INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATIONS AND ROBOTICS
ISSN 2320-7345

# A DEEP CONVOLUTIONAL ANALYSIS ON FACE FOR HUMAN AGE CLASSIFICATION

**K.Archana [1], R.Logeswari Saranya[2]**

[1]PG Scholar, Department of IT, PSG College of Technology, Coimbatore, India,
archana.krishnan095@gmail.com.
[2]Assistant Professor, Department of IT, PSG College of Technology, Coimbatore, India,
logessaran@gmail.com.

**Abstract: -** In recent days, the age classification has become predominant in social platforms.
This paper focus on the problem of age classification using facial age database. The hypothesis has been implemented to classify images into young, middle, old age group of an image. The CNN model is implemented to train the database and to classify the age group of a person. This work performs the classification with the huge set of facial images and evaluates the performance of the training model and calculates the error rate and accuracy. This model encounters the age prediction based on facial images task to an acceptable degree of classification accuracy.

**Keyword:** Age classification, convolutional neural network, age group.

## I. INTRODUCTION

Age plays an important role in computer vision, human computer interaction and business intelligence. The age classification is important component in the business intelligence field. For example, in business intelligence the age related information can be applied for marketing study, such as which age group is visiting the shopping mall, stores frequently. It has become an active topic not only in computational platforms, also in other disciplinary such as psychology, social science and public health. This system may be used to prevent vending machines from selling products, e.g., alcohol, tobacco, to underaged individuals.

The main difficulty in the classification of age is different aging pattern of different people. It depends on various factors such as race, characteristics, way of living etc., So, the actual age of the person is difficult to predict. Therefore, the main goal is of the computer vision is to assign the age label of the face image and classifying the age group it belongs. In this thesis, the main focus is on age group of the person rather than predicting the exact age of the person. This problem has received main attention in academia and industry.

Over the last decade, in most of the methods the features and statistical model are manually designed to estimate age. The machine learning methods has been employed to this system for the classification of face images to predict the age group. In recent years, due to its good feature extraction ability, CNN has been highlighted in machine learning and pattern recognition fields. It has achieved state-of-the-art performance in image recognition and can automatically extract the features. These deep learning approaches are followed in this model to classify the different face images .Without manual filtering ,to classify unconstrained age task automatically, the CNN has been introduced and it can obtain significant performance.CNN has a greater advantage on processing these unconstrained tasks, so that accuracy of the classification through the network structure is highly improved.

The major contributions of this paper are summarized as follows:

- The age classification aiming at image tasks is proposed using the CNN method to improve the performance.
- The experiments are done with the popular dataset IMFD in order to verify and evaluate the performance.

## II. RELATED WORK

### Age Classification

Recently, age a classification has received huge attention, which provides direct and quickest way for obtaining implicit and critical social information. The challenging problem in the age classification process is that extracting the features from the face images. The texture features and the geometry features were adopted first in [20].Later, in the age estimation process Biologically Inspired Features were proposed [15].For an improvement over BIF, recently Scattering Transform (ST) [2] was proposed to add filtering routes .The k Nearest Neighbors [13], Multilayer Perceptrons [21], and Support Vector Machines (SVM) [15] were used as the classification model in the process of age estimation. In [6] the Active Appearance Model was employed to extract appearance features and shape from facial images. AdaBoost algorithm with Local Binary Patterns [1] was proposed by Yang and Ai [17]. Li et al. [26] adopted a method based on ordinal discriminative feature learning. [8], developed a Probabilistic Fusion Approach (PFA) to produce higher performance of age estimation and combines regression and classification process. In [15], for age estimation on various datasets BIF features were shown to be effective. To achieve better performance the learning algorithm were incorporated. In [14], Guo et al. proposed to use aging manifold with locally adjusted robust regressor.

