Age classification using convolution neural networks using a local dataset

Syed Jawad Hussain[a], GulMar Khan[b], Umer Farooq[c], NZ Jhanjhi[d], M N Talib[e], and G.Suseendran[f]

[a] Associate Professor, Department of Computer Science, Barani Institute of Information Technology, Pir Mehr Ali Shah Arid Agriculture University Rawalpindi, Rawalpindi, Pakistan.
[b] Post Graduate Student, Department Of Computer Science &IT, The University of Lahore, Islamabad, Pakistan
[c] Department of Computer Science &IT, Sir Syed Case Institute of Technology, Islamabad, Pakistan,
[d] School of computer science and Engineering, SCE, Taylor's University, Malaysia,
[e] Papua New Guinea University of Technology, Lae, PNG
[f]Assistant Professor, Department of Information Technology, Vels Institute of Science, Technology & Advanced Studies, India

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 20 April 2021

Abstract: Human faces reveal various information including gender, age and ethnicity. A challenging problem for existing computer vision system is to estimate human age effectively. The age of a person can be estimated based on how the features tend to appear at a certain age. This can help develop applications and devices based on age classification. To develop and deploy such an application or device within Pakistan there is a need to identify if age classification will work with the local features. The existing dataset was not covering the local features regarding ageing for people of Pakistan. In this study, a dataset comprising of 9920 images from different areas of Pakistan is created. Features extracted from specific areas of the face are feed for training the models. Feature extraction and classification were done by using a convolutional neural network (CNN). The results help in concluding that CNN performed reasonably well for Pakistani face dataset for age classification where the model accuracy came up to 92%. Various applications like age-restricted item vending machines or access to hospitals for under-age visitors can be developed by using the model trained in this work.

Keywords: Age; Face; Classification; CNN.

1. Introduction

Humans have always been interested in researching ageing and the impact of ageing on human psychology. Different age classification methods have been investigated in the field of psychology. Psychological studies have aimed to examine the influence of the subject's age and the influence of ageing on humans (Chopik et al., 2018). Research in the domain of medical science, psychology and a better understanding of the human perception of vision has contributed to how the human brain recognizes and classify faces. This understanding only helped the researchers in efforts for developing automated systems which could classify people into various age groups. This capability can help develop and deploy services and products specific to the user's age.

Face detection has been a well-researched domain whereas there is still a need to develop better models and algorithms for age classification based on the face images. This is still an active research area with varied prospective application domains.

For humans, to detect subtle variations in the face in real life especially for unknown or rarely seen people is a difficult task. For example, if a person living in Pakistan see another person only once in life, who is of Chinese origin, it is very difficult to recognize him/her in a group of Chinese people and the same is true in reciprocal.

This becomes a complex problem when an automated age classification system needs to be developed which can work with different races. Face detection techniques improved over time and systems are being deployed which can detect, recognize and take relevant actions. Human face image-based age estimation was not the main focus of most of this work. Accuracy reported in earlier works is on the lower side. The datasets of human face available for face detection and recognition mostly haven't included people from sub-continent. Even within Pakistan people from different areas ageing happens differently. Hence, when the same datasets are being used for age classification the trained models do not generalize. It was obvious that there is a need to collect a dataset of people from all over Pakistan.

Development of an age classification based application can be deployed for preventing minors from accessing prohibited items like cigarettes and alcohol from vending machines. Age-restricted website access and premises access is another area where such applications can help improve facility management. The probable area of application of such capability is not restricted to the above-mentioned domain but can be extended to old age house care or medical facilities. Another application of age classification can be to filter the search results from a search engine, especially images, based on the detected age of the user.

Also, age classification and human face recognition have a crucial role in the following fields: age estimation, defence, security, border control, age and gender classification, forensic research, human-computer interaction, and social media.

Recently, research on CNN-based deep learning methods, such as AlexNet, VggNet as was used by (Can Malli et al., 2016), and Inceptionv3 classifiers, has used datasets for image classification, image detection, fruit classification, lung cancer classification, and leaf detection. Deep learning approach for feature extraction, automatic detection and classification worked well for large datasets. AlexNet, VggNet, and Inception v3 have been used for multiclass classification of face dataset age classification. We want to test CNN for age classification problem with the Pakistani face dataset that we have collected.

The first part of this work is to collect the images into a dataset from different parts of Pakistan. Preprocess the dataset and perform feature extraction. Machine learning algorithm i.e. CNN for feature extraction and classification for the dataset was used.

