AGE DETECTION USING FACIAL IMAGES

Sreegeethi Devaguptam
191IT110
NITK
Surathkal
India 575025
8500847505

Thanmai
191IT127
NITK
Surathkal
India 575025
6305382826

Leela Akshaya.T
191IT226
NITK
Surathkal
India 575025
8374707104

A. Indhu Lekha
191IT207
NITK
Surathkal
India 575025
9449790733

Abstract

Age progression is generally indicated by skin texture, colour, shape of face. The skin texture changes with the age progression of a person. The same happens with facial shape too. Skin complexion does not change with age much, but it is also an important factor as different skin colours age differently. In contrast to other facial variations, aging variation presents several unique characteristics which make age estimation a challenging task. This project provides a technique to estimate the age of the above factors i.e. wrinkles, complexion and shape of the face. Wrinkle areas are detected and features are extracted from face image. Complexion is also managed by using RGB filters. Face shape is extracted by augmentation. Based on feature values of each face image is processed using **Convolution Neural Networks Model** (Deep Learning). Labelled data sets are imported and ages are divided into classes where each face image is assigned to a class. The results we get from this technique will help us in real- world applications.

Keywords: Convolution Neural Network, age progression, augmentation

1 INTRODUCTION

From a human face image, we can extract many features such as gender, skin complexion, approximate age, expression, etc., Among these, age detection can help the technological world for enhancing security, screening, intelligence gathering and many. It's high time that we concentrate on age detection as we can finish many time taking tasks effortlessly. Estimating human age from face images is still a challenging problem because, the aging process is complex, the process is almost personal. Importantly,this process depends also on factors, such as health, lifestyle, location and weather conditions. Many models(Aging Pattern Subspace, image based regression, Weighted Appearance Specific, etc.,) have been developed previously, mainly using deep learning techniques. But concerned area was the accuracy of the models, they were not substantial.

From this attempt, we were trying to bring out a age detection model which is both accurate and efficient. We may not get the accurate age of a person but we can know in which age range will the person falls under. This paper proposes a fully automatic age estimation system using Convolution Neural Network to represent aging progress. The system we proposed has four main modules: 1) Data gathering Exploratory Data Analysis (EDA); 2) Feature extraction using canny edges; 3) Usage of RandomForest Classifier and SVC classifier; 4) Deep Learning Classification Modelling; 5)Improving the Model. The input image comes from a camera frame or image file.

Contributions of our project

- Human forensics use the technology to reconstruct the facial tissues in order to identify a dead person
- Biometric identification and verification

2 LITERATURE SURVEY

International Journal for Advanced Research tried to create an age estimator in 2012. They proposed a new framework for automatic age estimation of face images. They concluded that SVMs have considerable potential as classifiers of sparse training data and provide robust generalization ability.

Senior IEEE members proposed a new method AGES related to facial age detection and estimation. They claim that AGES is not only significantly better than that of the state-of-the-art algorithms, but also comparable to that of the human observers. Besides age estimation, AGES can be utilized in other computer vision tasks. For example, with the ability to simulate facial aging effects, AGES can be used for face recognition across ages, which has been tested in the experiment conducted by them.

IEEE Society of Information Forensic and Security described a. model which was bases on robust facial alignment technique, and on iterative estimation of the uncertainties of facial feature localization. But their tests leave room for future work. This is evident by considering the drop in performance exhibited when using our benchmark compared to previous ones. It is also evident when considering failed results – all of which are be easy for a human to correctly classify, but are still very challenging from a computer vision perspective.

3 PROBLEM STATEMENT

Age Classification using facial features solved using Deep learning method and Traditional machine learning method.

3.1 Objectives

- The main objective is to estimate the age of people using facial features.
- To extract features from a person's face.
- To use deep learning and optimization techniques to increase the accuracy of the model.

4 METHODOLOGY

Imported Facial-age dataset and UTKFace dataset for this project. These images are cleaned and labelled. Facial-age dataset contains 9778 images. UTKFace dataset contains 23708 images. Combined dataset is formed by combining facial-age dataset and UTKFace dataset, converting it into JPG format.