Local binary patterns (LBP) [1] and Gabor[16] features were used in [7] along with a hierarchical age classifier composed of Support Vector Machines(SVM) to classify the input image to an age-class followed by a support vector regression to estimate approximate age. [5], developed a system consisting of Face Detection using Haar-like features and the Convolution Neural Network (CNN), Normalization of face, Subspace Projection for noise reduction by a combination of algorithms Spectral Regression and Principal Component Analysis and Support Vector Machine for the age group prediction. For age estimation process, all these methods have proven effective performance on small and constrained benchmark . In [4], used image alignment method for locating 52 features points and constructed an active appearance model (AAM) . Then, texture features are sent to the SVM after facial image warping to predict the level of each group. In [10]at the general development technique presented by employing LBP descriptor variations and dropout –SVM classifier on this benchmark .

## III. METHODS

### Dataset Collection

The accuracy is tested in the Unconstrained Indian Movie Face Database. The database consists of 34512 images of 100 Indian actors which are collected from 100 videos.

**Figure 1 which represents the face images of Indian movie actors**

The accuracy is tested in the Unconstrained Indian Movie Face Database. The database consists of 34512 images of 100 Indian actors which are collected from 100 videos. With the high degree of variability interms of scale, pose, expression, illumination, age, occlusion, makeup and resolution all the images are selected manually and it is cropped from the video frames. IMFDB contains brief description of each image .The IMFDB is the first face database that provides a detailed description of every image in terms of age, pose, gender, expression and type of occlusion .It is more useful which may help other face related applications. The IMFDB is created with the same goals of as Public Figures (PubFig) face datasets and Labeled Faces in the Wild (LFW) which helps to improve in many ways.

Unlike LFW and PubFigs where the images are collected from Internet sources, the IMFDB face images are collected from the Indian Movie Videos. From the last two decades the videos are collected from different kind of people in age differences is compared to the images collected from Internet through a search query. With the tight bounding box the faces are cropped with a tight bounding box. The heuristic of cropping the face from forehead to chin is followed in order to maintain consistency across the face images

These databases are built by considering following things:

- Selection of movies and actors
- Selection of frames from videos
- Cropping of faces
- Pruning the database and
- Annotation

The annotation is provided for following attributes for every image

Illumination : Bad, Medium, High
Pose : Frontal, Left, Right, Up, Down
Age : Child, Young, Middle and Old
Occlusion : Glasses, Beard, Ornaments, Hair, Hand, None, Others
Makeup : Partial makeup, Over-makeup
Gender : Male, Female

Expressions: Anger, Happiness, Sadness, Surprise, Fear, Disgust

## Deep Learning

Deep Learning is a field of Computer Science that involves use of deep networks for learning features from a dataset. Convolutional Neural Networks (CNN) is the neural networks that have more than one hidden layer. The number of hidden layers required to solve a problem has always been a topic of research. A deep learning is a multiple layer of processing unit. The feature representations of supervised and unsupervised learning in each layer forms a hierarchy from low-level to high level. The deep learning approach which uses the trainable feature extractors.

## Convolutional Neural Network

The convolutional neural network (convnets) is made up of neurons with learnable weights and bias. These weights and bias are known as filter or kernels. Each neuron which receives the several inputs and takes a weighted sum then pass it through an activation function and output is calculated.

CNN consists of input, output and multiple hidden layers. A hidden layer which consists of convolutional layer, pooling layer, fully connected layer and normalized layer

### Convolution layer

The Convolution layer applies the convolution operations to the input and results passes to the next layer. The convolutional operation is done with the number of filters.

### Stride and Padding

Stride value which determines the convolution filters to move the larger number of pixels the larger slide value is given. It controls how the filter convolves around the input volume. Padding preserves the size of the input image. If a single zero padding is added, a single stride filter movement would retain the size of the original image. Zero-padding which surrounds a matrix with zeroes. The features at the edges of the original matrix are preserved and the size of the output feature map is controlled.

