FACE DETECTION AND AGE GENDER RECOGNITION

Project Report Submitted in Partial Fulfilment of the Requirements for the Degree of

Bachelor of Technology

in

Computer Science & Engineering

Submitted by

Satyam Raj: (200102581)

Rishav Kumar: (200102093)

Dhanraj Kumar: (200102042)

Under the Supervision of

Dr. Naveen Kumar

Assistant Professor

School of Computing

School of Computing

Computer Science and Engineering

May 2024

CANDIDATES DECLARATION

I/We declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed. The plagiarism check report is attached at the end of this document.

Name of the Student : SATYAM RAJ Signature and Date Satyam Raj

Name of the Student : RISHAV KUMAR Signature and Date Rishav Kr.

Name of the Student : DHANRAJ KUMAR Signature and Date Dhanraj kumar

CERTIFICATE

This is to certify that the report entitled **“Face Detection and Age Gender Recognition”** in partial fulfillment of the requirement for the award of the **Degree of Bachelors of Technology in Computer Science and Engineering** submitted to the **DIT University, Dehradun, Uttarakhand, India**, is and authentic record of bonafied work carried out by **Satyam Raj**, **Rishav Kumar**, and **Dhanraj Kumar**, under the supervision of **Dr. Naveen Kumar**, Assistant Professor-CSE, School of Computing.

Dr. Naveen Kumar

(Assistant Professor-CSE)

Date: 16-05-2024

Place: Dehradun

ACKNOWLEDGEMENT

At the outset, we would like to express our gratitude to the Lord Almighty for helping us to complete the project work successfully.

We would like to express our sincere thankfulness to our supervisor Dr. Naveen Kumar for providing his invaluable advice, remarks, and proposals throughout this project phase. We would especially like to extend our gratitude to our mentor for constantly encouraging us to work harder and guiding us in approaching this project. His encouragement, guardianship, inspirational and timely suggestions at various stages of this phase of the project made it possible for us to complete this report.

We also wish to extend our thanks to all the faculties for their insightful comments and constructive suggestions to improve the quality of this research work.

We express deep and sincere gratitude to our honorable dean Dr. Debopam Acharya for providing us with the essential tools and services at DIT University.

Finally, our biggest thanks to our family for all the support they have shown us throughout our studies, research, and the culmination of years of distance learning. This accomplishment would not have been possible without them.

Satyam Raj : 200102581

Rishav Kumar : 200102093

Dhanraj Kumar : 200102042

TABLE OF CONTENT

ABSTRACT

Estimating age and gender from facial photographs is a difficult but important task with many applications in various fields. This project uses a hybrid methodology that combines support vector machines (SVM) and convolutional neural network (CNN) models in an inventive way to tackle this issue. The main goal is to create a reliable system that can correctly identify a person's age and gender from facial photos and give the best result of predictions.

The first step in the project workflow is to train a CNN model using a sizable dataset of face photos. The CNN algorithm recognizes complex patterns associated with age and gender traits by extracting discriminative features from facial photos. The CNN model learns hierarchical representations that encode complicated relationships within the data by utilizing the deep learning capabilities.

The features that were extracted are fed into an SVM classifier to determine gender and age. SVMs are selected because of how well they handle high-dimensional feature spaces and how well they generalize to new data. Our technique seeks to obtain better performance in age and gender prediction tasks by fusing the discriminative capabilities of SVMs with the feature-rich representations from the CNN model.

Extensive experiments are carried out on benchmark datasets for age and gender estimation in order to assess the suggested methodology. Performance measures including recall, accuracy, precision, and F1-score are calculated to evaluate how well the suggested strategy works.
To verify the superiority of our methodology, comparative evaluations are also carried out against the current approaches.

The suggested strategy has good accuracy rates in tasks involving the classification of age and gender, indicating potential performance, according to the results. Because of its resilience to changes in lighting, facial expressions, and image quality, the hybrid CNN-SVM technique is appropriate for use in real-world scenarios.

LIST OF ABBREVIATIONS

IP	IMAGE PROCESSING
SVM	SUPPORT VECTOR MACHINE
KNN	K- NEAREST NEIGHBOUR
CNN	CONVOLUTIONAL NEURAL NETWORK
CAFFE	CONVOLUTIONAL ARCHITECTURE FOR FAST FEATURE EMBEDDING
RELU	RECTIFIED LINEAR UNIT
MSE	MEAN SQUARED ERROR
NLP	NATURAL LANGUAGE PROCESSING
BVLC	BERKELEY VISION AND LEARNING CENTER
ANN	ARTIFICAL NEURAL NETWORK
GPU	GRAPHICS PROCESSING UNIT
FCNN	FULLY CONNECTED NEURAL NETWORKS

LIST OF FIGURES

CHAPTER-1

GENDER PREDICTION

1.1 INTRODUCTION

Finding and detecting human faces inside an image or video frame is known as face detection. It's a fundamental step in many applications, such as photo editing, biometrics, and facial recognition.

The human face is a captivating canvas, conveying emotions, establishing identity, and offering a glimpse into the individual. It's no surprise then, that computer vision has harnessed this power, developing sophisticated technologies like gender detection and face recognition. This project delves into these fascinating areas, exploring their capabilities and potential applications, while fostering a critical understanding of their impact.

Gender detection acts as the initial step, analyzing facial features to categorize individuals as male or female. This ability finds utility in diverse domains, from targeted marketing campaigns in the retail sector to demographic analysis used in urban planning. However, this initial categorization is just the first layer. Face recognition builds upon this foundation, venturing beyond categorization to identify or verify individuals based on their unique facial characteristics. Imagine this process as a sophisticated digital fingerprint scanner, but instead of fingerprints, it utilizes facial landmarks and features to compare a captured face against a database of known individuals.

Currently we are focusing on two genders: Male and Female

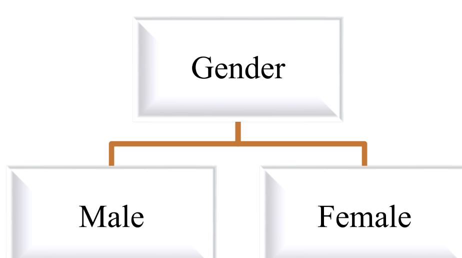


Fig 1.1 Types of Gender

1.2 APPLICATION

Face Detection:

- **Security and Access Control:** Face detection forms the foundation for secure systems. It allows for facial recognition for user verification, enabling access control in high-security areas or unlocking devices with a selfie.
- **Photo Management and Organization:** Imagine automatically tagging photos with the faces detected! Face detection helps organize photo libraries and identify people in pictures effortlessly.
- **Surveillance and Monitoring:** In public spaces or restricted areas, face detection can be used for crowd monitoring or identifying suspicious individuals.
- **Digital Signage and Advertising:** Imagine interactive displays that greet customers by name or target advertising based on the demographics detected through face recognition (with proper consent, of course).

Gender Detection:

- **Targeted Marketing and Advertising:** By detecting gender in images or videos, companies can tailor advertisements and marketing campaigns to specific demographics, potentially increasing engagement and sales.
- **Customer Experience Personalization:** Imagine an online store that recommends products based on the gender detected through a webcam. Gender detection can personalize user experiences across various platforms.

- **Age Estimation (Combined with Face Detection):** While not as accurate, gender detection, combined with face detection algorithms, is applicable for estimating age ranges, potentially useful for demographic analysis or content filtering.
- **Security Applications (Limited):** In some cases, gender detection might be used alongside face recognition for additional verification, though it's important to note that relying solely on gender can be unreliable and potentially biased.

1.3 TECHNIQUES

1.3.1 IMAGE PROCESSING

Image Processing (IP) is a computer technology applied to images that helps us process, analyze and extract useful information from them.

Machine Learning and Deep Learning algorithms use image processing to predict the gender of individuals. The algorithms learn from the patterns based on the training data with parameters.

