MVJ COLLEGE OF ENGINEERING, BENGALURU- 560067

(Autonomous Institution Affiliated to VTU, Belagavi)

SOFTWARE DEVELOPMENT CLUB

INTERNSHIP REPORT

ON

AGE ESTIMATION USING AI

Submitted by

Ishaan S	**1MJAI017**
Sumeetkumar Katke	**1MJ21CD061**
Protik Chowdhury	**1MJ21AI039**
Ullas HM	**1MJ21AI056**

Under the guidance of

Mrs Thejaswini

Associate Professor

Department of Computer Science and Engineering

MVJ College of Engineering

In Partial fulfillment for the award of degree of

Bachelor of Engineering

In

Computer Science Engineering

2022-2023

MVJ COLLEGE OF ENGINEERING, BENGALURU- 560067

(Autonomous Institution Affiliated to VTU, Belagavi)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

Certified that the internship titled **"Age Estimation"** was carried out by **Ishaan S (1MJ21IA017), Sumeetkumar Katke (1MJ21CD061), Protik Chowdhury (1MJ21AI039)** and **Ullas HM(1MJ21AI056)** in partial fulfillment for the award of degree of Bachelor of Engineering in Information Science, Computer Science of the Visvesvaraya Technological University, Belagavi during the year 2022-2023. It is certified that all corrections / suggestions indicated during the internal assessment have been incorporated in the internship report deposited in the department library. The internship report has been approved as it satisfies the academic requirements in respect of internship work prescribed by the institution for the said degree.

Signature of Guide | Signature of Head of the Department | Signature of Principal

MVJ COLLEGE OF ENGINEERING, BENGALURU- 560067

(Autonomous Institution Affiliated to VTU, Belagavi)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

DECLARATION

We, **Ishaan S (1MJ21IA017), Sumeetkumar Katke (1MJ21CD061), Protik Chowdhury (1MJ21AI039)** and **Ullas HM(1MJ21AI056)** students of Second Semester B.E., Department of Information Science, Computer Science, MVJ College of Engineering, Bengaluru - 560067, hereby declare that the Internship Titled **"Age Estimation"** has been carried out by us and submitted in partial fulfillment for the award of the degree of Bachelor of Engineering in Information Science, Computer Science during the year 2022-2023.

Further we declare that the content of the report has not been submitted previously by anybody for the award of any degree or diploma to any other University.

Place: Bengaluru

Date:

Name **Signature**

1. **Ishaan S (1MJ21AI017)**
2. **Sumeetkumar Katke (1MJ21CD061)**
3. **Protik Chowdhury (1MJ21AI039)**
4. **Ullas HM (1MJ21AI056)**

ACKNOWLEDGEMENT

.

The satisfaction and euphoria that accompany a successful completion of any task would be incomplete without the mention of people who made it possible, success is the epitome of hard work and perseverance, but steadfast of all is encouraging guidance.

So, with gratitude we acknowledge all those whose guidance and encouragement served as beacon of light and crowned our effort with success.

We are thankful to our principal Dr. Mahabaleshwarappa, for his encouragement and support throughout the project work.

We consider it a privilege and honour to express our sincere gratitude to our guide Ms. Thejaswini, Associate Professor, Dept. of CSE for her valuable guidance throughout the tenure of this project work, and whose support and encouragement made this work possible.

It’s also a great pleasure to express our deepest gratitude to all faculty members of our department for their cooperation and constructive criticism offered, which helped us a lot during our project work.

Finally, we would like to thank all our family members and friends whose encouragement and support was invaluable.

