approaches take a long time to train and also lack robustness [12]. To train deep neural networks, researchers have turned to the ResNet framework, which was released last year. The method can

employ shortcut identity connections to avoid overfitting in this network's numerous layers. When used for image classification, ResNet is capable of achieving good performance in prediction [1].

The Shortcut Identity Connection of ResNet50, as shown in Fig. 1, is a convolutional neural network is used to estimate an individual's age. The age estimation problem may be solved using a minimal face aging dataset with the proposed approach. Overfitting architectural designs have been solved through shortcut identity connections. In addition, the input data size can be matched to the design using this technique. Different images are trained and tested on various datasets to establish a shortcut identity connection in the ResNet50 architecture. As a final step, the algorithm used the testing dataset and a live image taken by the webcam to estimate age using an estimator. Detailed explanations of the algorithm can be found in the third section.

When deep networks were seen to be converging, a degradation issue arises [1]. As a result of this, the gradients between neurons were completely wiped out [2]. No additional layers or overfitting will solve this gradient problem [1]. Active functions like the sigmoid active function and the tanh active function [3] are to blame for this.

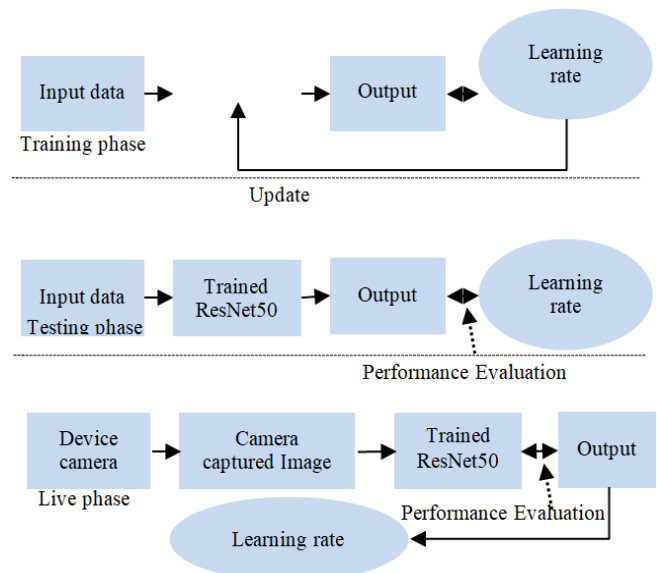


Fig. 1. Proposed Algorithm.

The following are the main points of this manuscript: The proposed shortcut identity connection method demonstrates the high capability of resnet50 by freezing layers from the structure; the presented method can accurately estimate the ages of global and regional people. The proposed age estimation process is immediate and uses an image taken directly by the device webcam.

II. RELATED WORK

A. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) is a feed-forward artificial neural network approach [34]. Individual neurons are structured such that they can respond to the overlapping areas of the visual field. In 1968, Hubel and Wiesel [4] studied the cat's visual cortex and discovered that the cortex had a complex arrangement of neurons. In order to cover as much of the subregional visual field as possible, these neurons are placed in a grid-like pattern. It is these neurons that perform the filtering of the input image and work with respect to the natural face photo. Each layer of the CNN design uses a differential function to turn the input volume into the output volume. The architecture of CNN is divided into several separate tiers. Fig. 2 displays the CNN layers.

The work of CNN layers [13] has given below:

The convolution layer serves as the foundation upon which CNN is built, and it is the first step in the learning process. CNN parameters can calculate by multiplying each filter's width and height (w and h), the preceding layer's filter (F), and the current layer's filters (n). $P = (w \times h \times f) \times n$.

During the training session, filters and parameters will learn. While processing the image classification, each filter generates an activation map. In the pooling layer, there are no parameters. The CNN pooling layer does not require any kind of learning. To avoid overfitting during training sessions, it is also beneficial to use the pooling layer. So, the pooling layer p parameter is zero.

A Fully Connected Layer is made up of feed-forward neural networks. The "Fully Connected Layer" is the network's final layer and is referred to as such because of its connectivity. An input is flattened from the final Pooling or Convolutional Layer. A fully connected layer has the most factors to consider when comparing it to other layers. The ideal parameters for a fully linked layer may be found easily because each neuron is connected to every other neuron in the system.