4.1 Classes for Age Classification

Combined the different age labels into classes of age-ranges so as to prepare the datasets for age classification modelling later. It is necessary to identify a few constraints to this process. The following constraints are identified:

- **Classes need to be balanced** - Classes of age-ranges should be established according to the number of images available per age label, such that the classes are as balanced as possible. This will ensure that the classification model learns to classify each age-range equally.
- **No. of classes need to be sufficient** - The number of classes of age-ranges needs to be chosen appropriately. Too many classes will result in very narrow age-ranges, which may badly affect the model performance (it is generally more difficult to predict someone's age exactly down to the year). Too few classes will result in very wide age-ranges, which may not serve the purpose of the age classification modelling itself. Initially 11 classes of age-ranges are taken.
- **Each class needs to have sufficient data** - The number of classes of age-ranges will also depend on the resulting number of images available for each class in the dataset. Too many classes will result in narrow age-ranges, thus reducing the no. of images available for training the model on each class. However, this may not be a significant factor since data augmentation techniques are readily available and, if needed, can be employed to increase the no. of images available per class.

4.2 Train and Test splitting

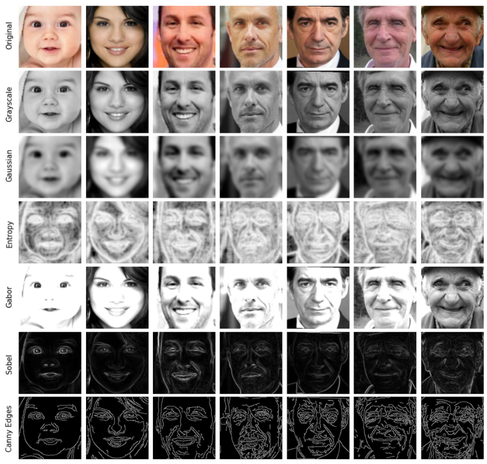
Before performing any form of classification or feature extraction on the images, it is necessary to split the combined face dataset into training and testing datasets. In this project it is divided into 70% training and 30% testing datasets

4.3 Traditional ML: Feature Extraction using Image Filters

There are a variety of techniques available out there that deal with feature extraction from images for classification modelling. In order to extract features from the facial images, applied a few different filters on some images from the dataset and try to visually spot any significant differences between them. The faces in the above plot increase in age as we move from left to right (1–90 years old). As each filter is applied to the images, as it is seen how different features in the faces are highlighted and whether or not they would be useful in differentiating between faces of different ages.

Table 1: SUMMARY OF LITERATURE SURVEY

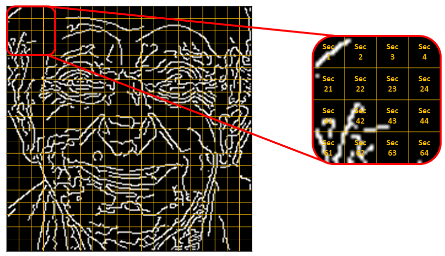
Authors	Methodology	Merits	Limitations	Additional Details
Geethi	CNN Modelling (and RGB Coloring)	Diversified the domain of input as people with different complexion age differently	No Significant increase in accuracy when compared to Gray-scale images	Optimised the CCN model built
Thanmai	Improvising Model (and Data-Set augmentation)	Could easily identify the different face structures and postures	Data becomes extremely huge and takes time for training	Facial images rotated clock-wise and anti clock-wise
Leela	Classification Modelling (and age class redistribution)	Solves the problem of different people aging at different paces	Exact age will not be found, instead only the age range with probability is given	Age range divided into 11 classes
Indhu	Data set Preparation (and Extraction of essential features)	Makes it easier to process the images as training is focused only on a particular feature	Relationships with other features may be lost	Canny Edge Extraction is used where facial lines are extracted



It can be concluded that Canny Edges may be the most useful filter for feature extraction in this case as we move from younger to older faces, the density of Canny Edges in the images seem to increase.