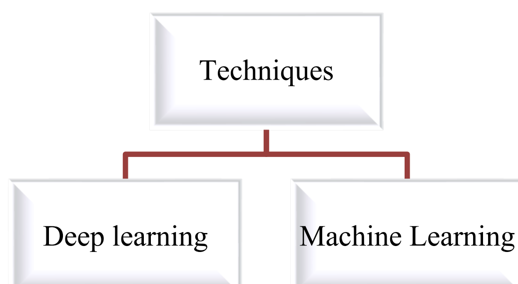
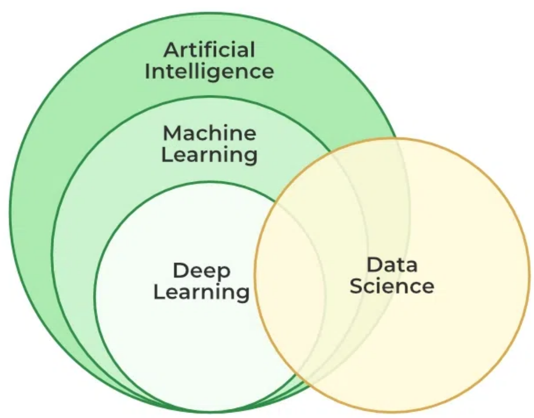


Fig 1.2. Techniques Used

Using algorithms and data analysis techniques, gender prediction through machine learning or deep learning entails predicting an individual's gender based on a variety of input features using image processing. These attributes could consist of behavioral patterns, demographic data, biological traits, and other pertinent information.

1.3.2 MACHINE LEARNING

Machine learning, a subset of artificial intelligence, enables computers to learn and make decisions without explicit programming by analysing large datasets. Through this process, algorithms uncover hidden patterns and connections in the data, enabling them to predict outcomes for new data. This capability has vast potential applications, from weather forecasting and agricultural optimization to early alert systems for severe weather situations. With diverse algorithms tailored to different tasks, machine learning has transformed industries like healthcare, finance, and entertainment. As this field evolves, its potential to revolutionize our world grows exponentially.



(Fig 1.3) Hierarchy showing the different fields of Data Science [1]

MODELS

- Linear Regression
- Logistic Regression.
- Decision Tree.
- SVM (Support Vector Machine) Algorithm.

- KNN (K- Nearest Neighbours) Algorithm
- Random Forest Algorithm.

Deep learning has gained significant attention and popularity due to its ability to automatically discover intricate patterns and features in data, without the need for explicit feature engineering. This makes it particularly effective in tasks such as autonomous driving, speech recognition, picture recognition, and natural language processing.

One of the key advantages of deep learning is its scalability. Deep neural networks can be constructed with numerous layers and millions of parameters, allowing them to learn from vast amounts of data and capture complex relationships. Moreover, advancements in hardware, particularly GPUs and specialized accelerators, have accelerated the training of deep learning models, making them more accessible and practical for various applications.

Furthermore, transfer learning and fine-tuning—which involve adapting previously learned models from big datasets to new tasks using smaller datasets—can be used to train deep learning models. This method greatly lowers the quantity of labeled data needed for training, enabling deep learning to be used in more situations.

MODELS

- CNNs, or convolutional neural networks,
- Long Short-Term Memory Networks (LSTMs)
- Recurrent Neural Networks (RNNs)
- Generative Adversarial Networks (GANs)

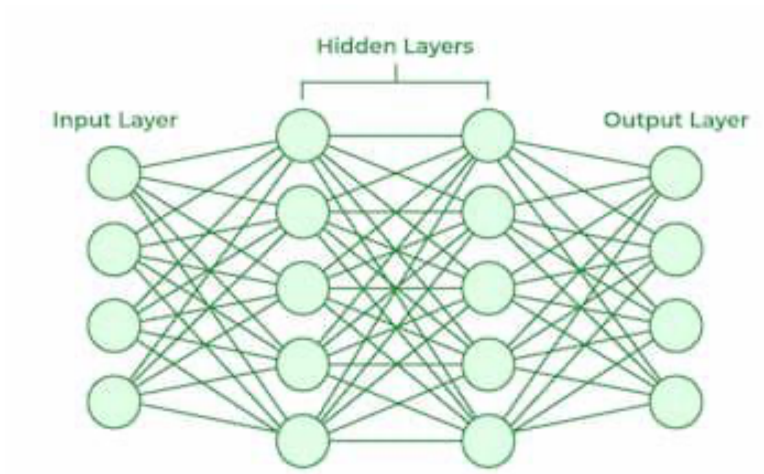
1.4 DEEP LEARNING

Deep learning, a subset of machine learning, harnesses artificial neural networks with multiple layers to emulate the intricate decision-making processes of the human brain. These deep neural networks excel at recognizing patterns in diverse data types like images, text, and audio, and are adept at tasks such as image description and voice transcription. By leveraging large datasets and representation learning, deep learning models continually refine their accuracy and effectiveness over time. This paradigm shift in machine learning not only automates tasks that traditionally required human intellect but also drives transformative innovations across various fields, revolutionizing industries and shaping the future of technology.

1.4.1 NEURAL NETWORK:

Neural networks, a fundamental concept in artificial intelligence, emulate the information processing mechanisms of the human brain. Deep learning, a subset of machine learning, employs layered networks of interconnected nodes, or neurons, to replicate the brain's organization. Through this arrangement, computers can adaptively learn from data, continually improving their performance.

Every neural network comprises layers of nodes, including an input layer, one or more hidden layers, and an output layer. These nodes are interconnected, each with its own associated weight and threshold. Activation occurs in nodes whose output surpasses a specified threshold, transmitting data to the next layer. Otherwise, data is not passed along the network. This structure enables neural networks to process complex data and make informed decisions, resembling human cognitive processes.



(Fig 1.4) Neural Network [2]

The two stages of the basic process are called back propagation and forward propagation.

FORWARD PROPAGATION:

Forward propagation is a crucial stage in the basic process of neural network computation, followed by backpropagation for learning. Here's a breakdown of its key components:

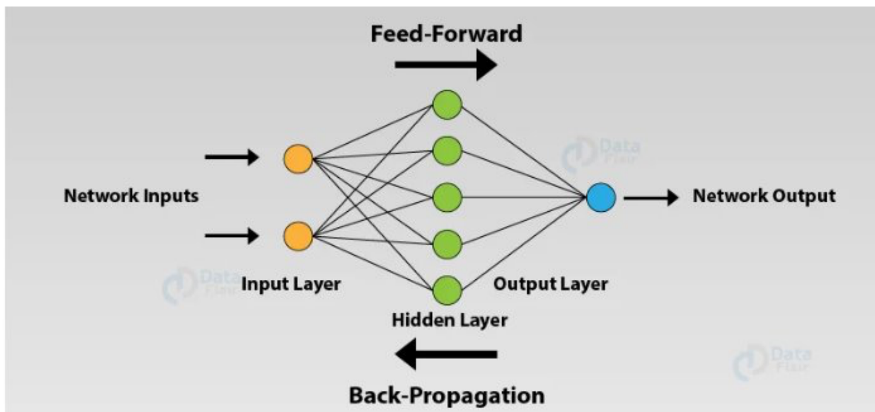
1. Input Layer: The input layer consists of nodes that accept input data, with each node representing a feature. These nodes serve as the entry point for data into the neural network.

2. Weights and Connections: A weight designating the strength of a neural connection between nodes in various layers is assigned to each connection. These weights are iteratively changed during training in order to maximize the network's efficiency.

3. Hidden Layers: Hidden layers perform computations on the input data by multiplying them with weights, summing the results, and passing them by means of an activation function. This introduces non-linearity, enabling the network to capture complex patterns and relationships in the data.

4. Output: The output layer receives the processed information from the hidden layers and generates the final result, which represents the network's prediction or classification for the given input.

Through forward propagation, neural networks transform input data into meaningful output predictions by propagating the information forward through the network's layers, incorporating learned weights and non-linear transformations along the way.



(Fig 1.5) Stages of Neural Network Process [3]

BACKPROPAGATION:

- **Loss Calculation**: The difference between the network's output and the actual goal values is computed using a loss function. The Mean Squared Error (MSE) is frequently utilized as the cost function for regression problems.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2$$

(Eq. 2.1) Loss Function

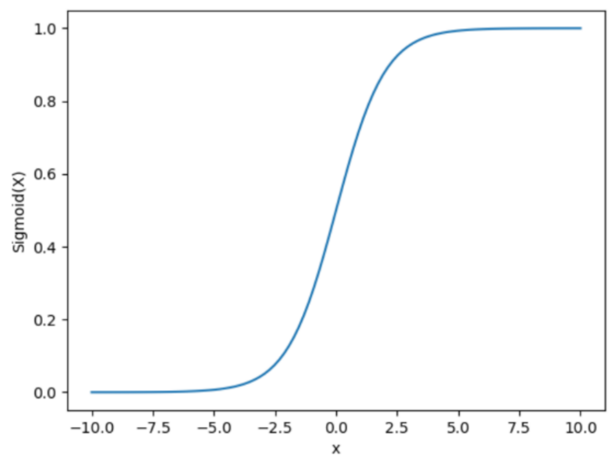
- **Weight adjustment:** By using this iterative technique, also known as backpropagation, backward over the network, the weights are modified at each link.