2. Related work

The following articles discuss the latest advancement in the domain of age classification using face image dataset.

2.1 Age estimation based on face images and pre-trained CNNs

Age and gender classification based on images is studied in this work. For such tasks, efficient architecture development was achieved by the work done in (Anand et al., 2017). The study aimed to improve the results. Previously published architectures were tweaked in terms of depth of the network, the number of parameters in the network, and modifications to the parameters of the network. To take advantage of gender-specific age characteristics inherent to images, this work focuses on coupling the architectures for age and gender classification. Gender classification, which has more prominent intragender facial variations and fewer number of potential classes, is an easier task compared with age classification. Age classification has improved by training different age classifiers for each gender. The results were improved from 56.26% to 82.35%, as compared with the baseline model.

2.2 Age-invariant face recognition using CNNs

Branch Convolutional Neural Network (B-CNN) have been trained on two different datasets. A CNN architecture, which consists of special and temporal networks, has been proposed. The network was trained on human motion database called HMDB-51 and UCF-101. Results show that if the dense optical flow was used in CNN training, then performance is enhanced. State-of-the-art results were obtained from both the datasets (Wen et al., 2016) (Y. Wang et al., 2018) (Li et al., 2018).

2.3 Age classification using CNN with multiclass

A CNN model for multiclass classification has been proposed for developing Sighthound's automated age and gender classification system. An important feature is that the system, consisting of deep CNNs, which provide improved results on competitive datasets, make computations efficient. Most tag datasets were collected between semi-supervised pipelines for efficient annotation. The trained model performed with 76.1% accuracy for public datasets (Liu et al., 2018).

2.4 Deep CNN for age estimation on VGG-face models

Automatic age estimation from real-world and unconstrained face images is rapidly gaining importance. In this work, a deep CNN model that was trained on a database for face recognition tasks is used to estimate the age information on an audience database. This project has reported the following contributions to the body of knowledge: (1) this work proves that a CNN model, which is trained for face recognition tasks, can be utilized for age estimation and can improve performance. (2) Overfitting can be overcome by using a pre-trained CNN on a large database for face recognition tasks. (3) In addition to the number of training images and the number of subjects in a training database, the pre-training task of the used CNN influences the performance of the age estimation model (Qawaqneh et al., 2017).

Figure 1 Dataset sample images

Some other researchers also discussed the issues of working in the domain of age classification using face images. Zhang et al. implemented local deep learning neural network (LDNN) by face detection and facial landmark acquisition which were converted to patches and then LDNN were trained on the patches (Zhang & Xu, 2018). Rothe et al. worked with single image without facial landmarks (Rothe et al., 2018). (J. Wang et al., 2016) used CNN-RNN for multi-label image classification and the discussion in (Sharif Razavian et al., 2014) showed how CNN features could be used for establishing a baseline for recognition. The work done by (Levi & Hassner, 2015) also used CNN for age and gender classification. In (Chen et al., 2017) instead of multi-label approach they used ranking-CNN which gave smaller estimation errors. Bianco showed that using deep CNN pre-trained model for face recognition help achieve better accuracy (Bianco, 2017). Ranjan et al. also favoured the use of CNN for face analysis. In (KoK et al, 2019) and (Fatima et al, 2019) used the machine learning for the same.

3. Materials and methods

To train a model which can effectively work for people of sub-continent generally and for Pakistan specifically this work made an effort towards generating a local face dataset. Images from various parts of Pakistan were collected where gender ratio was also considered and age variation was also kept in focus.

The dataset used in this study is a collection of images from different cities in Pakistan as shown in Figure 1. This dataset is divided into four classes, which are named as follows: Male Adult (MA), Male Young (MY), Woman Adult (WA), and Woman Young (WY). The dataset consists of 9920 images which are divided into testing, validation and training subsets. There are 7600 coloured images (200 × 200 pixels each) in the testing and validation subsets, whereas there were 2320 images in the testing subset. Each subtest contains four classes i.e. class 0, class 1, class 2 and class 3 as shown in

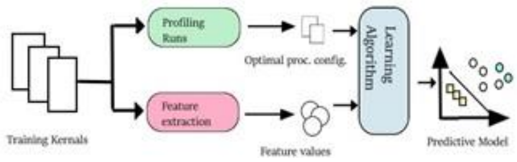
Table ***1***. It also shows the distribution of gender in the dataset.