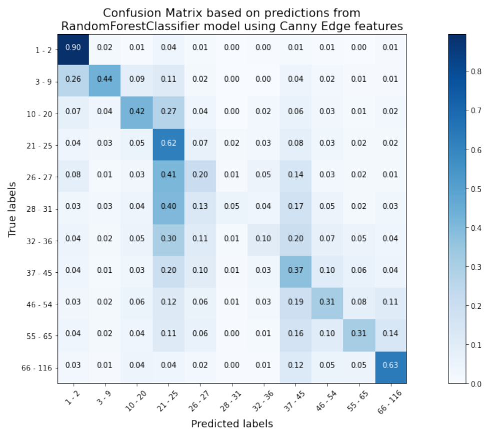
Canny Edges Feature Extraction

Broken down each 200x200 pixels image into sections of 10x10 pixels each. For each of the 400 resulting sections, then calculated the mean and standard deviation of the pixel values. This will result in 800 unique scalar values for each image, which then will be tabulated into a dataframe to be used as features in a machine learning classifier. Separate feature extraction is done for the train (23,440 images) and test (10,046 images) datasets.



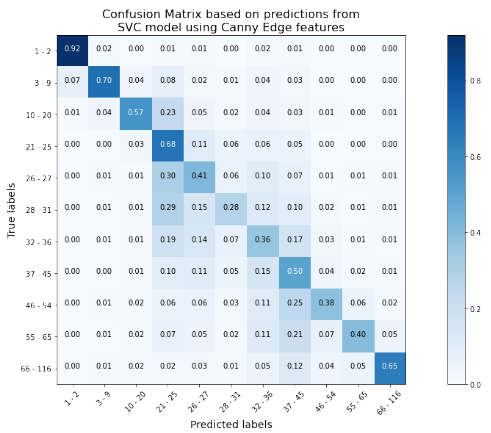
RandomForestClassifier model

The RandomForestClassifier model, with a training accuracy of 66.8% and a testing accuracy of 39.8%, clearly shows that it is over-fitting and not generalizing well on unseen testing data. The normalized confusion matrix below also show this clearly — even though the accuracy values are somewhat high for the younger age-ranges (of 1–2, 3–9, 10–20 and 21–25) and for the older age ranges (of 66–116), there is a presence of significant misclassification for the middle age-ranges of 26–65.



SVC model

As with the RandomForestClassifier model above, the SVC model, with a training accuracy of 92.9% and a testing accuracy of 53.4%, also shows that it is over-fitting and not generalizing well on unseen testing data. Even though the training and testing accuracies are better with SVC than with RandomForestClassifier, the degree of over-fit is significantly worse than RandomForestClassifier. The normalized confusion matrix below also show the same trend — even though the accuracy values are somewhat high for the younger age-ranges (of 1–2, 3–9, 10–20 and 21–25) and for the older age ranges (of 66–116), there is a presence of significant misclassification for the middle age-ranges of 26–65.



As it is clear from the summary of scores(Table 2), modelling using the traditional machine learning methodology may not be the best way to approach this problem. Modelling with deep learning is an alternative method that will be tried to improve the accuracy over machine learning.

There are, of course, a multitude of methods that could still be utilized to improve the above accuracy scores and reduce the degree of over-fit in the models. For instance, better differentiating features could be extracted from the images using some other more complicated techniques, or other classifiers could be utilized to see whether they perform better in this case.

4.4 Deep Learning

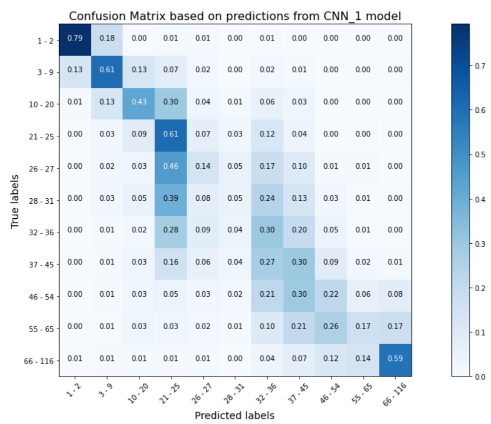
4.4.1 Importing the Datasets

In order to prepare the dataset of images to be passed into the neural network, first tried converting the images into a Pandas Dataframe but that resulted in a huge and heavy data frame which gave rise to "Out of memory" errors. To avoid these errors, adopted another approach which involved creating dataset pipelines using TensorFlow inbuilt.dataset API. This approach significantly reduced the RAM consumption.