Contents

CHAPTER-1
INTRODUCTION

Preamble: This chapter gives a brief overview of the project "Age Estimation using AI" and what to expect in order to make a program that can take input, compare it to a model, provide a detailed output and also carry out the execution of said program

1.1 INTRODUCTION

- In this project, we use
    - Computer Vision
    - Deep Neural Networking
    - Machine learning

  in order to estimate any person's age in real time. This being in real time not only gives a massive advantage to the company collecting the data as every stream of data is live and all the changes will be recognized instantly and recorded if need be.
- This project has endless utility and applications, be it in the spheres of marketing or healthcare or any site that requires age verification, i.e., once it has been fine-tuned and reaches a certain threshold of accuracy.
- This program uses a linear regression algorithm in order to estimate a person's age.
- The basic working of the program is,
    - It turns on the camera.
    - It recognises the face.
    - It compares the recognised face to a trained model.
    - It gives out a percentage-based report that contains the probability that the face falls under age groups between 0 – 100.
- In doing so, the program provides data that could be essential feedback to corporations in order to better their services or get a good idea about what age range is being attracted to what. For instance, let's say a multimedia sharing and streaming service such as YouTube, they can use said program to collect data regarding which age group is watching what kind of media on their site, and therefore, based on this feedback, alter

the recommended content for each pool of users simply based on a recognised pattern. Another example would be ecommerce sites such as Amazon, where the demographic data for each product is taken down and hence is recommended to the age group that views it the most.

- Not only does this contain the data to alter the recommended content to promote more sales or retention based on users, it also has the age of each user, therefore it also brings to the table another aspect of efficiency, distributing the result of the entire recommendations in a well-organized manner.

1.2 What is a Regression Model?

A regression model is a statistical model that estimates the relationship between one dependent variable and one or more independent variables using a line (or a plane in the case of two or more independent variables).

A regression model can be used when the dependent variable is quantitative, except in the case of logistic regression, where the dependent variable is binary.

In the case of our project the model is formed from a dataset. Now in layman terms, instead of making the computer look through each image in the dataset, the computer basically obtains a vague understanding of what each class looks like, and this is used to come to the desired outcome, i.e., an AR output and a report in the terminal.

CHAPTER-2
LITERATURE SURVEY

Preamble: This chapter contains the integrated information management system with vital information on destinations, products, services, transport facilities, etc., is very important.

Human face conveys a significant information about identity, age, gender, emotion, and ethnicity. It is a key demographic and a soft biometric trait for human identification. Ages are also important in the face-to-face communication between humans. Facial features influence one person's attraction to another. They can signal cues to fertility and health. Therefore, these factors can increase a person's productivity and success.

Age is one of the facial attributes which plays a significant role in helping or hindering communication. Like culture, beliefs, experience, language, age can affect both how we say what we mean, as well as how we interpret what others mean. It is a factor that influences how we communicate with each other, and can act as a barrier, along with many other factors. In a study which came out of the university of Pennsylvania and which analyses the vocabulary of 75000 facebook users, researchers showed that the vocabulary of a person can predict his/her age (Schwartz et al., 2013).

Raising the ability of a machine to recognize and interpret faces and facial traits such as age in real time can improve the interaction between humans and machines. Many researchers pay attention to the automatic interpretation of face images. Consequently, systems to identify faces and gender, estimate age and recognize emotions, have been developed.

However, faces change with age: as we get older, the skin becomes thicker and its colour and texture change, the tissue composition begin to be more sub-cutaneous and the facial skeleton lines and wrinkles appear. The process of ageing is very complicated and varies greatly for different individuals.

Thus, Automatic Age Estimation (AAE) from face images is a challenging topic because of the large facial appearance variations. It is due to a mixture of extrinsic and intrinsic factors. The extrinsic factors are mainly determined by living environment, health conditions, lifestyle,

etc., while intrinsic factors include physiological elements, such as genes. Robust AAE systems based on facial images should deal with facial expressions and appearance changes.

AAE systems have a wide range of applications in Human–Computer Interaction (HCI), in surveillance and web content filtering and in Electronic customer relationship management (E-CRM). They are needed mainly because humans fail to perform age estimation accurately. Thus, it is crucial to develop AAE systems that outperform human performance.

Different shallow learning surveys for AAE exist (Fu et al., 2010; Ramanathan et al., 2009). In fact, to the best of our knowledge, there is no comparative study that combines the most popular deep learning models, with the existing state-of-the-art CNN architectures for AAE. The main contributions of this work are as follows:

In this paper, an extensive comparative analysis of several frameworks for real AAE based on deep learning architectures is given. Several well-known CNNs and public datasets are used. The best configuration for each architecture based on Morph dataset is stated and different configurations are tested.