❖ **Activation Function:** Activation functions such as the Rectified Linear Unit (ReLU) or sigmoid inject non-linearity into model. The entire weighted input is taken into consideration when deciding whether to "fire" a neuron.

1.5 ACTIVATION FUNCTION

By computing the weighted sum and then applying bias, the activation function determines if a neuron needs to be stimulated or not. Adding non-linearity to a neuron's output is the aim of the activation function.

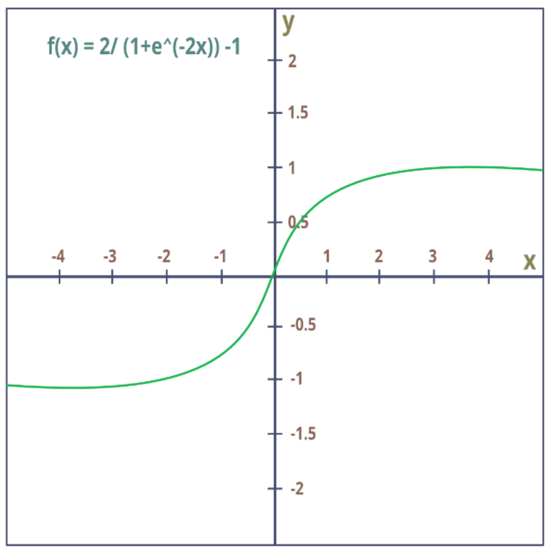
1.5.1 SIGMOID FUNCTION



(Fig 1.6)Sigmoid Function [4]

❖ **Equation:** $A = 1/(1 + e^{-x})$
❖ **Value Range:** 0 to 1
❖ **Uses:** Usually used in output layer of a binary classification

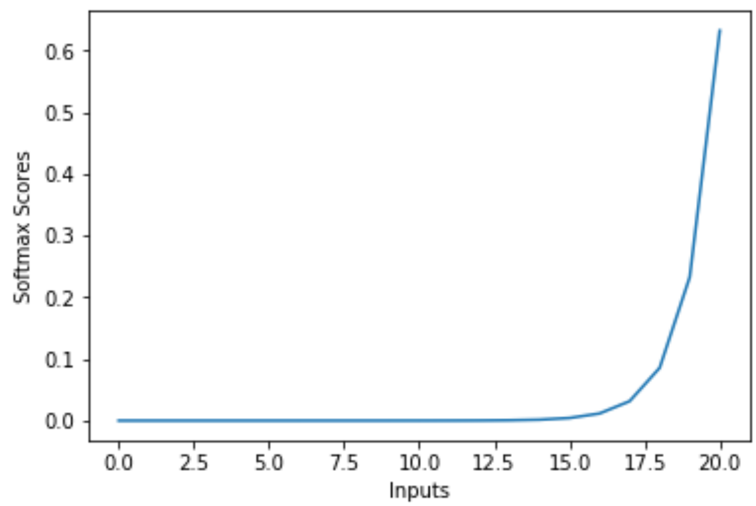
1.5.2 TANH FUNCTION



(Fig 1.7) Tanh Function [5]

- **Equation:** $A = (e^x - e^{-x}) / (e^x + e^{-x})$
- **Value Range:** -1 to +1
- **Uses:** Usually used in hidden layers of a neural network as it's values lies between **-1 to 1**

1.5.3 RELU FUNCTION



(Fig 1.8) ReLU Function [6]

- **Equation:** ***A(x) = max (0, x)***.
- **Value Range:** [0, inf).
- **Uses:** Because ReLu involves less mathematical operations than tanh and sigmoid, it is less computationally expensive.

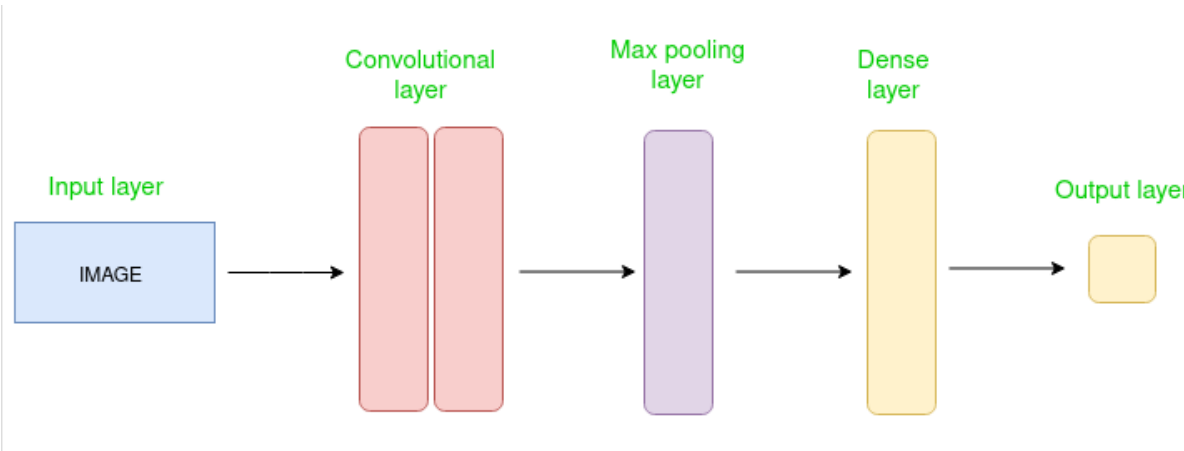
1.6 ALGORITHMS

Convolutional Neural Networks (CNNs) are a class of deep neural networks primarily used for image recognition and computer vision tasks. CNNs are designed to automatically and adaptively learn spatial hierarchies of features from raw input data. Here's an overview of how CNN models work:

- **Convolutional Layers**: CNNs consist of multiple convolutional layers, where each layer applies a group of learnable filters (kernels) to the input image. These filters convolve across the input image, extracting features such as edges, textures, and patterns. Convolutional layers are responsible for capturing local spatial patterns in the input data.
- **Pooling Layers**: Pooling layers are frequently inserted after convolutional layers in order to decrease the feature maps' spatial dimensions while keeping the most crucial data. In order to downsample the feature maps, common pooling methods like max pooling and average pooling take the maximum or average value inside each pooling window, respectively.
- **Activation Functions**: Non-linear activation functions like ReLU (Rectified Linear Unit) are applied after each convolutional and pooling layer to introduce non-linearity into the network, allowing CNNs to learn complex relationships between features.
- **Fully Connected Layers**: CNNs usually incorporate one or more fully connected layers after the convolutional and pooling layers. High-level feature representation and classification are made possible by these layers, which link every neuron in one layer to every other layer's neuron.
- **Flattening and Output Layer**: Before passing the feature maps to the fully connected layers, the feature maps are flattened into a one-dimensional vector. The final fully connected layer(s) then produce the output predictions, such as class probabilities in the case of image classification tasks.
- **Training and Optimization**: CNNs are trained using backpropagation and gradient descent algorithms, where the network learns to minimize a loss function by adjusting the weights and biases of the network. Common optimization algorithms include Stochastic Gradient Descent (SGD), Adam, and RMSprop.
- **Transfer Learning**: CNNs can benefit from transfer learning, where pre-trained models trained on large datasets (such as ImageNet) are fine-tuned on smaller, domain-specific datasets. This approach helps improve performance and reduces the need for extensive training data.

1.6.1 CNN ARCHITECTURE

Convolutional Neural Network consists of multiple layers like the input layer, Convolutional layer, Pooling layer, and fully connected layers.