### Pooling Layer

The pooling layer which is present next to the convolutional layer. This layer simplifies the output containing information from the convolution layer. Pooling is done for reducing the spatial size of the image. Pooling layer has max pooling, min pooling and average pooling. The max pooling is commonly applied in the pooling layer. The max pooling still retains the information of an image.

### The Output Layer

The output layer is said to be a fully connected layer which generates the final output. The fully connected layer is attached to the end of the network. In a dense layer, every node in the layer is connected to every node in the forthcoming layer. Every neuron from the last max-pooling layer is connected to every layer of the fully-connected layer. This layer makes classification regarding the object. The flatten layer is used to connect the pooling layer to the output layer for getting single tensor flow.

## IV. IMPLEMENTATION

The tensor flow makes it easy to understand the convolutional neural network. The Google Brain team developed the tensor flow for high computation in the concept of deep learning and machine learning concepts. It is an open source framework which gives higher performance in the numerical computation. This provides flexible architecture for the ease of deployment across various platforms which involves gpu, cpu and tpu and also from desktop, mobile

servers and edge devices. The tensor flow which supports for machine learning deep learning and other scientific and numeric computation. CNN has four layers. Each layers which performs its own operations. The layer module provides a high-level API in the tensor flow. It facilitates the creation of dense (fully connected) layers and convolutional layers add activation functions and apply dropout regularization. The CNN needs the large training datasets for good performance.

## V. EXPERIMENTS

### Preprocessing

The experiment is conducted on the IMFD dataset which is designed for age classification in an unconstrained environment. The data preprocessing is done as a initial step to get clear data from the raw data. The raw data may be incomplete, inconsistent and also may contain error. To solve these issue data preprocessing is performed on the raw image data. The deep learning techniques which expects the data in specific format. In this paper, rescaling and resizing is processed with the scikit-learn technique. The machine learning algorithms can benefit from rescaling the attributes to all have the same scale, when the data is comprised of various scales of attributes.

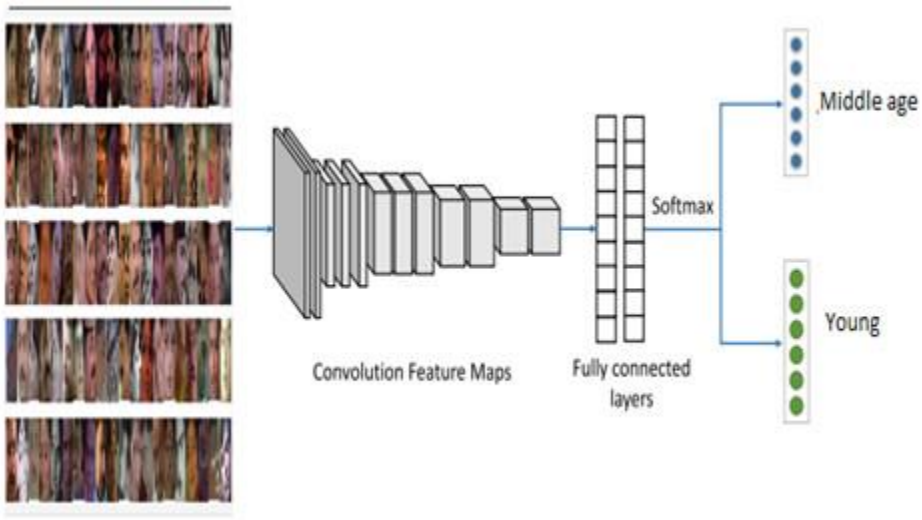


**Figure 2 Architecture of the age classification system**

The next step is the network is pertained with the large set of dataset. Here the Indian movie Face Database is used to train the network. The tensor flow and the keras model are imported to test and train the movie dataset. For the network structure, each convolutional layer is followed by a Rectified Linear Unit (ReLU), a max pooling operation. The number of layers and the number of filters in each layer are determined during network design. Each layer in the CNN which performs its own operation and validate the results. The fully connected layer is gets the output from ReLU and a dropout layer. Finally, a softmax operation is done and the training loss is calculated. The predicted output is classified into a normalized value. For this model, the metrics such as accuracy, loss and entropy is calculated.
For training the network, batch size is set to 50, the convolution filter Is set to 5 and hidden units involved. The training stops after the 50 epochs. To summarize the work,