The used CNNs are pre-trained on ImageNet to solve object category classification. This first initialization leverages and transfers the knowledge from object recognition domain to facial age estimation domain. An evaluation framework of knowledge transfer from face recognition task across AAE is performed to study the performance of knowledge transfer from related tasks.

The AAE from face images is a challenging topic because of a mixture of extrinsic and intrinsic factors. Thus, the robustness of the best architecture is evaluated under expression changes, “crossing” ethnicity and “crossing” gender. The structure of our proposed work is organized as follows: the next section illustrates the related work. Section 4 represents the global schema and the several studied CNN architectures. In Section 5, more details about the datasets and the different experiments are given. Then, the performance of real age estimation of the several frameworks are studied. After discussing the different results and presenting a comparison with the existing approaches, section 6 concludes the paper.

CHAPTER-3
FEATURES AND REQUIREMENTS

Preamble: This chapter shows us what features the program offers along with the basic hardware and software requirements needed in order to run the program succesfully

3.1 Features

- All the functions happen in real time.
- A visual result/output.
- The age the person is most likely to be is presented.
- A percentage output is presented for each age group a person can be from age 0-100.
- Multiple such outputs can be presented simultaneously for multiple faces.
- The output comes in a well-structured and comprehensive manner.

3.2 Requirements

- Processors: Intel Atom processor or Intel Core i3 processor.
- RAM: At least 2 GB.
- Disk space: 1 GB.
- Camera.
- Operating systems: Windows* 7 or later, macOS, and Linux.
- Python version 3.6 or above. With OpenCV-Python installed

3.3 Choosing what Python module to use.

The program depends on 3 very specific functions

1, Computer Vision.

2, Machine Learning.

3, Deep Neural Networking.

The go to module for anything computer vision related is “OpenCV” as it contains everything relevant to computer vision. It most importantly contains files that facilitate and enable face detection which is an essential for the functioning of our project.

When it comes to machine learning, OpenCV also has the means to convert a dataset into a trained regression model.

Also, with respect to deep neural networking, OpenCV can produce the required output after comparing the detected face with the trained model.

Therefore, everything relevant to this project falls under one singular python module, i.e., OpenCV.

CHAPTER-4
BASIC CONCEPTS

Preamble: This chapter contains all the necessary information to understand how the program works and what theories and concepts are used in the making of this project. This is a set of prerequisite knowledge needed to recognize it's functioning

4.1 Computer Vision

Computer vision is an interdisciplinary scientific field that deals with how computers can gain high-level understanding from digital images or videos. The computer vision that this program utilizes is from an imported library called opencv or CV2. Computer vision covers everything from utilizing the camera to recognizing and matching pattens to much more.

4.2 Machine Learning

Machine learning is a field of inquiry devoted to understanding and building methods that 'learn', that is, methods that leverage data to improve performance on some set of tasks. There are multiple types of machine learning.

- Supervised learning.
    - Housing price prediction
    - Medical imaging
- Unsupervised learning.
    - Customer Segmentation
    - Market basket analysis
- Semi-supervised learning.
    - Text Classification
    - Lane finding on GPS data

- Reinforcement learning.
    - Optimized marketing
    - Driverless cars

Supervised learning: As its name suggests, Supervised Machine Learning is based on supervision. It means in the supervised learning technique, we train the machines using the "labelled" dataset, and based on the training, the machine predicts the output. Here, the labelled data specifies that some of the inputs are already mapped to the output.

More preciously, we can say; first, we train the machine with the input and corresponding output, and then we ask the machine to predict the output using the test dataset.

Unsupervised Learning: It is different from the Supervised learning technique; as its name suggests, there is no need for supervision. It means, in unsupervised machine learning, the machine is trained using the unlabeled dataset, and the machine predicts the output without any supervision.