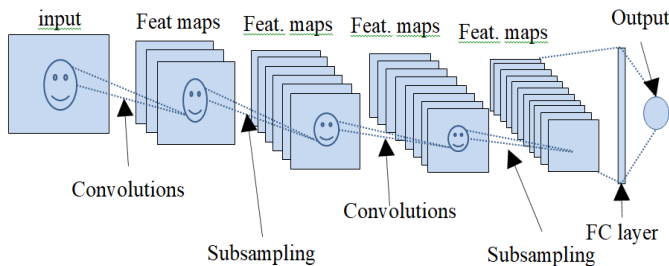


Fig. 2. Neural Networks are constructed using a Common Method known as Convolution. There are Several Separate Layers in CNN Topologies, Including an Input Layer, Convolution Layers, and a Fully Connected Layer.

The generated image is presented in the figure above in the top right corner (Fig. 2), is known as the "Feature Map" [31]. It's the number of confusing features. As a result, the input image information is reduced. When the number of confusing features is 2, the algorithm reduces the image. If the number of confusing features is 4, the algorithm further reduces the image, which makes it easier to process. The question is whether the filter function detector causes information loss. The higher the number in the feature map, the better the filtering process, which implies that the algorithm is not missing many features. Essentially, researchers create all of the feature maps that are required, as well as all of the filters that are required (e.g., edge detection, blur detection, bump detection). Using the Softmax function, the method can assure that all of the estimation will add up to 1. The cross-entropy function and the Softmax function are frequently used interchangeably [32]. After applying the Softmax function in CNN, the next step is to test the reliability of the model using the cross-entropy function to maximize the performance of the neural network [35]. Using the cross-entropy function has several advantages. One of the best features is that if the output value is much smaller than the actual value at the start of the back-propagation, the gradient descent will be very slow. Since Cross-Entropy uses the logarithm, it helps the network evaluate even the largest errors.

B. Residual Network (ResNet)

VGG networks [5] encourage the use of simple residual networks. [5] In order to reduce layer complexity, the number of filters is doubled when the feature map's size is only half-complete [1]. Residual connections were introduced in [1], where they proposed theoretical and practical proof for the use of signal merging for image recognition. This is the original residual connection, and this is an improved version that decreases the computational cost of 11 convolutions (Fig. 3).

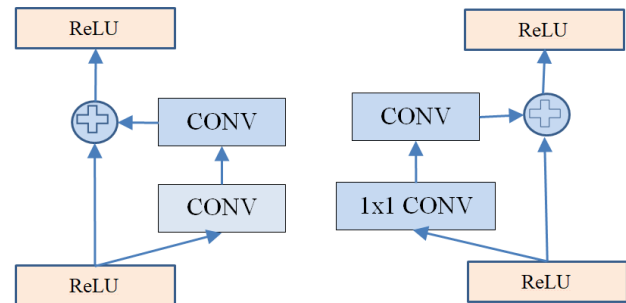


Fig. 3. Residual Connection.

The 50-layer of residual network architecture [1] detail has given below:

It is possible to get a stride size of 2 with a kernel size of 7×7 and 64 unique kernels in the first convolution layer. With a stride length of two, which is the maximum pooling layer. As a result of the three levels being repeated three times, the network now contains nine layers after 1×1 , 64 kernels, and 3×3 , 64 kernels. By the time the procedure reaches the 12th layer, the model has a 3×3 -layer kernel, a 1×1 , 128-layer kernel, and a 1×1 -layer kernel that has been replicated four times in total. There is one more layer in this level, making the total number of layers in the kernel $1 \times 1 \times 1$, 1024. There are

now a total of 18 tiers to pick from after the procedure's sixth iteration. It now has a total of nine levels of kernels: $3 * 3$, 512, and $1 * 1$, 2048. In the end, a Softmax function and a 1000-node connected layer reduce the algorithm to a single layer. So, this is a typical pool, no? The activation function and the maximum and the average pooling layer are not calculated. Adding all of the numbers together, they get a convolutional network with 50 layers: $1 + 9 + 12 + 18 + 9 + 1$.

Age estimation can be seen as both non-linear and non-category regression problems. In the current age estimation frameworks, the use of images of aging and age estimation modules is common. Shape and texture data from facial photos are commonly used in aging trait techniques. They fall under the categories of anthropometric models, AAM, AGES, age distribution, and appearance models. Regression methods or age-related institutional categories can be used to estimate an individual's age from this point onward. In the most recent research, categorical and regression methodologies are experienced jointly in the most recent research, resulting in a hybrid structure [29]. To create age estimates based solely on facial features, the robust multi-sample regression mastering rule set is also used [30]. Table I presents existing ways for estimating a person's age.