Table 1 Data label distribution in the dataset

	Label	Training set	Test set	
Class 0	MA	2100	650	2750
Class 1	MY	2100	650	2750
Class 2	WA	1700	510	2210
Class 3	WY	1700	510	2210
		7600	2320	9920

4. Methodology

In this section, we describe the proposed architecture for age classification as shown in Figure 2. The first part discusses the experimental setup for this research. Then a discussion regarding the feature extraction and model



selection is included.

Figure 2 Architectural flow diagram

4.1 Experimental setup

Dataset was prepared for the implementation of CNN and was divided into different subsets of training, testing and validation. The organization and class-wise details are shown in **Error! Reference source not found.**.

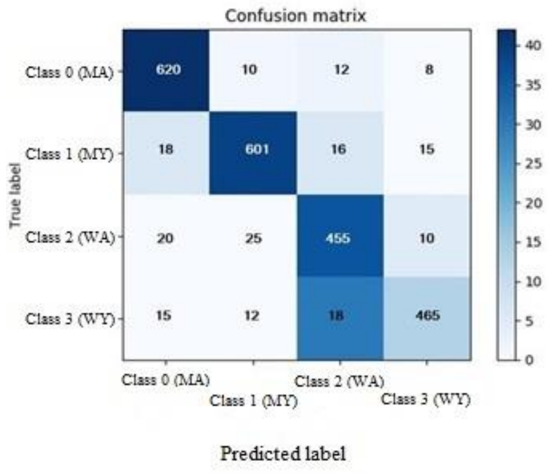
Table 2 Dataset description

Dataset	Main Folder	Sub Folder	Images
4-Class	Train	MA	2100
		MY	2100
		WA	1700
		WY	1700
	Test	MA	650
		MY	650
		WA	510
		WY	510

The CNN model was trained on the dataset created from Pakistani face images and the settings of the models are shown in Table 3. Keras was used for preprocessing and loading of the data into memory. The training log was saved to a file so that a graph showing the training results could be plotted. These generators were passed to the training module of the Keras to start the training procedure.

Table 3 Parameter Settings

Learning algorithm	Settings
Dropout probability for input/hidden units	0.025/0.5/0.65
Initial/final momentum	0.5/0.99
Number of hidden units	512
Number of hidden layers	3
Activation function	ReLU



5. Experiments and Results

Figure 3: Confusion Matrix

Error! Reference source not found. shows the confusion matrix of prediction results produced by CNN. The number of correct and incorrect predictions are summarized with count values and broken down by each class. The confusion matrix shows the ways how the classification model has correctly and incorrectly classified values. **Error! Reference source not found.** shows four classes, namely classes 0, 1, 2, and 3. The diagonal values of each class show the correct classification. A total of 620 out of 650 values in class 0 are correctly classified. Similarly, classes 1, 2, and 3 have 601 out of 650, 455 out of 510, and 465 out of 510 correctly classified values,

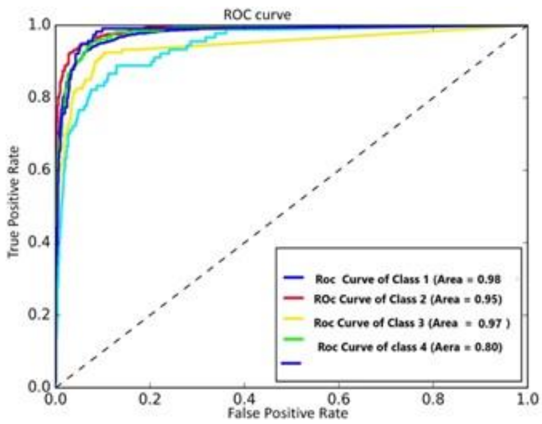


Figure 3 ROC Curve

respectively.

The receiver operating characteristic (ROC) shows the true positive and false-positive rates for four different classes. The dotted line represents the threshold value between the classes. If any class area value is greater than this threshold value, then it indicates that the model has a huge chance to predict accurately in that class. As shown in Figure 4, the area under the curve value of class 1 is 0.98, which shows that this class has a 98% chance to predict the correct value of that class. Similarly, that of classes 2, 3, and 4 is 0.95, 0.97, and 0.80, respectively.