4.4.2 Classification Modelling

After correctly importing the datasets, the next step was to build a basic Convolutional Neural Network (CNN) model that performs with reasonable accuracy on the given data and with the given number of total parameters. The idea was to get an initial benchmark on the model's performance, and then incrementally try different techniques to see whether they improve the performance from that point or not.

Earlystopping as a callback is used while training the CNN model to monitor the validation loss so as to avoid overfitting and stop the model from training for further epochs if the validation loss started to increase continuously.The images are also converted from RGB coloured to grayscale before fitting into this model. After fitting the model for 30 epochs, the following loss and accuracy scores were obtained(Table 3).



The normalized confusion matrix showed that the accuracy values are somewhat high for the younger age-ranges and for the older age ranges, when compared to traditional ML approach, there is a presence of significant misclassification for the middle age-ranges of 26–65. This may be attributed to the fact that people's facial appearances (in general) do not change as much during these middle age ranges as they do during the younger and older age ranges.

The next step in this process was to improve the model's performance using a few different techniques. In order to improve the accuracy of the neural network, tried the following strategies:

Table 2: Traditional ML accuracy

	GridSearchCV best Score (cv = 5)	Train Accuracy	Test Accuracy
RandomForestClassifier	39.3%	66.8%	39.8%
SVC	49.0%	92.9%	53.4%

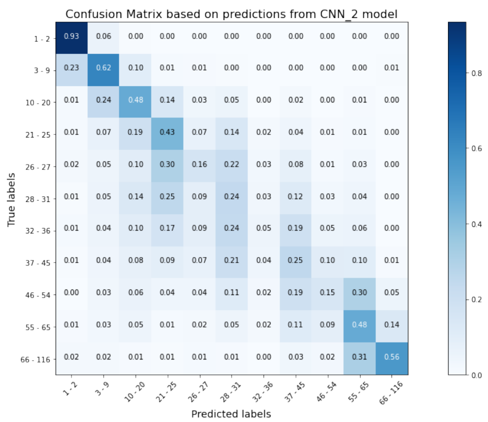
Table 3: Loss and accuracy scores for basic CNN model for 30 epochs

Model Description	Epochs	Train Loss	Validation Loss	Train Accuracy	Validation Accuracy
CNN with Grayscale images	28-30	1.5355	1.6252	41.40%	38.34%

Using RGB coloured images instead of grayscale images :

The rationale behind this is that maybe adding the colour data in the images may bring out some features within the CNN model which may enhance the overall performance of the model.

The overall accuracy(from Table 4) for the CNN model using RGB coloured images is still quite low (40%). It is slightly better than the previous CNN model using grayscale images, although with a slightly higher degree of overfitting. Observing from the normalized confusion matrix below, it is evident that the problem of misclassification for the middle age-ranges of 26-54 still persists.

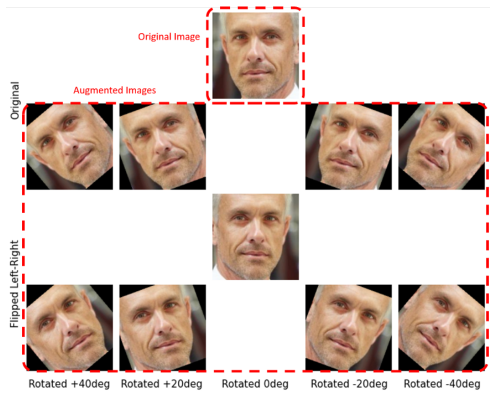


Since using RGB coloured images instead of grayscale images did not really add any significant improvements in terms of accuracy, it can be concluded that using RGB coloured images instead of grayscale images may not improve the model performance and may lead to more over-fit.

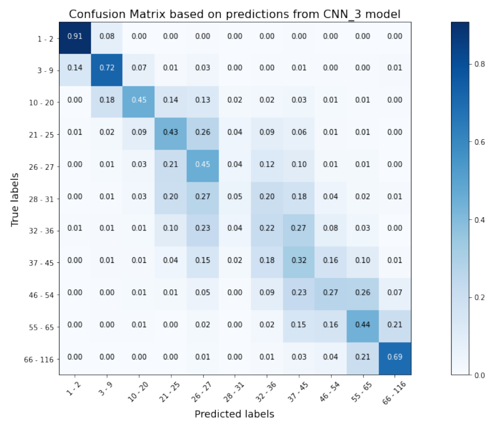
Augmenting the images in training dataset :

The rationale behind this is that increasing the amount of data for the model to train with would help to increase the variance in the dataset. This may improve the model's accuracy whilst decreasing the possibility of over-fit.