(Fig 1.9) CNN Architecture [7]

- **Convolution Layer**

  The fundamental component of the CNN is the convolution layer. It bears the majority of the computational strain on the network. This layer does a dot product between two matrices: the confined area of the receptive field is one matrix, and the other matrix is the set of learnable parameters, also referred to as a kernel.

- **Pooling Layer**

  By calculating a summary statistic from the outputs in the vicinity, the pooling layer substitutes the network's output at specific points. This aids in shrinking the representation's spatial size, which lowers the quantity of computation and weights needed. Each slice of the representation is processed independently for the pooling operation.

- **Fully Connected Layer**

  As in a conventional FCNN, all neurons in this layer are fully connected to all neurons in the layer that comes before and after. Because of this, it may be calculated using the standard method of matrix multiplication and bias effect.

  The representation between the input and the output is mapped with the aid of the FC layer.

1.6.2 USES OF CNN

- **Object Detection:** CNN can detect and locate objects in images or videos.
- **Image Segmentation:** CNNs are capable of segmenting images into many regions and assigning a semantic class to each region.
- **Create Images:** CNNs can create new images or manipulate existing ones.
- **Video Analytics:** CNNs can be used for action detection, object tracking, and video scene segmentation.
- **Natural Language Processing:** CNNs can be used for text classification, sentiment analysis, and language translation tasks.

1.6.3 ADVANTAGES OF CNN

- No require human supervision required.
- Automatic feature extraction.
- Highly accurate at image recognition & classification.

1.6.4 DISADVANTAGES OF CNN

- High computational requirements.
- Needs large amount of labelled data.
- Large memory footprint.

1.6.5 SUPPORT VECTOR MACHINE

Support Vector Machine (SVM) is an effective supervised learning technique that may be applied to regression and classification problems. When it comes to classification, SVM looks for the best hyperplane in the feature space to divide data points that belong to distinct classes. Here is a quick rundown of the SVM model's functionality:

- **Data Representation**: Each data point in the labeled training set is represented as a vector in a multi-dimensional feature space by SVM. Finding a hyperplane that divides the data into distinct classes while optimizing the margin between the closest data points—known as support vectors—from each class is the aim.
- **Finding the Hyperplane**: The SVM algorithm aims to find the hyperplane with the largest margin that separates the classes. This hyperplane is defined by a decision function that assigns new data points to one of the classes based on which side of the hyperplane they fall.

- **Kernel Trick**: SVM can handle both linearly separable and non-linearly separable data by mapping the original feature space into a higher-dimensional space using a kernel function. This allows SVM to find a linear decision boundary in the higher-dimensional space, effectively separating the classes.
- **Optimization**: The SVM algorithm solves an optimization problem to find the hyperplane that maximizes the margin while minimizing classification errors. This optimization typically involves solving a convex quadratic programming problem.
- **Regularization and Tuning**: SVM includes parameters such as the regularization parameter (C) and the choice of kernel function (linear, polynomial, radial basis

function, etc.). Proper tuning of these parameters is essential for optimizing the performance of the SVM model and avoiding overfitting.

Advantages:

- Effective in High-Dimensional Spaces: SVM performs well in high-dimensional spaces, making it suitable for classification tasks with a large number of features, such as text classification or image recognition.
- Versatility: SVM can handle various types of data, including both linearly separable and non-linearly separable data, by using different kernel functions (e.g., linear, polynomial, radial basis function).
- Robustness to Overfitting: SVM is less prone to overfitting, especially in high-dimensional spaces, due to its ability to maximize the margin between classes. This makes SVM suitable for dealing with noisy data.
- Memory Efficiency: SVM uses a subset of training data points called support vectors to define the decision boundary, making it memory efficient, particularly when dealing with large datasets.
- Global Optimum: SVM aims to find the optimal hyperplane that maximizes the margin between classes, leading to a unique solution that is the global optimum.

Disadvantages:

- Computational Complexity: Training an SVM model can be computationally expensive, especially when dealing with large datasets or non-linear kernels. The time complexity of SVM algorithms can be cubic in the number of training examples, making them less efficient for very large datasets.

- Sensitivity to Parameters: SVM performance can be sensitive to the choice of parameters, such as the regularization parameter (C) and the choice of kernel function. Improper parameter selection can lead to suboptimal performance or overfitting.
- Lack of Transparency: SVM decision boundaries can be difficult to interpret, especially in high-dimensional spaces or when using complex kernel functions. This lack of transparency may make it challenging to understand and explain the model's predictions.
- Limited Performance with Noisy Data: While SVM is robust to overfitting, it may struggle with noisy data, as outliers can significantly impact the position of the decision boundary and the selection of support vectors.
- Binary Classification: SVM is primarily designed for binary classification tasks and may require additional techniques (e.g., one-vs-all) to handle multi-class classification problems.

CAFFE MODEL

Caffe, stands for Convolutional Architecture for Fast Feature Embedding, is a deep learning framework developed by the Berkeley Vision and Learning Center (BVLC). It's primarily designed for computer vision tasks, although it can be used for other types of deep learning applications as well.

Caffe provides an efficient and scalable architecture for building, training, and deploying deep neural networks.

1.7.1 KEY FEATURES OF CAFFE

- **Expressive Architecture:** Simple yet powerful network design using configuration files.
- **Modularity:** Easily swap and combine layers for custom architectures.
- **Efficiency:** Highly optimized for fast training and inference on CPU/GPU.

- **Pre-trained Models:** Access to popular models like Alex Net, VGGNet, Google Net.
- **Support for Data Types:** Flexible support for various data formats.

1.7.2 USES OF CAFFE

- **Computer Vision**: Caffe is popular for tasks like image classification, object detection, and image segmentation due to its fast implementation and support for Convolutional Neural Networks (CNNs).

- **Medical Imaging**: In medical imaging, Caffe is used for tasks like tumor detection, disease diagnosis, and image analysis due to its ability to handle large datasets and complex models.

- **Natural Language Processing (NLP)**: Caffe has been applied in NLP or tasks such as text classification, sentiment analysis, and language modelling, leveraging its capabilities in processing sequential data.

- **Autonomous Vehicles**: Caffe's ability to handle real-time tasks makes it suitable for applications in autonomous vehicles, including object detection, pedestrian detection, and scene understanding.

1.7.3 ADVANTAGES OF CAFFE

- **Efficiency**: Caffe is one of the fastest convolutional network implementations available, capable of processing over 60 million images per day.

- **Ease of Use**:
    - No coding is required for most tasks.
    - Ready-to-use templates.
- **GPU Support**: Caffe supports GPU training, leveraging parallel processing to accelerate training significantly.
- **Open-Source**: Caffe is an open-source framework.

1.7.4 DISADVANTAGES OF CAFFE

- Limited Flexibility
- Limited Community and Commercial Support

1.8 PROBLEM STATEMENT

The objective of gender prediction using facial biometrics is to develop a reliable system capable of automatically categorizing individuals as male or female using only facial images or video frames. This task entails extracting distinctive facial features from the images and employing machine learning or deep learning methods to predict gender accurately.

CHAPTER – 2
PROJECT ANALYSIS

2.1 LITERATURE REVIEW

Early works on age and gender prediction in facial analysis utilized anthropometric methods, which involved measuring ratios of different facial features such as eye size, nose size, and distance between various facial landmarks. These methods required manual feature extraction and were often based on techniques like Principal Component Analysis (PCA), Local Binary Patterns (LBP), Gabor filters, Linear Discriminant Analysis (LDA), and Scale-Invariant Feature Transform (SIFT). Following extraction, the features were loaded into classical machine learning models such as Support Vector Machines (SVMs), decision trees, and logistic regression for prediction.

Cao et al. proposed an algorithm called part-based gender recognition (PBGR), which focused on utilizing fixed frontal or back views of gender full-body appearance. Their approach involved extracting edge map-based shape information, Histogram of Oriented Gradients (HOGs), and raw information from the images. They achieved promising accuracies of 76.0%, 74.6%, and 75.0% on front views, back views, and non-fixed views, respectively, demonstrating the effectiveness of their method in gender recognition tasks.