The high value represented means that both classes have a low false-positive rate as shown in Table 4***Table 4***. Similarly, the recall of class 1 is 0.90, which shows that this class has a low false-negative rate. The F1 value of class 2 is 0.97, which shows that this class mostly has correct predictions.

Table 4 CNN classification report

	Precision	Recall	F1-Score	Support
Class 0 (MA)	0.95	0.88	0.91	650
Class 1 (MY)	0.78	0.90	0.83	650
Class 2 (WA)	0.95	1.00	0.97	510
Class 3 (WY)	1.00	0.91	0.95	510
	0.92	0.92	0.92	2320

Table 4 shows the precision, recall, and F1 measurements of each class. The precision of classes 0 and 2 is 0.95. Figure 4 shows the training and validation accuracy against the number of epochs or iteration. This accuracy increases with the increase in the number of epochs. A training accuracy of 97% and a validation accuracy of 92% are obtained, and the curve indicates that the results are not improving anymore. Besides, further training leads to the overfitting of the model. Hence, further training must be stopped.

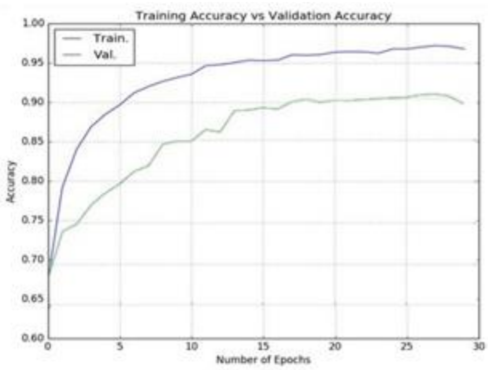
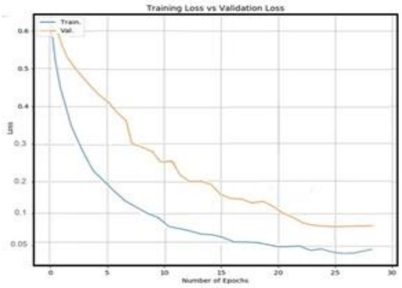


Figure 4 Training and validation accuracy for CNN

Figure 5 shows that the training and validation losses decrease with the increase in the number of epochs. A training loss of 0.03 and a validation loss of 0.08 are obtained, and the curve indicates that the results are not



improving anymore. Also, further training leads to the overfitting of the model.

Figure 5 Training and validation loss for CNN

6. Conclusion and Future Work

In this work, a Pakistani face image dataset of 9920 images was created from all over Pakistan. This dataset contained male and female faces. CNN was implemented on the dataset and validation accuracy of 92% was achieved. Several applications for local use can be created by using the model trained in this work.

To further enhance this work we aim to implement and compare several algorithms. Dataset also needs to be enhanced in two different ways. Initially, by including more images in the existing classes and then including images for more age classes.

References

1. Anand, A., Labati, R. D., Genovese, A., Muñoz, E., Piuri, V., & Scotti, F. (2017). Age estimation based on face images and pre-trained convolutional neural networks. *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, 1–7.
2. Arunkarthikeyan, K. and Balamurugan, K., 2021. Experimental Studies on Deep Cryo Treated Plus Tempered Tungsten Carbide Inserts in Turning Operation. In Advances in Industrial Automation and Smart Manufacturing (pp. 313-323). Springer, Singapore.
3. Balamurugan, K., Uthayakumar, M., Sankar, S., Hareesh, U.S. and Warrier, K.G.K., 2018. Modeling and surface texturing on surface roughness in machining LaPO4–Y2O3 composite. Materials and Manufacturing Processes, 33(4), pp.405-413.
4. Balamurugan, K., Uthayakumar, M., Sankar, S., Hareesh, U.S. and Warrier, K.G.K., 2018. Effect of abrasive waterjet machining on LaPO 4/Y 2 O 3 ceramic matrix composite. Journal of the Australian Ceramic Society, 54(2), pp.205-214.
5. Balamurugan, K., Uthayakumar, M., Sankar, S., Hareesh, U.S. and Warrier, K.G.K., 2017. Mathematical modelling on multiple variables in machining LaPO4/Y2O3 composite by abrasive waterjet. International Journal of Machining and Machinability of Materials, 19(5), pp.426-439.
6. Balamurugan, K., Uthayakumar, M., Sankar, S., Hareesh, U.S. and Warrier, K.G.K., 2018. Preparation, characterisation and machining of LaPO4-Y2O3 composite by abrasive water jet machine. International Journal of Computer Aided Engineering and Technology, 10(6), pp.684-697.
7. Bianco, S. (2017). Large age-gap face verification by feature injection in deep networks. *Pattern Recognition Letters*, *90*, 36–42.
8. Can Malli, R., Aygun, M., & Kemal Ekenel, H. (2016). Apparent age estimation using ensemble of deep learning models. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 9–16.
9. Chen, S., Zhang, C., Dong, M., Le, J., & Rao, M. (2017). Using ranking-cnn for age estimation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 5183–5192.