TABLE I. EXISTING AGE ESTIMATION METHODS

Network Architecture	Publishing Year	References
Multi-purpose CNN	2017	[14]
Diagnosing DLM	2017	[15]
Deep cumulatively and comparatively LM	2017	[16]
AAFs of CNN	2017	[17]
DMTL	2017	[18]
VGG-NET-GPR	2018	[19]
ELM	2018	[20]
DAG-CNN	2018	[21]
The CNN triplet ranking	2019	[22]
DeepAge	2019	[23]
Multitasks-AlexNet	2019	[24]
SADAL & VDAL	2020	[25]
LRN	2020	[26]
CR-MT	2020	[27]
MA-SFV2	2020	[28]

C. Shortcut Identity Connection

Adding more layers in the architecture of a deep neural network allows it to better recognize multiple layers of features in an image. Due to over-exercise, additional layers may not always yield precise results. The problem can be alleviated by establishing a shortcut identity connection between residual blocks. The layer of the neural network where most accuracy will be achieved must be identified [36]. Excessive fitness reduces accuracy, so don't add any more layers after you've found the one with the highest accuracy. The "shortcut identity connection" refers to the identity link between the beginning residual block and the correct stage of a residual block. When

the shortcut identity connection is used, all remaining blocks in the architecture will be eliminated [33]. The residual building block [1] is defined as:

$$b = F(a, \{W_i\}) + a, \tag{1}$$

There are two levels here, a is input, and b is output level [37]. The function represents the remaining knowledge. According to Eqn. (1), the size of input a and F must be equal. Use a linear projection if the dimensions of the input image and F don't match up. This issue can solve with a shortcut identity connection if they aren't.

$$b = F(a, \{W_i\}) + W_c a, \tag{2}$$

The experiments in this paper involve varying dimensions of images to estimate age. The shortcut identity connection is effective for the degradation problem, and there is a need to match input dimensions. Age estimation from different dimensional input images can be robust in real-life experience.

III. AGE ESTIMATION METHODOLOGY

The ResNet50 residual network is used in this paper to present a new age estimate approach called "age estimation using the shortcut identity connection of CNN" [33]. Data preprocessing, model creation or choosing a pre-made model, learning rate estimation, and model training are all part of the algorithm's process of determining age estimation. These parts work together to figure out how old you are.

A. Data Preprocessing

This module consists of input data from GitHub (ImageNet) and a loaded data set from the device (FG-NET dataset by Yanwei Fu). The model's expectations for data preprocessing are using ResNet50-selected auto normalization pixel values. Several steps are involved in the data preprocessing modules: acquiring the ImageNet data set, importing all libraries required for the ResNet50 model, identifying missing values, encoding the face aging data, splitting the dataset for the training phase, the testing phase, and the feature scaling.

There are several steps involved in data acquisition before the model may be built and developed. The algorithm needs to create a data set accurately, and then the algorithm uses a shortcut identity connection to match dimensions. A number of graphs and figures were created using the Matplotlib Python program for the experiment. In addition to importing datasets from multiple sources, and data preparation also requires the import of datasets. Setting the data set's current directory was required before import. For the experiment's training and testing phases, the algorithm needed data set splitting to calculate the mean absolute error (MAE), validation error, loss, and validation loss, as well as delete a particular row or column when more than 80% of the values were missing. When dividing a dataset in half, the ratio is 80:20. In the end, the data preparation module applies feature scaling. Within a particular range of values, the algorithm had to summarise several sorts of variables in a dataset. These things are very important in age estimation experiments of any significance: the epochs, loss functions, error functions, and validation functions for each epoch are very important as well.

B. Model Creation

The residual network has 50 layers in this experiment. In brackets, you'll see the residual building blocks (Table II). It takes 50 layers of the residual network to achieve 3.8 billion FLOPs per second [1].

The algorithm used shortcut identity connections for identity mapping to use different dimensional images and set up a freezing process to ignore layers from the model to avoid overfitting. To wrap up, the algorithm presents an alternative CNN design based on shortcut identity connections (Fig. 4).