- import various packages
- resizing all the images into arrays
- defining hyper parameters
- converting target variables to the required size
- defining the model

## VI RESULTS:

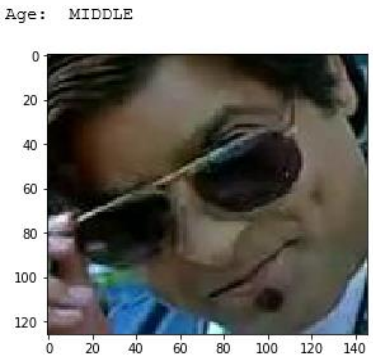


**Figure 3 predicts the age group of this person**

```
MIDDLE    0.542751
YOUNG     0.336883
OLD       0.120366
Name: Class, dtype: float64
```

**Figure 4 shows the age group and the class type**

Fig 4, represents, this model has an input layer 4 convolutional layers and max pooling layer, a flatten layer and 2 dense layer is built. From total of 3,122,691 parameters, all the parameters are trained in the network. The number of parameters can be reduced by adding convolutional and pooling layer.

| Layer (type) | Output Shape | Param # |
|---|---|---|
| input_1 (InputLayer) | (None, 32, 32, 3) | 0 |
| conv2d_1 (Conv2D) | (None, 30, 30, 96) | 2688 |
| max_pooling2d_1 (MaxPooling2 | (None, 15, 15, 96) | 0 |
| conv2d_2 (Conv2D) | (None, 14, 14, 256) | 98560 |
| max_pooling2d_2 (MaxPooling2 | (None, 7, 7, 256) | 0 |
| conv2d_3 (Conv2D) | (None, 6, 6, 256) | 262400 |
| max_pooling2d_3 (MaxPooling2 | (None, 3, 3, 256) | 0 |
| conv2d_4 (Conv2D) | (None, 2, 2, 384) | 393600 |
| flatten_1 (Flatten) | (None, 1536) | 0 |
| dense_1 (Dense) | (None, 1536) | 2360832 |
| dense_2 (Dense) | (None, 3) | 4611 |

Total params: 3,122,691
Trainable params: 3,122,691
Non-trainable params: 0

**Figure 5 represents layers, filters and parameters**

```
Epoch 46/50
13934/13934 [==============================] - 76s 5ms/step - loss: 0.1501 - acc: 0.9432 - val_loss: 1.0588 - val_a
cc: 0.7244
Epoch 47/50
13934/13934 [==============================] - 77s 5ms/step - loss: 0.1231 - acc: 0.9565 - val_loss: 1.1247 - val_a
cc: 0.7316
Epoch 48/50
13934/13934 [==============================] - 76s 5ms/step - loss: 0.0968 - acc: 0.9705 - val_loss: 1.1945 - val_a
cc: 0.7289
Epoch 49/50
13934/13934 [==============================] - 76s 5ms/step - loss: 0.0847 - acc: 0.9738 - val_loss: 1.2600 - val_a
cc: 0.7302
Epoch 50/50
13934/13934 [==============================] - 76s 5ms/step - loss: 0.0829 - acc: 0.9739 - val_loss: 1.2782 - val_a
cc: 0.7251

<keras.callbacks.History at 0x1294f5b00>
```

**Figure 6 defines the epochs, accuracy and loss**

On this system, the entire process of feature extraction, training the neural network, and evaluation took a total of 1m 15s with each epoch taking less than 0 seconds to complete. At the end of the 50th epoch, accuracy on testing and training data which is evaluated is of 97%.The loss which is limited to be at the range of 0.1501-0.08.As far now, the highest accuracy is 97% by utilizing the convolutional neural network on classification**.**