This is done by rotating the image in different angles and feeding them into the dataset. Nearly 10 more augument images are created for the original image.



As seen (from Table 5), the overall accuracy for the CNN model using the augmented training dataset improved slightly, while the degree of over-fitting reduced even after training for double the number of epochs (60 epochs instead of 30). So, it can be concluded that augmenting the images in the training dataset may help to improve the model's accuracy whilst decreasing the overfit.



Re-distributing the age-range classes :

As we have seen above, even with data augmentation, the accuracy score did not improve sig-

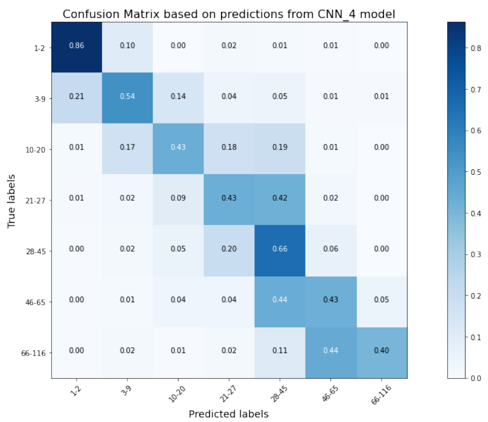
Table 4: Loss and accuracy for grayscale and RGB colored images

Model	Epochs	Train Loss	Validation Loss	Train Accuracy	Validation Accuracy
CNN(Grayscale images)	Early stop at 28 out of 30	1.5355	1.6252	41.40%	38.34%
CNN(RGB colored images)	30	1.4672	1.5971	43.56%	39.46%

Table 5: Loss and accuracy using augmentation and grayscale model

Epochs	Train Loss	Validation Loss	Train Accuracy	Validation Accuracy
30	1.4710	1.4727	42.51%	42.52%
60	1.3793	1.4208	45.45%	44.85%

nificantly. The primary issue is still the fact that there is a presence of significant misclassification for the middle age-ranges of 26–65. So, to avoid this problem, we will re-distribute the age-ranges into classes again. This time, however, instead of just looking at the available number of images per age-range, we will also take into account the human intuition factor (likely age groups that we, as humans, would classify a person into) and the accuracy scores for the individual classes shown in the confusion matrices above.Thus, the age-ranges will be redistributed into the 7 classes as follows (age in years): 1-2, 3-9, 10-20, 21-27, 28-45, 46-65 and 66-116.

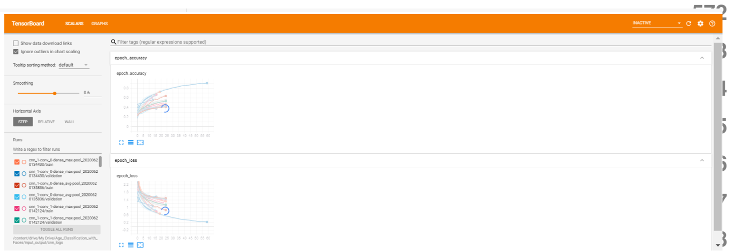


As seen from Table 6, the accuracy scores show a similar degree of over-fit as the initial CNN model fit on grayscale images (which is expected), but the accuracy scores themselves are significantly higher.The normalized confusion matrix also shows that we may have addressed the problem of misclassification for the middle age-ranges of 26–65 (at least to some extent if not completely).

Optimizing the CNN model architecture :

After exploring the above techniques of manipulating the dataset to improve the preformance of the CNN models, optimized the CNN model architecture itself to enhance the overall performance. The idea here is to design multiple models of different architectures and compare their performances in terms of loss and accuracy values using TensorBoard.

TensorBoard as a callback is used while training the multiple CNN models so as to be able to compare their performances in interactive (and definitely more intuitive) plots.