C. Estimate Learning Rate

During these trials, the algorithm used the range of learning rates to estimate the optimal learning rate for the algorithm and data set under consideration. A maximum learning rate of 0.0001 has been employed in the one-cycle policy [6] (1 e-4). Algorithm training sessions make use of the term "epoch term." There is a count of how many times the algorithm has been trained. A batch size of 512, ResNet50, and momentum equal to 0.9 was used, and the weight decrease was equivalent to 1e-4 [6]. Since it has a rate of learning close to 1, it can be put to good use. CIFAR10 and ResNet56's training loss, validation accuracy, and validation loss are illustrated in Fig. 5's characteristic plot (a). Fig. 5 shows the age estimation method's log scale of learning rate (b).

D. Training Model

In this experiment, the auto-fit method employed a learning rate schedule with an automatic initial stop on the plateau and a

reduction of the maximum learning rate [7]. The auto-fit strategy, when supplied with the fit_onecycle parameter, reduces the learning rate of each cycle using cosine. To estimate age, the algorithm has used the Softmax image classifier with the ResNet50 model pre-trained on ImageNet.

TABLE II. 50 LAYERS OF THE RESIDUAL NETWORK

CNN layer name	Output size of layer	Layer
Convolution1	112×112	7×7, 64 stride 2
Convolution2_n	56×56	3×3 max pool, stride 2
		$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
Convolution3_n	28×28	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
Convolution4_n	14×14	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$
Convolution5_n	7×7	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	Avg. Pool, 1000-d fc, Softmax
Floating point operations per second (FLOPs) [1]		3.8×10^9

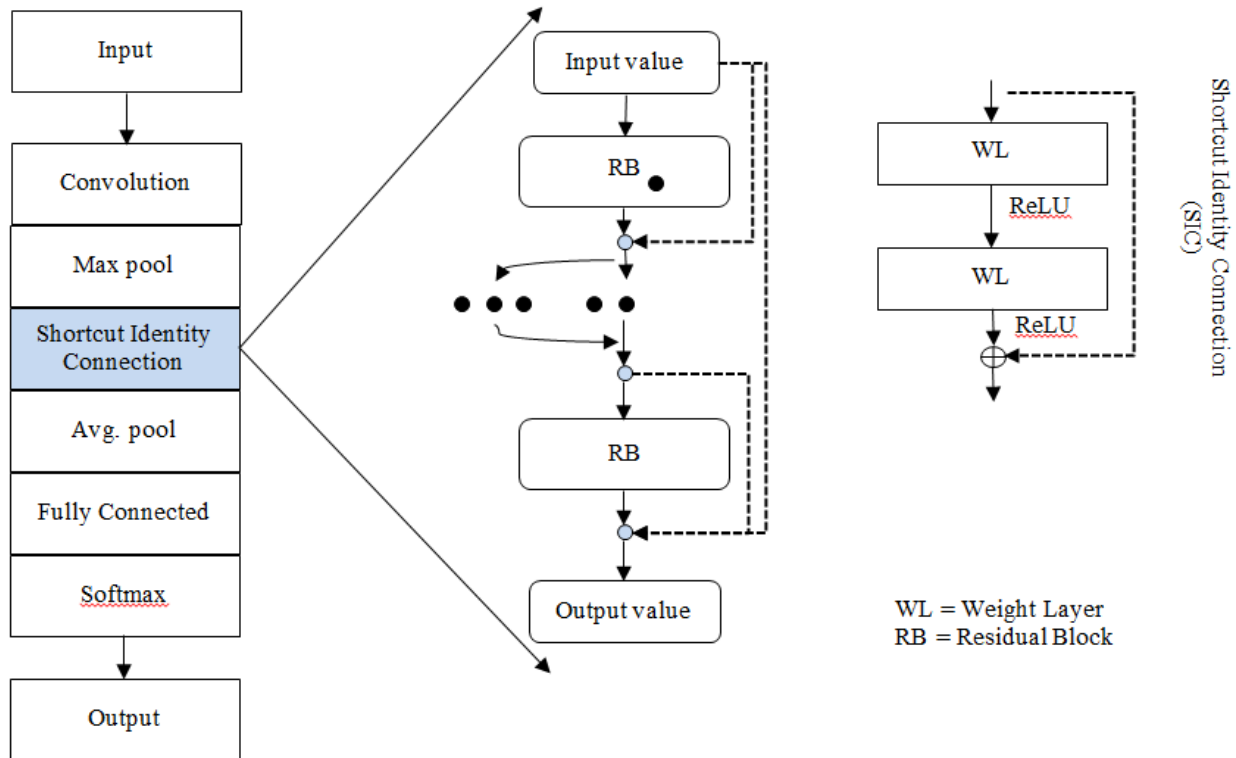
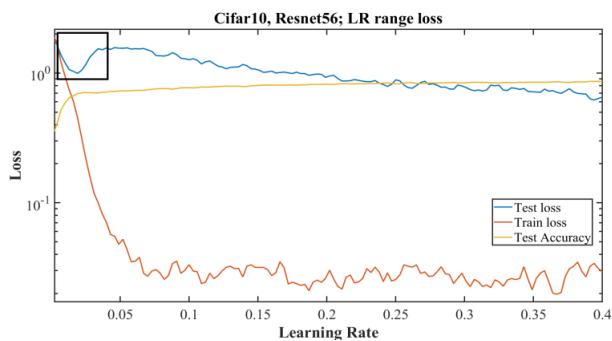
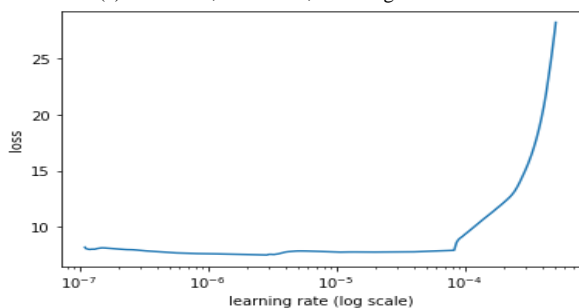