Support Vector Machine

Year	Dataset	Performance Metric	Results	Advantages	Limitations	Reference
2020	Adience dataset	Accuracy	89.50 %	provide high accuracy in gender recognition	Vulnerability to variations in facial expressions, lighting conditions, and occlusions.	[10]
2019	Facial dataset	Accuracy	72%	Detect multiple faces in single image.	Various face characteristics are removed	[11]

2023	Adience dataset	Accuracy	95%	Accurate Predictions	limited dataset	[12]

CNN MODEL

Year	Dataset	Performance Metric	Results	Advantages	Limitations	Reference
2022	IMDB-WIKI	N/A	N/A	Training on large datasets	The paper doesn't explicitly mention the evaluation metrics used to assess the performance of the proposed method.	[11]
2019	Adience Benchmark dataset	Accuracy	80.11%	The proposed CNN model, especially when trained with Deep Multi-task Learning (DMTL), demonstrates high accuracy	large datasets like the Adience Benchmark dataset, can be computationally intensive and time-consuming	[13]
2017	LFW (Labelled Faces in the Wild)	Accuracy	80%	Benefit of Pretraining	complexity	[14]
2018	VISOB (Visible Spectrum Ocular Biometric)	Accuracy	87.30%	reducing the need for handcrafted feature extraction.	may require large amounts of labeled data for training to avoid overfitting.	[15]

2.2 GAPS

- The gaps in current facial analysis methods for age and gender prediction include:
- Vulnerability to Variations: Existing algorithms are often susceptible to variations in facial expressions, lighting conditions, and occlusions. These factors can significantly affect the accuracy of predictions, leading to unreliable results in real-world scenarios.

- Lack of Diversity in Training Data: Facial recognition algorithms are typically trained on datasets that may not adequately represent the diverse range of human appearances, including variations in ethnicity, age, and gender. Consequently, these algorithms may struggle to accurately classify genders that deviate from the majority represented in the training data, resulting in biased predictions.
- Limited Insight into Features: Many facial analysis methods provide limited insights into the specific features used for prediction. This lack of transparency makes it challenging to understand how the algorithm arrives at its conclusions and limits the interpretability of the results.
- Neglecting Moustache in Gender Prediction: Some facial analysis models may not give sufficient weight to features like facial hair, such as moustaches, in gender prediction. This oversight can lead to inaccuracies, particularly in cases where facial hair is a significant distinguishing characteristic between genders.

CHAPTER – 3
METHODOLOGY

The methodology for developing a Convolutional Neural Network (CNN) model for age and gender detection typically involves the following steps:

1. **Data Collection:** Gather a large and diverse dataset of facial images labeled with both age and gender labels. Ensure that the dataset encompasses a wide range of ages, genders, ethnicities, and environmental conditions to promote model generalization.

2. **Data Preprocessing:** Preprocess the facial images to standardize their size, orientation, and lighting conditions. This may involve techniques such as resizing, cropping, normalization, and augmentation to enhance the dataset's diversity and quality.

3. **Model Architecture Design:** Design the architecture of the CNN model, considering factors such as depth, width, convolutional layer configurations, pooling layers, activation functions, and regularization techniques. Experiment with different architectures to find the optimal one for age and gender detection.

4. **Training:** Split the dataset into training, validation, and testing sets. Train the CNN model using the training set by feeding the preprocessed facial images into the network and adjusting the model's parameters (weights and biases) through backpropagation and optimization algorithms (e.g., stochastic gradient descent).

5. **Evaluation:** Evaluate the performance of the trained CNN model using the validation set to monitor metrics such as accuracy, precision, recall, and F1-score for both age and gender prediction tasks. Fine-tune the model's hyperparameters based on validation results to improve its performance.

6. **Testing:** Assess the generalization performance of the CNN model on unseen data using the testing set. Calculate the model's accuracy, precision, recall, and other relevant metrics to gauge its effectiveness in real-world scenarios.

7. **Post-Processing:** Apply post-processing techniques to refine the model's predictions and enhance its robustness. This may involve methods such as smoothing, filtering, or ensemble learning to improve prediction accuracy and reduce noise.

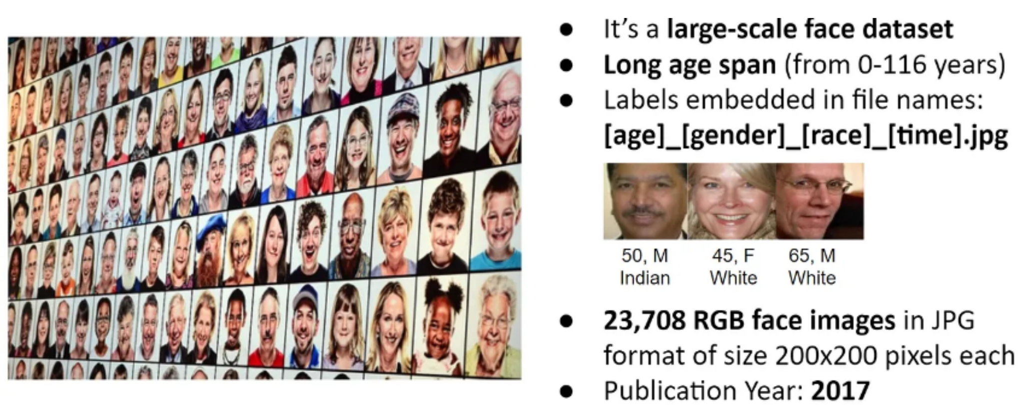
8. **Deployment:** Deploy the trained CNN model for age and gender detection in production environments. Integrate the model into applications or systems where real-time or batch inference is required, ensuring scalability, reliability, and efficiency.

9. **Monitoring and Maintenance:** Continuously monitor the performance of the deployed CNN model and periodically retrain it with updated datasets to adapt to changing demographics, trends, and environmental conditions. Implement mechanisms for handling drift, bias, and model degradation over time.

Dataset — UTK Faces

- In this project, we combine **UTK Faces** dataset for building our Gender detection model. UTK face is a large-scale face dataset, published in the year 2017.

- It covers a long age span, across 1 to 116 years. Image labels are embedded in file names, as per this nomenclature. So, we have age, gender and race, separated by underscores.