Based on the performances of all the 18 models built above with varying architectures, and after analyzing all of them using TensorBoard, the final CNN model that will be built would have the following architecture:

1. An input Conv2D layer (with 32 filters) paired with an AveragePooling2D layer.
2. 3 pairs of Conv2D (with 64, 128 and 256 filters) and AveragePooling2D layers.
3. A GlobalAveragePooling2D layer.
4. 1 Dense layer with 132 nodes.
5. An output Dense layer with 7 nodes.

Deep Learning: Building Final CNN Model :

Below is the summary of all the CNN models built so far. After understanding the effects of all the improvement and optimization techniques on the CNN model performance, the final CNN model can be defined and trained:

1. With gray-scale images instead of RGB coloured images.

Table 6: Loss and accuracy using re-distributed age-ranges and grayscale model

Epochs	Train Loss	Validation Loss	Train Accuracy	Validation Accuracy
30	1.0265	1.1075	57.58%	54.17%

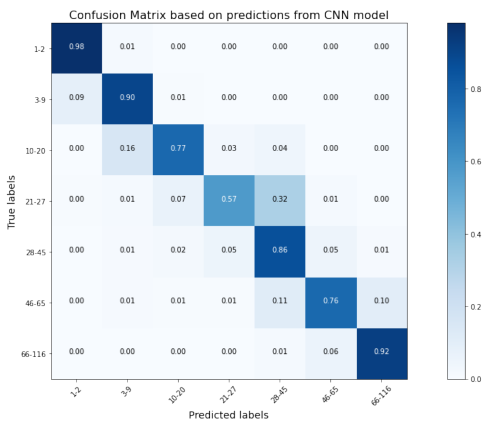
2. With augmented training data-set (234,400 images) instead of original training data-set (23,440 images).

3. for 60 epochs.

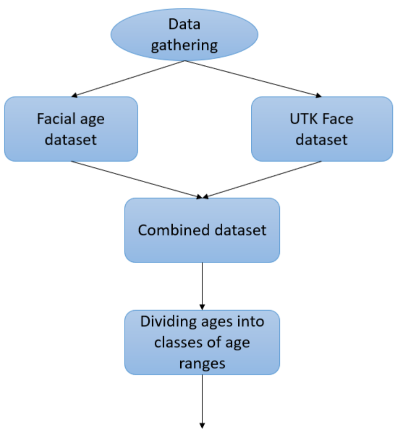
4. for re-distributed classes of age-ranges.

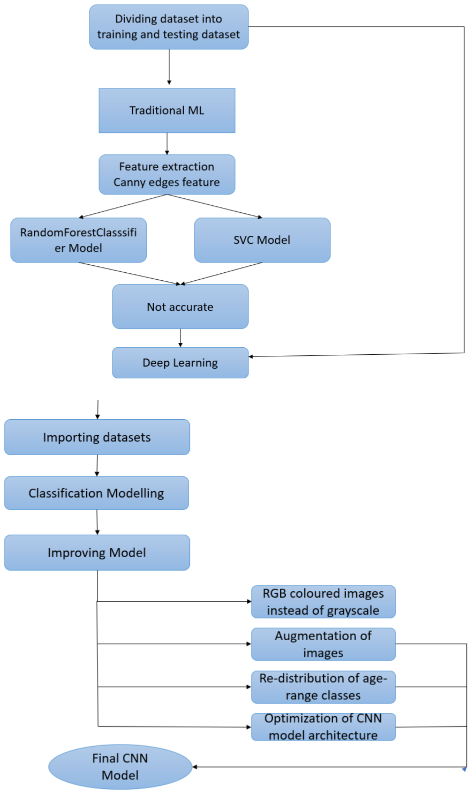
5. with an optimized architecture, comprising of:

   - An input Conv2D layer (with 32 filters) paired with an AveragePooling2D layer.
   - 3 pairs of Conv2D (with 64, 128 and 256 filters) and AveragePooling2D layers.
   - A GlobalAveragePooling2D layer.
   - 1 Dense layer with 132 nodes.
   - an output Dense layer with 7 nodes.