Fig. 4. Proposed an Alternative CNN Architecture using Shortcut Identity Connection.



(a) CIFAR10, ResNet50; Learning Rate Loss.



(b) Age Estimation Approach Learning Rate.

Fig. 5. Learning Rate Log Scale.

IV. EXPERIMENTS AND RESULTS

To obtain these results, the algorithm used the ImageNet and FG-NET datasets together. The experiments use a variety of epochs to evaluate the outcomes. Epoch specifies how many times the algorithm runs during the training process. Each sample of the training data set is only transmitted through the neural network once during a single epoch in a single neural network [33]. One epoch is made up of one or even more batches. The facial photos of Malaysian and Bangladeshi citizens are used to test the suggested approach. Images from the ImageNet and FG-NET databases are shown in Fig. 6. Another experiment was performed on the FG-NET database utilizing the ResNet50 design. To see how the shortcut identity connection worked, the algorithm used images of different sizes and resolutions.

A total of 21690 photos from the ImageNet collection were utilized to train the algorithm, and a further 2411 images that have been verified were used to conduct the testing. To obtain the most accurate results while using the fewest facial photos possible, we have sought to reduce the number of images used. The ImageNet data set can be used to estimate an age with an average accuracy of 81.75 percent and MAE of 2.27 in the training session (Fig. 7). The evaluation method of Mean Absolute Error (MAE) [8] is shown in Eqn. (3).

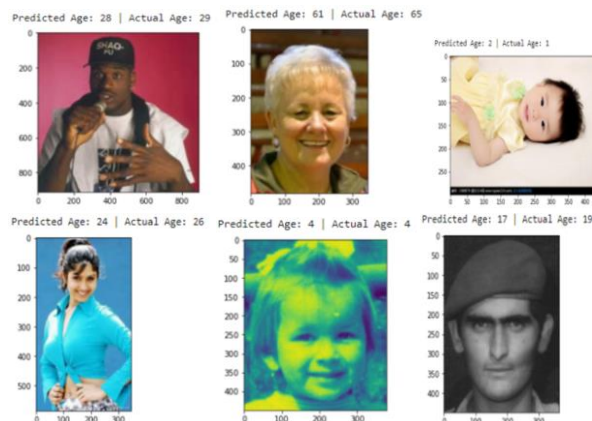
$$MAE = \sum_{i=1}^n |\hat{l}_i - l_i| / N_t, \quad (3)$$

Where: \hat{l}_i age is recognized, i is a testing sample, and the total number of testing samples is N_t .