10. Chopik, W. J., Bremner, R. H., Johnson, D. J., & Giasson, H. L. (2018). Age Differences in Age Perceptions and Developmental Transitions. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00067
11. Eleyan, A., & Demirel, H. (2005). Face recognition system based on PCA and feedforward neural networks. *International Work-Conference on Artificial Neural Networks*, 935–942.
12. Fatima-tuz-Zahra, N. Jhanjhi, S. N. Brohi and N. A. Malik, "Proposing a Rank and Wormhole Attack Detection Framework using Machine Learning," 2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS), Karachi, Pakistan, 2019, pp. 1-9, doi: 10.1109/MACS48846.2019.9024821.
13. Kok, S.H., Abdullah, A., Jhanjhi, N.Z. and Supramaniam, M., 2019. A review of intrusion detection system using machine learning approach. International Journal of Engineering Research and Technology, 12(1), pp.8-15.
14. Latchoumi, T.P., Loganathan, J., Parthiban, L. and Janakiraman, S., 2016, August. OFS method for selecting active features using clustering techniques. In *Proceedings of the International Conference on Informatics and Analytics* (pp. 1-4). https://doi.org/10.1145/2980258.2982108
15. Latchoumi, T.P., Reddy, M.S. and Balamurugan, K., 2020. Applied Machine Learning Predictive Analytics to SQL Injection Attack Detection and Prevention. European Journal of Molecular & Clinical Medicine, 7(02), p.2020.
16. Levi, G., & Hassner, T. (2015). Age and gender classification using convolutional neural networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 34–42.
17. Li, Y., Wang, G., Nie, L., Wang, Q., & Tan, W. (2018). Distance metric optimization driven convolutional neural network for age invariant face recognition. *Pattern Recognition*, *75*, 51–62.
18. Liu, W., Chen, L., & Chen, Y. (2018). Age classification using convolutional neural networks with the multi-class focal loss. *IOP Conference Series: Materials Science and Engineering*, *428*(1).
19. Qawaqneh, Z., Mallouh, A. A., & Barkana, B. D. (2017). Deep convolutional neural network for age estimation based on VGG-face model. *ArXiv Preprint ArXiv:1709.01664*.
20. Ranjan, R., Sankaranarayanan, S., Castillo, C., & Chellappa, R. (2019). *All-in-one convolutional neural network for face analysis*. Google Patents.
21. Ranjeeth, S., Latchoumi, T.P. and Paul, P.V., 2020. Role of gender on academic performance based on different parameters: Data from secondary school education. Data in brief, 29, p.105257.
22. Rothe, R., Timofte, R., & Van Gool, L. (2018). Deep expectation of real and apparent age from a single image without facial landmarks. *International Journal of Computer Vision*, *126*(2–4), 144–157.
23. Sharif Razavian, A., Azizpour, H., Sullivan, J., & Carlsson, S. (2014). CNN features off-the-shelf: An astounding baseline for recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 806–813.
24. Wang, J., Yang, Y., Mao, J., Huang, Z., Huang, C., & Xu, W. (2016). Cnn-rnn: A unified framework for multi-label image classification. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2285–2294.
25. Wang, Y., Gong, D., Zhou, Z., Ji, X., Wang, H., Li, Z., Liu, W., & Zhang, T. (2018). Orthogonal deep features decomposition for age-invariant face recognition. *Proceedings of the European Conference on Computer Vision (ECCV)*, 738–753.
26. Wen, Y., Li, Z., & Qiao, Y. (2016). Latent factor guided convolutional neural networks for age-invariant face recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4893–4901.
27. Zhang, Y., & Xu, T. (2018). Landmark-Guided Local Deep Neural Networks for Age and Gender Classification. *Journal of Sensors, 2018*.