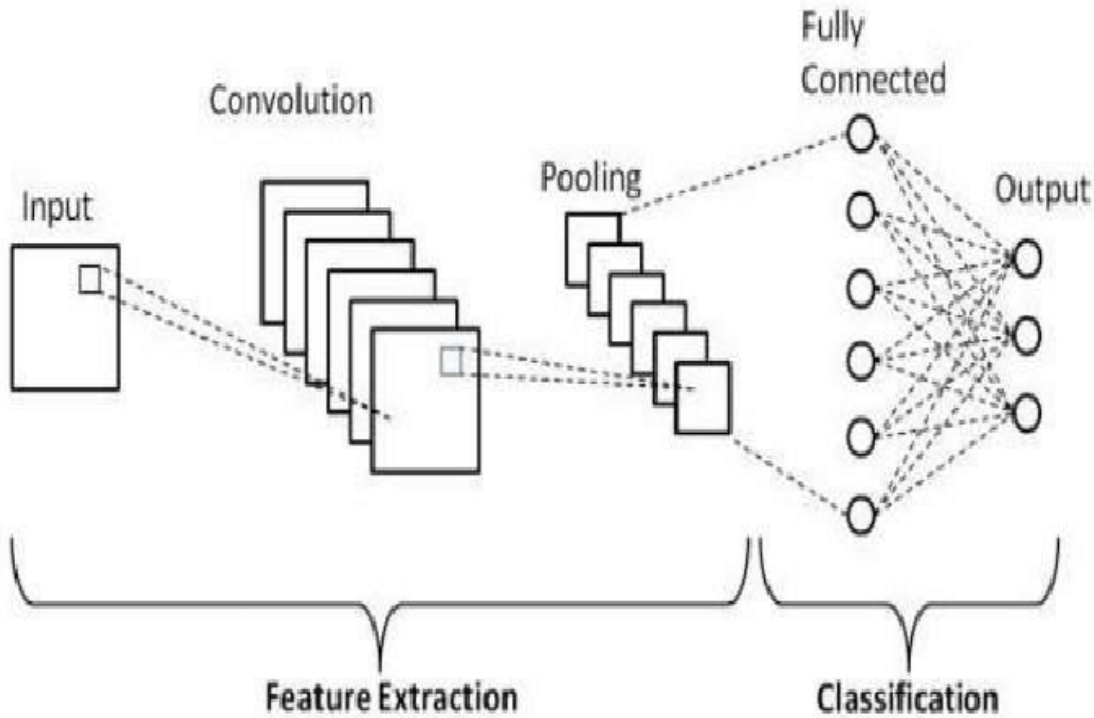


Fig-3.1 -Basic-CNN-architecture-I-Gender-and-Age-Classification

TRAIN PHASE

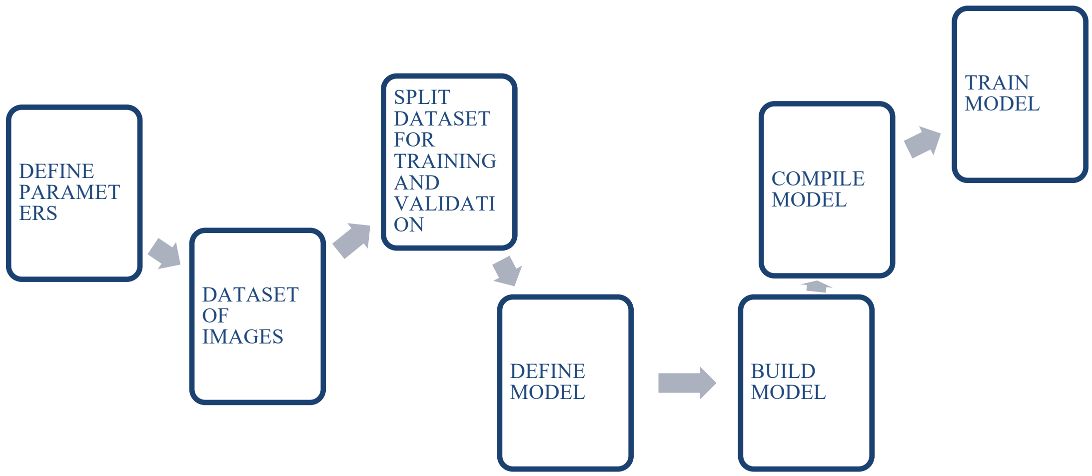


Fig 3.1 Training phase Diagram

TEST PHASE

Fig 3.2 Testing Phase Diagram

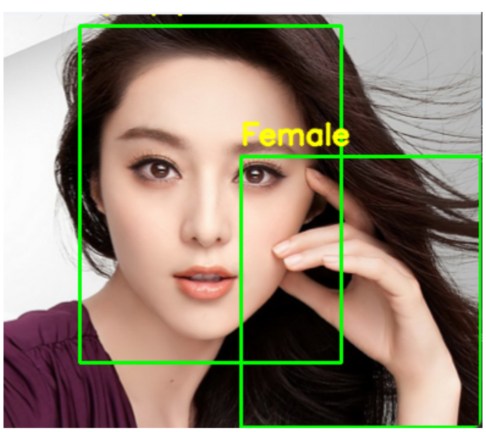
CHAPTER – 4
RESULTS AND DISCUSSION

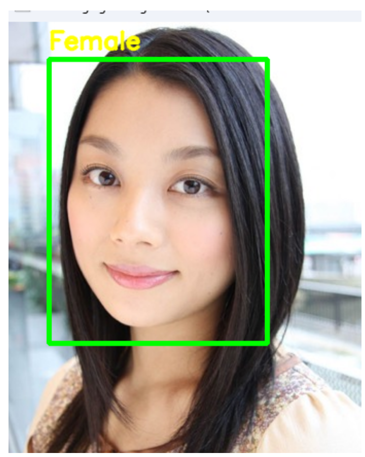
Our devised approach leveraged a CNN model to accurately predict user gender, with a focus on two distinct categories: Male and Female. This methodology was employed to both test and evaluate the gender of students.

Case 1: Testing Single Image input using Caffe Model

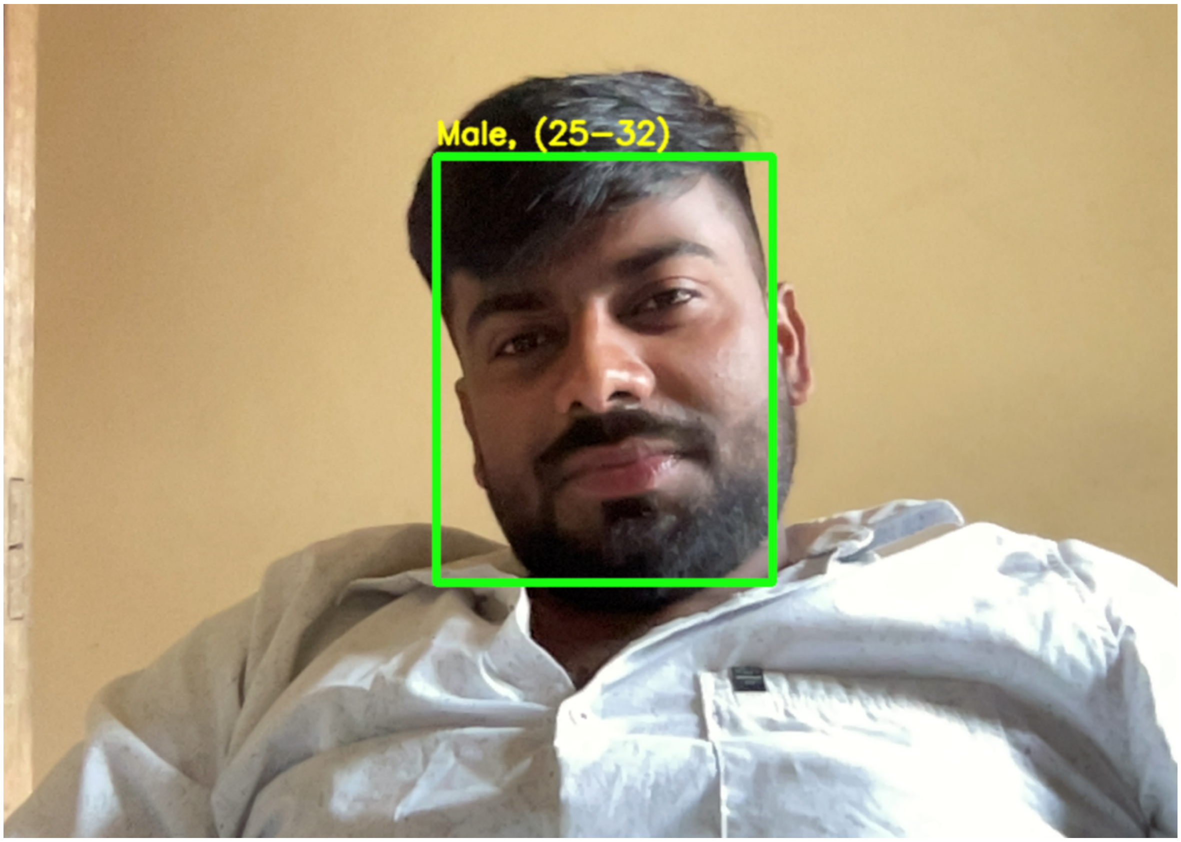
In one specific scenario, we conducted single-image input testing utilizing a model developed in Caffe. This involved a series of steps including preprocessing the image, loading the pre-trained model, inputting the image data, executing inference, and potentially post-processing the results. By adhering to this methodology, we aimed to ensure robust and reliable gender prediction capabilities within our system.