In unsupervised learning, the models are trained with the data that is neither classified nor labelled, and the model acts on that data without any supervision.

Semi-Supervised Learning: Semi-Supervised learning is a type of Machine Learning algorithm that lies between Supervised and Unsupervised machine learning. It represents the intermediate ground between Supervised (With Labelled training data) and Unsupervised learning (with no labelled training data) algorithms and uses the combination of labelled and unlabeled datasets during the training period.

Although Semi-supervised learning is the middle ground between supervised and unsupervised learning and operates on the data that consists of a few labels, it mostly consists of unlabeled data. As labels are costly, but for corporate purposes, they may have few labels. It is completely different from supervised and unsupervised learning as they are based on the presence & absence of labels.

Reinforcement Learning : Reinforcement learning works on a feedback-based process, in which an AI agent (A software component) automatically explore its surrounding by hitting & trail, taking action, learning from experiences, and improving its performance. Agent gets rewarded for each good action and get punished for each bad action; hence the goal of reinforcement learning agent is to maximize the rewards.

In reinforcement learning, there is no labelled data like supervised learning, and agents learn from their experiences only.

The Reinforcement Learning process is similar to that of a human being; for example, a child learns various things by experiences in his day-to-day life. An example of reinforcement learning is to play a game, where the Game is the environment, moves of an agent at each step define states, and the goal of the agent is to get a high score. Agent receives feedback in terms of punishment and rewards

4.3 Deep Neural Networking

Deep learning is part of a broader family of machine learning methods based on artificial neural networks with representation learning. As our program produces an output in the form of a probability percentage table based on each age group between ages 0 to 100, it is a deep neural network that brings forth that output after comparing the detected face to the model and coming to a suitable output.

CHAPTER- 5
WORKING OF THE PROGRAM (METHODOLOGY)

Preamble: This chapter shows how the program is split into 3 parts, the input, process and output.

5.1 The Camera and its Usage

The main Module being used in our project is OpenCV aka CV2. This module permits us to utilize the maximum extent of the computer vision available in the system.
So a loop is established with the cv2.videocapture() command. Within this loop exists a waitkey delay of 1ms in order to prevent lag and to ensure proper and uninterrupted working of the camera.
The exit key is assigned to “Q”, meaning if the “Q” key is pressed while the camera is operating, the established loop will encounter a break and the camera will stop its function.

5.2 Detection of the face

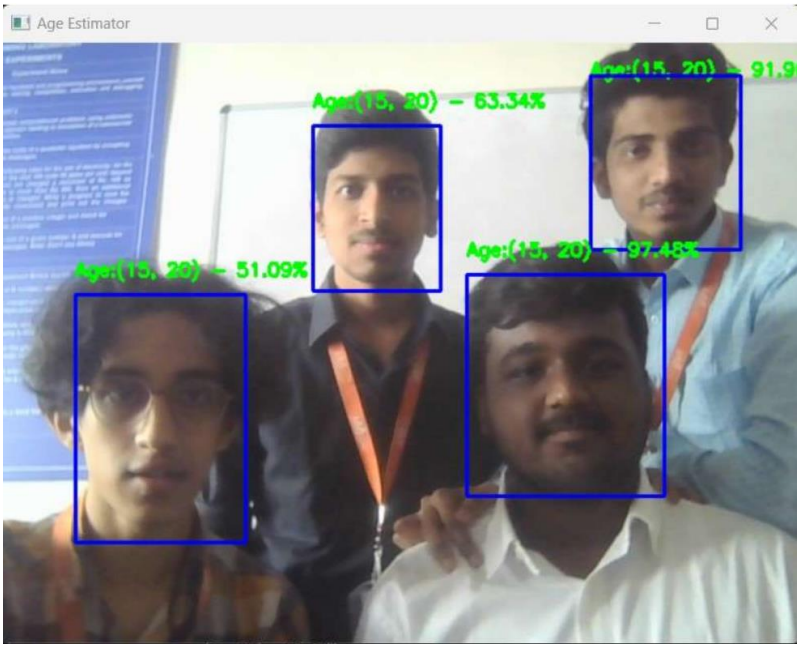
Along with computer vision, CV2 also provides .xml files called haarcascade files, they are used for detection of several attributes, for example, frontal face, hand, lips, glasses, masks, etc. For our project, the “frontal face” file was used in order to detect any faces that the camera sees in real time.
How the haarcascade works is very similar to that of a model, each haarcascade file contains a set of true images and false images, now what it does is, it compares what it sees in the camera (with a certain minimum size threshold, that way it won’t detect facial features at a pixel level) with the two sets of true and false images and puts a rectangle around what matches the true image, hence, detecting the face.
This file does not have the property of explicitly differentiating between different faces, like a facial recognition software or anything, it can not say whose name the face belongs to, nor can it store any sort of additional references on its own in order to