Only 352 facial photos from the FG-NET database were used throughout the training period to test how this method worked with a smaller number of facial aging cases. Fig. 8 shows the FG-NET dataset age estimation accuracy [23]. In the

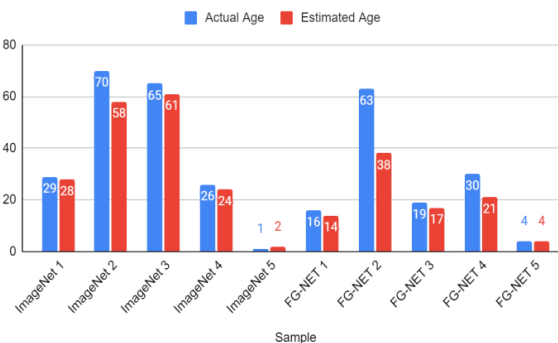
training session, MAE for the FG-NET dataset was 2.38. The accuracy rate is 57% across the experiment.

A parameter includes added to estimate age automatically from the camera-captured facial image. The system will automatically turn on the device's camera at the end of the training and testing session of the algorithm. It helps to capture the live facial image to estimate age. After capturing the photo, the estimator will estimate the age. The accuracy of the camera-captured image is 93.75%. A sample has been shown in Table III. Only images from the ImageNet database have been used in training to figure out how old people are from their live webcam captured images, so far.



(a) Sample Facial Images.

Actual Age and Estimated Age



(b) Age Estimation from Facial Images.

Fig. 6. (a) Shows Sample Facial Images (b) with an Estimated Age.

Result evaluation of ImageNet images

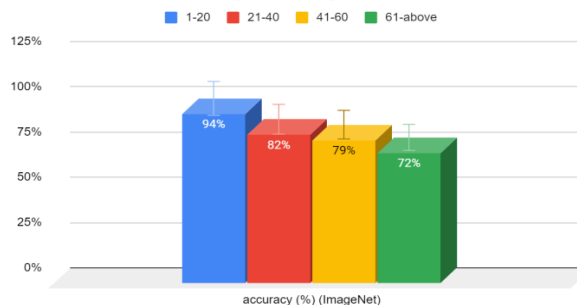


Fig. 7. The Testing Result Evaluation of ImageNet Images.

Result evaluation of FG-NET images

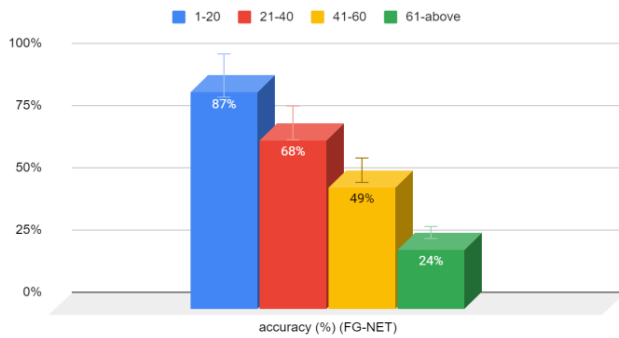


Fig. 8. Testing Results Evaluation of FG-NET Images.

TABLE III. RESULTS OF CAMERA CAPTURED IMAGES

Actual Age	Estimated Age	Facial Images
48	45	
19	20	
26	25	
19	23	
29	29	

Table IV compares the proposed approach's performance to other methods for estimating an age from facial photos and finds that the suggested method outperforms them all. ImageNet testing results have not been found to compare with the proposed algorithm's results.

TABLE IV. THE COMPARISON OF AGE ESTIMATION RESULTS

Methods	Database	MAEs
WAS [9]	FG-NET	8.06
AAS [9]	FG-NET	14.83
KNN [10]	FG-NET	8.24
BP [10]	FG-NET	11.85
SVM [10]	FG-NET	7.25
AGES [9]	FG-NET	6.77
AAM + 2D-DCT [8]	FG-NET	5.39
CMT learning [11]	MORPH-II, FGNET	3.04
AEUSICR (proposed)	ImageNet, FG-NET	2.27, 2.38

V. CONCLUSION AND FUTURE WORK

The proposed ResNet50-based CNN model for estimating age from a facial image has the best performance of any model that's already out there. Use the shortcut identity connection method to come up with a new CNN architectural design. Using different dimensional images and freezing many layers of CNN architecture, this method has been effectively implemented. The image is classified using the softmax classifier, and the positive value is normalized. Improved calculation speed and accuracy are two benefits of shortcut identity connection use in CNN systems. The proposed method achieves MAE 2.27 and MAE 2.38 using the ImageNet and FG-NET datasets. In the future, the algorithm will consider the proposed technique to determine gender and ethnicity and improve the accuracy of the age estimate method by adjusting the ResNet method. Age estimation thus relies on the capacity to classify face images. The proposed method can apply to small and big databases for facial image classification.