## VII CONCLUSION AND FUTURE WORK:

In this work, age classification model has been implemented to predict the age group of a specific person from the input facial images. From this work, it can be concluded that the implemented method has significally improved the accuracy when compared to the existing age classification methods. This system processes include preprocessing of input image, filtering, features extraction, train the classifier by sending extracted features by the CNN model to the softmax classifier and finally, testing is done for the data by passing it to

classifier in order to obtain the results. From the experimental results achieved, it can be concluded that the CNN classifier produces better for the age-group classification. The future work is to predict the exact age with a face and gesture of a person.

## VIII REFERENCES

[1] T. Ahonen, A. Hadid, and M. Pietikainen. “Face description with local binary patterns: Application to face recognition”. IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(12):2037–2041, 2006.
[2] K.-Y. Chang and C.-S. Chen. “A learning framework for age rank estimation based on face images with scattering transform”. IEEE Transactions on Image Processing, 24(3):785– 798, 2015
[3] T. Ahonen, A. Hadid, and M. Pietikainen. “Face description with local binary patterns: Application to face recognition”. Trans. Pattern Anal. Mach. Intell., 28(12):2037–2041, 2006.
[4] Hu, Z., Wen, Y., Wang, J., Wang, M., Hong, R., & Yan, S.” Facial Age Estimation With Age Difference”. IEEE Transactions on Image Processing, 26(7), 3087-3097,2017.

[5] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa. "Natural language processing (almost) from scratch". Journal of Machine Learning Research, 12:2493–2537, Nov. 2011.
[6] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(6):681–685, 2001
[7] S. E. Choi, Y. J. Lee, S. J. Lee, K. R. Park, and J. Kim. "Age estimation using a hierarchical classifier based on global and local facial features". Pattern Recognition, 44(6):1262–1281, 2011.
[8] Y. Fu, G. Guo, and T. S. Huang, "Age synthesis and estimation via faces: A survey",IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 11, pp. 1955–1976, Nov. 2010
[9] Chang, K. Y., & Chen, C. S. "A learning framework for age rank estimation based on face images with scattering transform". IEEE Transactions on Image Processing, 24(3), 785-798,2015
[10] E. Eidinger, R. Enbar, and T. Hassner. "Age and gender estimation of unfiltered faces". Trans. on Inform. Forensics and Security, 9(12), 2014
[11] A. Gunay and V.V. Nabiyev. "Automatic age classification with lbp. In Computer and Information Sciences". ISCIS'08. 23rd International Symposium on, pages 1–4. IEEE, 2008
[12] G. Guo, Y. Fu, C. R. Dyer, and T. S. Huang. "Image-based human age estimation by manifold learning and locally adjusted robust regression". IEEE Transactions on Image Processing, 17(7):1178–1188, 2008.
[13] G. Guo, G. Mu, Y. Fu, and T. S. Huang." Human age estimation using bio-inspired features". In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 112–119, 2009.
[14] C. Liu and H. Wechsler. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition".Trans.ImageProcessing,11(4):467–476,2002.
[15] Z. Yang and H. Ai. "Demographic classification with local binary patterns". In International Conference on Biometrics, pages 464–473. Springer, 2007.
[16] S.Yan, M.Liu, and T.S.Huang. "Extracting age information from local spatially flexible patches". In Acoustics, Speech and Signal Processing, pages 737–740. IEEE, 2008.

[17] S. Yan, X. Zhou, M. Liu, M. Hasegawa-Johnson, and T. S. Huang. Regression from patch-kernel. "In Proc. Conf. Computer. Vision Pattern Recognition". IEEE, 2008.

[18] Y. H. Kwon and N. D. V. Lobo. "Age classification from facial images". In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pages 762– 767, 1994.
[19] A. Lanitis, C. Draganova, and C. Christodoulou. "Comparing different classifiers for automatic age estimation". IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 34(1):621–628, 2004.