recognize faces, all it knows, in plain human terms, is what something looks like, and in this instance, it knows what a face looks like and alerts the user when it captures one on the camera.

Another interesting feature about the haarcascade files is that one can create their own haarcascade files and include it under the same folder of their installed OpenCV library. By simply including enough true and false images of anything, a haarcascade file can be create to detect pretty much anything, with an appropriate accuracy.

5.3 Training a Dataset

The dataset we attempted to use was called “UTKface” which is the most standard dataset for age detection. The data in the dataset was structured in the form of positive and negative folders under each age class, in each positive folder is a collection of pictures of people who fall under that age group, and in the negative folder is a collection of images of people who do not fall in that age group.
In this dataset the negative folder in the dataset consists of all the other pictures in the dataset that don’t fall in the respective true folder of each class. The dataset was trained into a model using machine learning, this follows the linear regression theory. What this does is, it takes all the true images from each class and meshes it together to create a very generic image of each class.
This is a necessity as the detected face cannot be individually compared with each of the positive and negative images in order to arrive at an age estimate, instead it is compared with each class, therefore with the respective generic model image. This saves the system a lot of computing power, memory and also a lot of time. This is verified by using the test data that also comes with the dataset.



5.4 Presenting the output.

The output is not only presented in the form of what is displayed in the camera, it is also displayed in the form of a statistic. Let us visualize age groups 0-2,4-6,8-12,15-20,25-32,38-43,48-53,60-100 are buckets and the probability that the detected face falls in each of these buckets is displayed in a way that all of the probabilities are in the form of a percentage and that all of them add up to 100%

```
============================== Face 1 Prediction Probabilities ==============================
(0, 2): 0.02%
(4, 6): 1.15%
(8, 12): 39.35%
(15, 20): 55.00%
(25, 32): 4.20%
(38, 43): 0.21%
(48, 53): 0.05%
(60, 100): 0.03%
Age:(15, 20) - 55.00%
```

Also, another interesting feature of this program is that it can detect multiple faces and give specific results for each of those faces, this opens up more applications and boosts its utility infinitely more