ACKNOWLEDGMENT

I have my deepest gratitude to Dr. Hadi Affendy Bin Dahlan, the doctor who supervised me. Patience, time, and encouragement he has given me have been greatly appreciated. He taught me a lot about image processing and computer vision in particular. Hadi Affendy bin Dahlan, in particular, taught me a great deal about doing and writing a research report. I'd want to express my gratitude to the individuals who provided feedback on my review article, which allowed me to make the necessary corrections.

REFERENCES

- [1] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. 2016.
- [2] Bengio, Yoshua, Patrice Simard, and Paolo Frasconi. "Learning long-term dependencies with gradient descent is difficult." IEEE transactions on neural networks 5, no. 2 (1994): 157-166.
- [3] Szandala, Tomasz. "Review and comparison of commonly used activation functions for deep neural networks." In Bio-inspired neurocomputing, pp. 203-224. Springer, Singapore, 2021.
- [4] D. H. Hubel and T. N. Wiesel, "Receptive fields and functional architecture of monkey striate cortex," The Journal of physiology, vol. 195, no. 1, pp. 215-243, 1968.
- [5] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).

- [6] Smith, Leslie N. "A disciplined approach to neural network hyperparameters: Part 1--learning rate, batch size, momentum, and weight decay." arXiv preprint arXiv:1803.09820 (2018).
- [7] Maiya, Arun S. "ktrain: A low-code library for augmented machine learning." arXiv preprint arXiv:2004.10703 (2020).
- [8] Asuman Günay and Vasif V. Nابیev, "Age Estimation Based on AAM and 2D-DCT Features of Facial Images" International Journal of Advanced Computer Science and Applications (IJACSA), 6(2), 2015. <http://dx.doi.org/10.14569/IJACSA.2015.060217>.
- [9] T. Cootes, G. Edwards and C. Taylor, "Active appearance models", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 6, pp. 681-685, Jun 2001.
- [10] A. Lanitis, C. Taylor and T. Cootes, "Toward automatic simulation of aging effects on face images", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 4, pp. 442-455, April 2002.
- [11] Yoo, ByungIn, et al. "Deep facial age estimation using conditional multitask learning with weak label expansion." IEEE Signal Processing Letters 25.6 (2018): 808-812.
- [12] Maria Luiza C. Passini, Katuscia B. Estébanez, Graziela P. Figueredo and Nelson F. F. Ebecken, "A Strategy for Training Set Selection in Text Classification Problems" International Journal of Advanced Computer Science and Applications (IJACSA), 4(6), 2013. <http://dx.doi.org/10.14569/IJACSA.2013.040608>.
- [13] Hjelm, R. Devon, Alex Fedorov, Samuel Lavoie-Marchildon, Karan Grewal, Phil Bachman, Adam Trischler, and Yoshua Bengio. "Learning deep representations by mutual information estimation and maximization." arXiv preprint arXiv:1808.06670 (2018).
- [14] Ranjan, S. Sankaranarayanan, C. D. Castillo, & R. Chellappa, "An all-in-one convolutional neural network for face analysis," in 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), pp. 17-24, May, 2017.
- [15] J. Xing, K. Li, W. Hu, C. Yuan, & H. Ling, "Diagnosing deep learning models for high accuracy age estimation from a single image," in Pattern Recognition, vol. 66, pp. 106-116, 2017.
- [16] K. Li, J. Xing, W. Hu, & S. J. Maybank, "D2C: Deep cumulatively and comparatively learning for human age estimation," in Pattern Recognition, vol. 66, pp. 95-105, 2017.
- [17] L. Hou, D. Samaras, T. M. Kurc, Y. Gao, & J. H. Saltz, "Convnets with smooth adaptive activation functions for regression," in Proceedings of machine learning research, vol. 54, pp. 430, 2017.