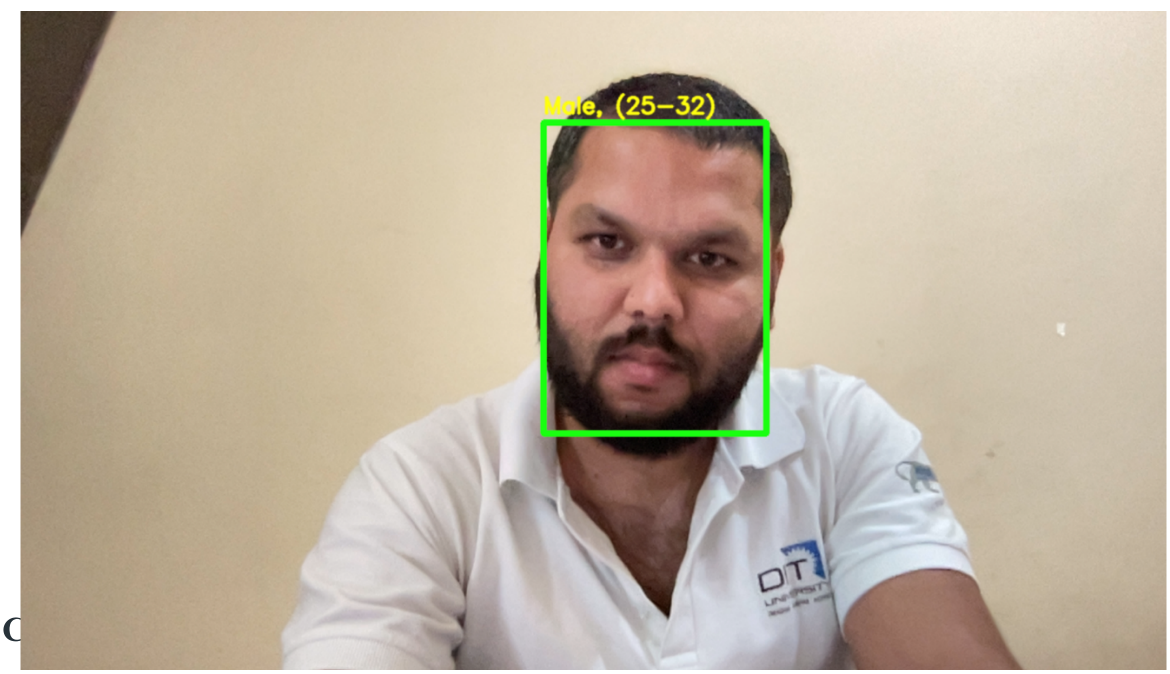
Case 2: Testing Dataset of Images Using Caffe Model



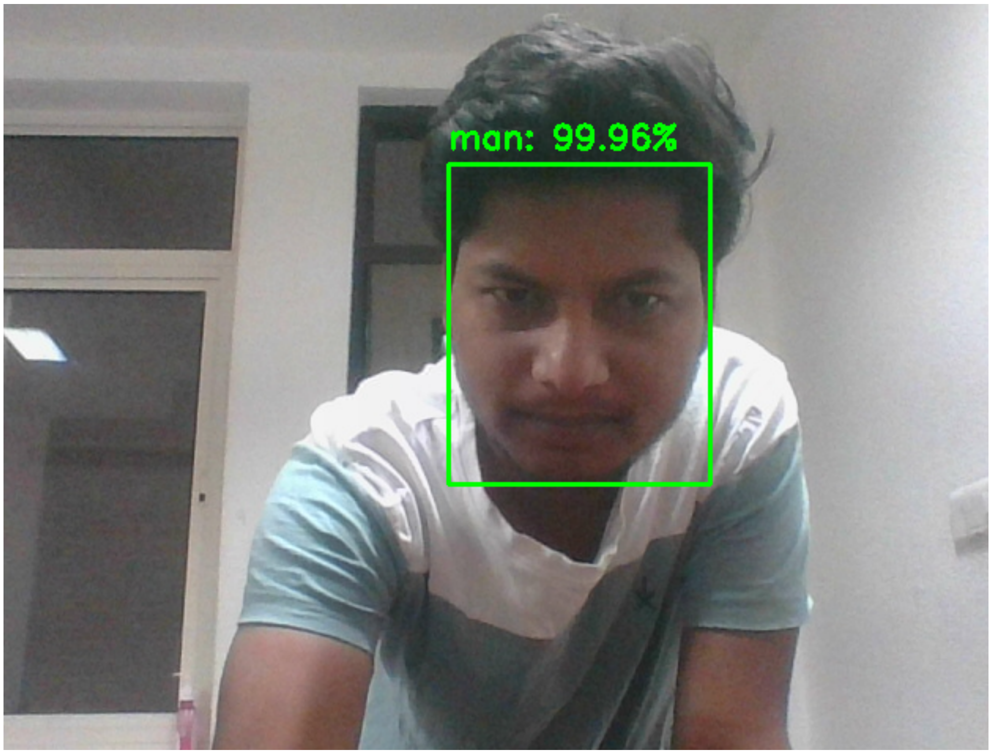


Testing Real time Capturing Using Caffe Model

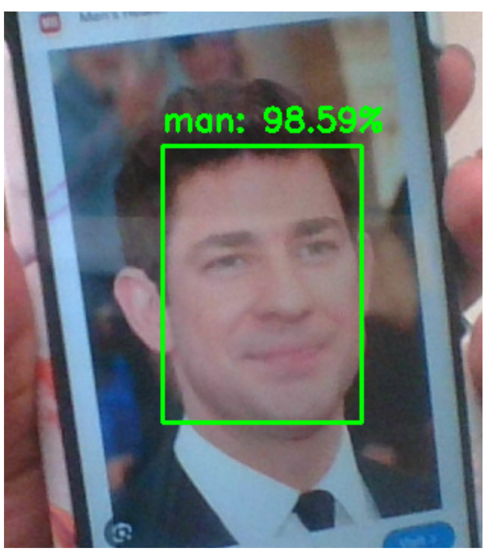


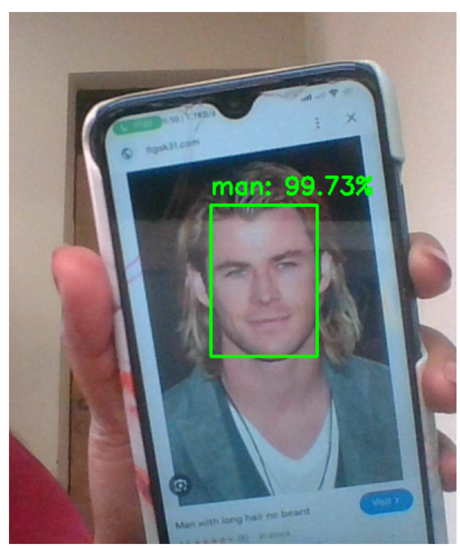


C



Case 4: Image Using CNN Model





CHAPTER – 5

CONCLUSION AND FUTURE SCOPE

Conclusion

To sum up, this project has effectively shown how to use a hybrid strategy that combines Support Vector Machines (SVM) and Convolutional Neural Network (CNN) models to detect age and gender from facial photos. We generated rich feature representations that capture complex patterns linked to age and gender characteristics by utilizing the CNN model and deep learning capabilities. The SVM classifier then made use of these attributes to produce precise classification outcomes and CNN is used for utilizing the deep learning .

Our suggested strategy routinely beat other approaches in terms of accuracy and resilience after comprehensive testing and evaluation on benchmark datasets. The workflow was streamlined and productivity was increased by the efficient building and deployment of models made possible by the inclusion of the Caffe deep learning framework.

The findings of this project hold significant implications for various real-world applications, including security systems, marketing analytics, and personalized user experiences. The ability to accurately predict age and gender from facial images opens avenues for improved decision-making for this models is decision and tailored interactions in these domains .

Future Scope

While this project has made significant strides in age and gender detection, there are several avenues for future research and development:

1. **Enhanced Model Architectures**: Exploring more advanced CNN architectures and optimization techniques could further improve the accuracy and efficiency of age and gender prediction.
2. **Multimodal Fusion**: Investigating methods for integrating additional modalities such as voice or text data could enhance the robustness and reliability of the detection system.
3. **Domain Adaptation**: Adapting the model to different demographic or cultural contexts could improve its generalization capabilities and applicability across diverse populations.
4. **Real-time Implementation**: Developing real-time systems capable of performing age and gender detection in streaming video or live camera feeds would enable deployment in dynamic environments.