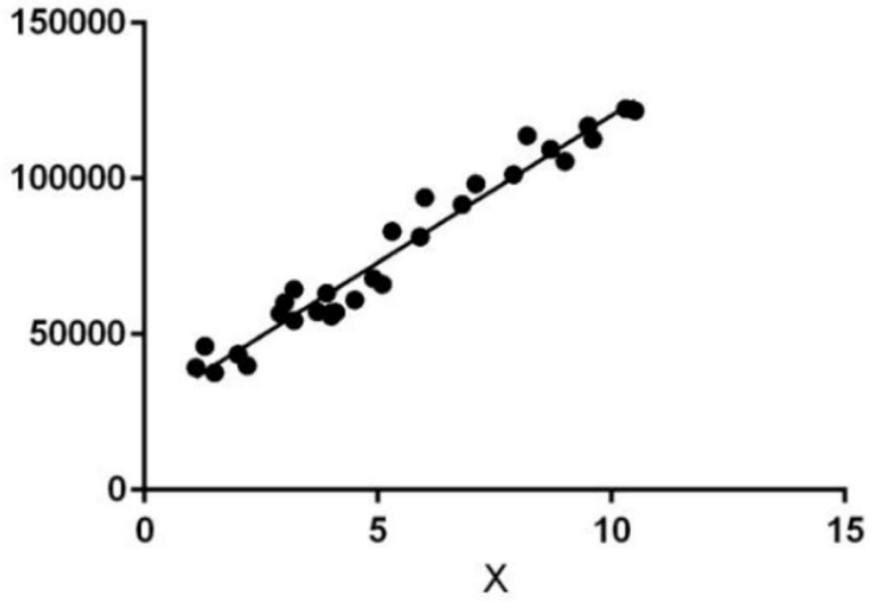
```
============================== Face 1 Prediction Probabilities ==============================
(0, 2): 0.00%
(4, 6): 0.00%
(8, 12): 12.04%
(15, 20): 61.44%
(25, 32): 26.49%
(38, 43): 0.02%
(48, 53): 0.00%
(60, 100): 0.00%
Age:(15, 20) - 61.44%
============================== Face 2 Prediction Probabilities ==============================
(0, 2): 0.02%
(4, 6): 0.16%
(8, 12): 99.31%
(15, 20): 0.23%
(25, 32): 0.27%
(38, 43): 0.00%
(48, 53): 0.00%
(60, 100): 0.00%
Age:(8, 12) - 99.31%
```

CHAPTER- 6
MODEL THEROY - LINEAR REGERESSION

Preamble: In the chapter, you get to know the final result from the travel blog website and after multiple discussions considering many countries' GDP and economy.

The basic definition of the term linear regression is using known values to predict/estimate the values of unknown variables.

For better understanding, consider an example straight line graph y=mx+c, the more y and x values we substitute and find points on the graph, the more defined that line becomes, and with more sample values we have, it becomes easier to predict where in this model an unknown value will fall. In other words, the more samples of each age group we have, the more defined the model becomes. Therefore the bigger the dataset the better, because the more samples it has the more likely it is to give the correct output.



A feature of creating a model with linear regression, is that in order to predict where an unknown variable will fall in the model, it doesn't compare it to each and every known variable, an easy way to visualize it is, the computer has a general idea and can tell where in the model, the variable will fall without comparing specifically, and based on the given datasets for the model, it can also produce values of margin of error, therefore saving a lot of time and computing power, moreover it would be impossible to accomplish the program without a model.

.

CHAPTER 7
CHALLENGES

Preamble: This chapter mainly consists of all the problems one might face while creating such a program or while executing it.

UTKface, though being the most common dataset for anything age related, simply did not have enough data to compare with and therefore was inaccurate, and there was nothing that could be done to alter the dataset any more, therefore a pretrained caffemodel was used in order to come to an age estimate.

A waitkey was enabled for the videocapture because without it, the video screen just did not stay on

The whole program runs solely on OpenCV and if there come any major changes to the library, the program could face some serious problems. In a metaphorical sense, all our eggs are in the same basket.

The recursive loop produces a percentage report for every face each time it runs the loop, causing it to be sub optimal with respect to processor usage. This also causes an irregularity with the outputs, meaning it produces multiple outputs for the same input, i.e., face.

CHAPTER 8
FILES AND REFERENCES [WEB]

Preamble: This segment contains all the files needed to create this project and also all the references.

1, Age Model

https://drive.google.com/open?id=1kiusFljZc9QfcIYdU2s7xrtWHTraHwmW

2, Age Prototxt

https://drive.google.com/open?id=1kWv0AjxGSN0g31OeJa02eBGM0R_jcjIl

3, Face Model

https://raw.githubusercontent.com/opencv/opencv_3rdparty/dnn_samples_face_detector_20180205_fp16/res10_300x300_ssd_iter_140000_fp16.caffemodel

4, Face Prototxt

https://raw.githubusercontent.com/opencv/opencv/master/samples/dnn/face_detector/deploy.prototxt

5, Javapoint (Types of Machine Learning)

https://www.javatpoint.com/types-of-machine-learning