- [18] F. Wang, H. Han, S. Shan, & X. Chen, "Deep multitask learning for joint prediction of heterogeneous face attributes," in 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), pp. 173-179, May, 2017.
- [19] J. Wan, Z. Tan, Z. Lei, G. Guo, & S. Z. Li, "Auxiliary demographic information assisted age estimation with cascaded structure," IEEE transactions on cybernetics, vol. 48, no. 9, pp. 2531-2541, 2018.
- [20] M. Duan, K. Li, & K. Li, "An Ensemble CNN2ELM for Age Estimation," IEEE Transactions On Information Forensics And Security, vol. 13, no. 3, 2018.
- [21] S. Taheri, & Ö. Toygar, "On the use of DAG-CNN architecture for age estimation with multi-stage features fusion," Neurocomputing, vol. 329, pp. 300-310, 2019.
- [22] W. Im, S. Hong, S. E. Yoon, & H. S. Yang, "Scale-Varying Triplet Ranking with Classification Loss for Facial Age Estimation," in Computer Vision – ACCV 2018. Lecture Notes in Computer Science, vol. 11365. Springer, Cham, 2018.
- [23] O. Sendik, and Y. Keller. "DeepAge: Deep Learning of face-based age estimation," in Signal Processing: Image Communication, vol. 78, pp. 368-375, 2019.
- [24] Savov, M. Ngo, S. Karaoglu, H. Dibeklioglu, & T. Gevers, "Pose and Expression Robust Age Estimation via 3D Face Reconstruction from a Single Image," in Proceedings of the IEEE International Conference on Computer Vision Workshops, pp. 0-0, 2019.
- [25] H. Liu, P. Sun, J. Zhang, S. Wu, Z. Yu, & X. Sun. "Similarity-Aware and Variational Deep Adversarial Learning for Robust Facial Age Estimation," IEEE Transactions on Multimedia, 2020.
- [26] P. Li, Y. Hu, X. Wu, R. He, & Z. Sun. "Deep label refinement for age estimation," Pattern Recognition, vol. 100, p. 107178, 2020.
- [27] N. Liu, F. Zhang, & F. Duan. "Facial Age Estimation Using a Multitask Network Combining Classification and Regression," IEEE Access, vol. 8, pp. 92441-92451, 2020.
- [28] X. Liu, Y. Zou, H. Kuang, & X. Ma. "Face Image Age Estimation Based on Data Augmentation and Lightweight Convolutional Neural Network," Symmetry, vol. 12, no. 1, pp. 146, 2020.
- [29] 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 recognition, 44(6), pp.1262-1281.
- [30] Ni, B., Song, Z. and Yan, S., 2011. Web image and video mining towards universal and robust age estimator. IEEE Transactions on Multimedia, 13(6), pp.1217-1229.
- [31] Gong, S., Liu, C., Ji, Y., Zhong, B., Li, Y. and Dong, H., 2018. Advanced image and video processing using MATLAB (Vol. 12). Springer.
- [32] Myers Jr, L.R., Keller, J.G., Rogers, S.K. and Kabrisky, M., 1998, March. Evolution programs for Bayesian training of neural networks. In Applications and Science of Computational Intelligence (Vol. 3390, pp. 90-98). SPIE.
- [33] S. Pramanik and H. Affendy Bin Dahlan, "Age Estimation Using Shortcut Identity Connection of ResNet50 Based on Convolutional Neural Network," 2021 International Conference on Electrical Engineering and Informatics (ICEEI), 2021, pp. 1-7, doi: 10.1109/ICEEI52609.2021.9611146.
- [34] Pedrycz, Witold, and Shyi-Ming Chen, eds. Deep learning: Concepts and architectures. Springer International Publishing, 2020.
- [35] Al-Mushayt, Omar Saeed. "Automating E-government services with artificial intelligence." IEEE Access 7 (2019): 146821-146829.
- [36] Chen, Chun-Fu, Quanfu Fan, Neil Mallinar, Tom Sercu, and Rogerio Feris. "Big-little net: An efficient multi-scale feature representation for visual and speech recognition." arXiv preprint arXiv:1807.03848 (2018).
- [37] Jain, Anil K., Arun Ross, and Salil Prabhakar. "An introduction to biometric recognition." IEEE Transactions on circuits and systems for video technology 14, no. 1 (2004): 4-20.