4.5 Flow chart





5 RESULTS AND ACCURACY

Using the traditional machine learning approach in this scenario, though possible, did not yield models with very high accuracy scores. Furthermore, this approach requires a significant amount of domain knowledge and image data processing expertise for substantial features to be extracted from the images for use in the classification models. The deep learning and neural networks approach, on the other hand, does not require any significant domain knowledge and image processing expertise, as it does not require any image features to be extracted manually. This approach also yielded much better performing models with higher accuracy scores. Thus, this comparison between the two approaches in this project highlighted the true power and possibilities of deep learning with neural networks, and provided us with a better understanding behind their increasing popularity in the recent years. The final model is tested on the face detected from the input image using OpenCV libraries.

Table 7: Loss and accuracy values for final CNN model

Epochs	Train Loss	Validation Loss	Train Accuracy	Validation Accuracy
Peak at 54 of 60 epochs	0.2430	0.6052	90.44%	82.97%

Table 8: Accuracy Table

Methods	**Model**	**Train Accuracy**	**Validation Accuracy**
Traditional Machine	Random Forest Classifier	66.8%	39.8%
Learning	SVC	92.9%	53.4%
Deep Learning	Convolution Neural Network (CNN)	90.4%	83.0%

6 Conclusion

Developed a Convolution Central Network model using deep learning to detect age of a person by face. In this model it used grayscale images instead of RGB coloured images and augmented training dataset and also optimized architecture, comprising of: an input Conv2D layer (with 32 filters) paired with an AveragePooling2D layer, 3 pairs of Conv2D (with 64, 128 and 256 filters) and AveragePooling2D layers, a GlobalAveragePooling2D layer, 1 Dense layer with 132 nodes, and an output Dense layer with 7 nodes. Finally detected face using opencv libraries and estimated age using the above built model.

7 INDIVIDUAL CONTRIBUTION

7.1 Geethi

Worked on building the basic Convolution Neural Network Model using Tensorflow. Also kept a note of the validation loss to avoid over-fitting of training data by using Earlystopping. Generated confusion which specifically determines the accuracy for each age label.

7.2 Thanmai

Implemented data set augmentation in order to increase the model's validation accuracy whilst decreasing the possibility of over-fit. Original image was rotated 20°and 40°both clock wise and anti-clock wise and fed into the CCN model and was evaluated for the increase validation accuracy.

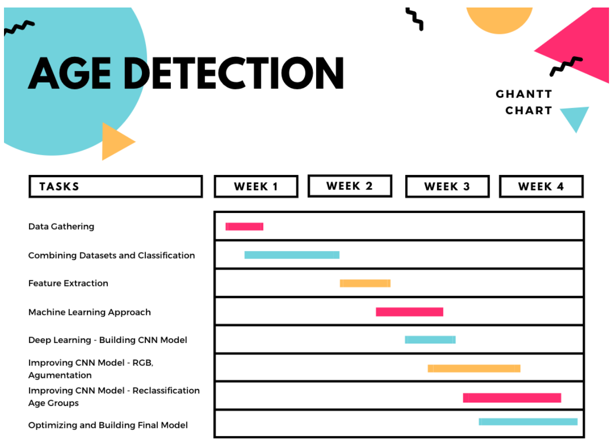
7.3 Leela

Performed re-distribution of age class ranges in order to increase the training validation accuracy without over-fitting the data-set. Also implemented Support Vector Model and Random Forest Classifier for Machine Learning approach initially on the data set.

7.4 Indhu

Worked on Data Gathering from different sources for viable data-set. Gathered a total of 25,000 images for feeding the mode. Performed feature extraction and concluded "Canny Edges" as the best parameter which forms the base of them model built by Machine Learning.

7.5 Gantt Chart



8 IMPLEMENTED/ BASE PAPER

1. Automatic Age Estimation System for Face Images by Chin-Teng Lin, Dong-Lin Li, Jian-Hao Lai, Ming-Feng Han and Jyh-Yeong Chang

9 REFERENCES

[1] Automatic Age Estimation Based on Facial Aging Patterns by Xin Geng, Zhi-Hua Zhou, Senior Member, IEEE, and Kate Smith-Miles, Senior Member, IEEE

[2] Age-Invariant Face Recognition by Unsang Park, Member, IEEE, Yiying Tong, Member, IEEE, and Anil K. Jain, Fellow, IEEE