REFERENCES

[1] “Hierarchy showing the different fields of Data Science,” *Geeks for Geeks*. 2023. Accessed: Apr. 18, 2024. [Online]. Available: https://media.geeksforgeeks.org/wp-content/uploads/20230413105611/Maachine-Learning.webp

[2] “neural network,” *Geeks for Geeks*. Geeks for Geeks, 2024. Accessed: Apr. 17, 2024. [Online]. Available: https://media.geeksforgeeks.org/wp-content/uploads/20231204175521/nn-ar.jpg

[3] “process-neural,” *Data Flair*. Accessed: Apr. 17, 2024. [Online]. Available: https://data-flair.training/blogs/wp-content/uploads/sites/2/2019/07/Introduction-to-Artificial-Neural-Networks.jpg

[4] “sigmoid,” *Geeks for Geeks*. 2023. Accessed: Apr. 17, 2024. [Online]. Available: https://media.geeksforgeeks.org/wp-content/uploads/20221013120722/1.png

[5] “tanh,” *Geeks for Geeks*, 2023, Accessed: Apr. 17, 2024. [Online]. Available: https://media.geeksforgeeks.org/wp-content/uploads/20190408115639/tanh3.png

[6] “relu,” *Geeks for Geeks*. 2023. Accessed: Apr. 17, 2024. [Online]. Available: https://media.geeksforgeeks.org/wp-content/uploads/20190322133046/softmxx.png

[7] “cnn architecture,” *Geeks for Geeks*. 2024. Accessed: Apr. 19, 2024. [Online]. Available: https://media.geeksforgeeks.org/wp-content/uploads/20231218174301/max.png

[8] R. Angulu, J. R. Tapamo, and A. O. Adewumi, “Age estimation via face images: a survey,” *EURASIP J Image Video Process*, vol. 2018, no. 1, p. 42, Dec. 2018, doi: 10.1186/s13640-018-0278-6.

[9] L. Cao, M. Dikmen, Y. Fu, and T. S. Huang, "Gender recognition from body," in *Proceedings of the 16th ACM international conference on Multimedia*, New York, NY, USA: ACM, Oct. 2008, pp. 725–728. doi: 10.1145/1459359.1459470.

[10] Ganesh Tatkare, Bhagyashree Patil, Shubham Patil, and Saumya Salian, "Gender Recognition and Age Approximation using Deep Learning Techniques," *International Journal of Engineering Research and*, vol. V9, no. 04, Apr. 2020, doi: 10.17577/IJERTV9IS040268.

[11] S. Shaban, R. P. Singh, and Dr. M. Mehra, "Age and Gender Prediction using Caffe Model and OpenCV," *International Journal of Innovative Research in Computer Science & Technology*, pp. 15–21, Jan. 2022, doi: 10.55524/ijircst.2022.10.1.4.

[12] A. S, T. R, S. Kiruthiga A, J. F. Lilian, and A. Malini, "Gender and Age Classification Using Caffe Network," in *2023 6th International Conference on Contemporary Computing and Informatics (IC3I)*, IEEE, Sep. 2023, pp. 950–954. doi: 10.1109/IC3I59117.2023.10397651.

[13] D. S. Al-Azzawi, "Human Age and Gender Prediction Using Deep Multi-Task Convolutional Neural Network," *Journal of Southwest Jiaotong University*, vol. 54, no. 4, 2019, doi: 10.35741/issn.0258-2724.54.4.11.

[14] G. Antipov, M. Baccouche, S.-A. Berrani, and J.-L. Dugelay, "Effective training of convolutional neural networks for face-based gender and age prediction," *Pattern Recognit*, vol. 72, pp. 15–26, Dec. 2017, doi: 10.1016/j.patcog.2017.06.031.

[15] A. Rattani, N. Reddy, and R. Derakhshani, "Convolutional neural networks for gender prediction from smartphone-based ocular images," *IET Biom*, vol. 7, no. 5, pp. 423–430, Sep. 2018, doi: 10.1049/iet-bmt.2017.0171.

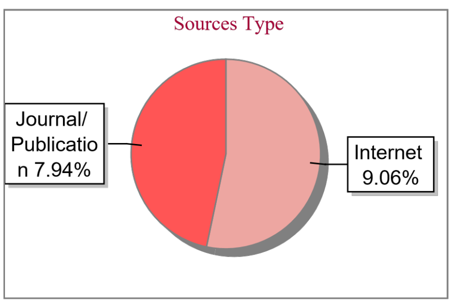
The Report is Generated by DrillBit Plagiarism Detection Software

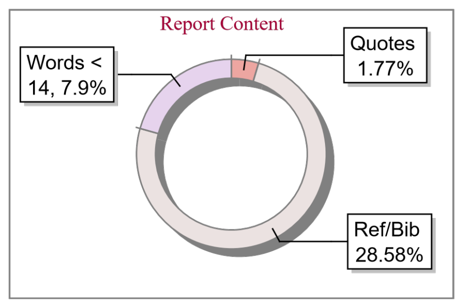
Submission Information

Author Name	SATYAM RAJ
Title	Face Detection and Age Gender Recognition
Paper/Submission ID	1779502
Submitted by	asha.negi@dituniversity.edu.in
Submission Date	2024-05-10 12:51:56
Total Pages	30
Document type	Project Work

Result Information

Similarity **17 %**





Exclude Information

Quotes	Excluded
References/Bibliography	Excluded
Sources: Less than 14 Words %	Not Excluded
Excluded Source	**0 %**
Excluded Phrases	Not Excluded

Database Selection

Language	English
Student Papers	Yes
Journals & publishers	Yes
Internet or Web	Yes
Institution Repository	No

A Unique QR Code use to View/Download/Share Pdf File

DrillBit

DrillBit Similarity Report

A-Satisfactory (0-10%)

B-Upgrade (11-40%)

C-Poor (41-60%)

D-Unacceptable (61-100%)

SIMILARITY %	MATCHED SOURCES	GRADE
17	46	B

LOCATION	MATCHED DOMAIN	%	SOURCE TYPE
1	www.ncbi.nlm.nih.gov	2	Internet Data
2	arxiv.org	1	Publication
3	fastercapital.com	1	Internet Data
4	dl.uswr.ac.ir	1	Publication
5	Thesis submitted to shodhganga - shodhganga.inflibnet.ac.in	1	Publication
6	www.weka.io	1	Internet Data
7	www.frontiersin.org	1	Internet Data

8	ejnpn.springeropen.com	1	Internet Data
9	Automatic Detection of Negative Symptoms in Schizophrenia via Acoustically Measu by He-2020	<1	Publication
10	moam.info	<1	Internet Data
12	www.geeksforgeeks.org	<1	Internet Data
13	www.ncbi.nlm.nih.gov	<1	Internet Data
14	journalofbigdata.springeropen.com	<1	Internet Data

16	www.dx.doi.org	<1	Publication
17	Thesis Submitted to Shodhganga Repository	<1	Publication
18	www.linkedin.com	<1	Internet Data
19	citeseerx.ist.psu.edu	<1	Internet Data
20	www.intechopen.com	<1	Publication
21	www.mdpi.com	<1	Internet Data
22	Dynamic classification for video stream using support vector machine, by Mariette Awad Yuic- 2008	<1	Publication
23	kth.diva-portal.org	<1	Publication
24	mdpi.com	<1	Internet Data
25	springeropen.com	<1	Internet Data
26	transsyst.ru	<1	Internet Data

27	www.ijcrt.org	<1	Publication
28	www.intechopen.com	<1	Publication
29	www.ncbi.nlm.nih.gov	<1	Internet Data
30	Discovery and validation of gene classifiers for endocrine-disrupting, by Wang, Rong-Lin Ben-	<1	Publication
31	publikationen.bibliothek.kit.edu	<1	Publication
32	www.frontiersin.org	<1	Publication
15	Computer image analysis for intramuscular fat segmentation in dry-cured ham slic by Muoz-2019	<1	Publication

33	IEEE 2018 Joint 10th International Conference on Soft Computing and	<1	Publication
37	springeropen.com	<1	Internet Data
40	adoc.pub	<1	Internet Data
41	arxiv.org	<1	Publication
42	eprints.covenantuniversity.edu.ng	<1	Internet Data
43	IEEE 2018 Condition Monitoring and Diagnosis (CMD)- Perth, WA (201, by Zhang, Qinqin Lin,- 2018	<1	Publication
44	nature.com	<1	Internet Data
45	Risk Upper Bounds for General Ensemble Methods with an application to Multiclass by Laviolette-2017	<1	Publication
46	Spectral Imagery Tensor Decomposition for Semantic Segmentation of Remote Sensin by Lpez-2020	<1	Publication
47	Thesis submitted to shodhganga - shodhganga.inflibnet.ac.in	<1	Publication
48	unctad.org	<1	Publication
49	VirPreNet a weighted ensemble convolutional neural network for the virulence pr by Yin-2020	<1	Publication
50	www.dx.doi.org	<1	Publication
51	www.dx.doi.org	<1	Publication