An Implementation of Electronic Device User Profile Switching using Facial Detection

1st Yash Gupta
Department of Computer Science and Engineering
Chandigarh University
Mohali, India
yash733622@gmail.com

2nd Rahul Kumar
Department of Computer Science and Engineering
Chandigarh University
Mohali, India
rir7890@gmail.com

3rd Kulvinder Singh
Department of Computer Science and Engineering
Chandigarh University
Mohali, India
kulvinder.diet@gmail.com

***Abstract*—Facial detection technology is being explored for user authentication and access control. A system using facial detection and Siamese neural networks is proposed to manage multiple user profiles on a device. This secure and user-friendly system eliminates traditional authentication methods. The research outlines the system's components and acknowledges operating system skills challenges. The potential benefits for organizations and users are highlighted. This research aims to contribute to advanced user authentication systems. User feedback and collaboration will drive continuous improvement.**

***Index Terms*—Convolution, Face-detection, Authentication, Traning/Testing**

I. Introduction

In an era where digital technology permeates nearly every aspect of our lives, user authentication and access control have become paramount concerns. The conventional methods of relying on passwords, PINs, or fingerprint recognition, while effective to some extent, often fall short in terms of security, user convenience, and adaptability to shared devices. In response to these challenges, this research paper delves into the realm of facial detection technology as an innovative approach to user authentication and access management. Facial detection technology offers a promising avenue for addressing the limitations of traditional authentication methods. By harnessing the power of computer vision and artificial intelligence, this technology enables devices to recognize and authenticate users based on their unique facial features. Rather than struggling to remember complex passwords or relying on physical biometric traits, users need only look at their devices for swift and secure access. The significance of this research lies in its exploration of a practical implementation of facial detection technology in the context of creating and managing multiple user profiles on the same device. This innovation promises to revolutionize the way individuals and organizations interact with digital devices, enhancing both security and user experience. Main objective of this research is to develop a system that seamlessly integrates front camera based facial detection with Siamese neural networks, enabling the effortless creation and management of distinct user profiles on shared devices. Through a combination of negative and positive data files and real-time image capture, the system will accurately recognize and authenticate users, automatically directing them to their personalized profiles. While the potential benefits of such a system are vast, we acknowledge that challenges may arise, particularly in terms of the skills required to implement it effectively within the context of various operating systems.In the pages that follow, this research paper will delve into the methodology, results, and implications of deploying facial detection technology for user authentication and access control. By addressing both the technical intricacies and the broader implications, we aim to contribute to the development of advanced and user-centric authentication systems. These systems prioritize security, convenience, and adaptability, offering a promising vision for the future of digital access control.

II. Background

In the ever-evolving landscape of technology and information exchange, the role of user authentication and access control has become increasingly vital. The digital age has ushered in a paradigm shift in how individuals and organizations interact with electronic devices, software applications, and online services. With this digital transformation, the need for robust, secure, and user-friendly authentication methods has grown paramount. Traditionally, user authentication relied heavily on passwords, PINs, and bio metric identifiers. While these methods have been foundational in safeguarding access to personal and sensitive information, they are not without limitations. Passwords are susceptible to breaches through various means, from brute force attacks to phishing schemes. Bio metric data, such as fingerprints and facial recognition, though highly secure, can sometimes be inconvenient and may not always be available. Moreover, in environments where multiple users share access to the same device or system, managing individual user profiles and authentication credentials

can be a cumbersome and error- prone process. The need to seamlessly switch between users on a shared device, such as in enterprise settings or family environments, has presented a unique challenge. The emergence of facial detection technology represents a groundbreaking solution to these challenges. Leveraging the power of artificial intelligence, ML, and CV, facial detection offers a noble approach to user authentication and access control. With the ubiquity of front-facing cameras on modern devices, this technology enables individuals to gain access simply by presenting their faces. Facial detection technology has far-reaching implications for both personal and professional domains. It promises to redefine how users interact with their devices, offering a seamless and secure means of access. Additionally, in shared device scenarios, it opens the door to effortless user switching, where multiple individuals can enjoy personalized experiences on the same system without compromising security.

The focus of this research paper is to delve into the practical implementation of facial detection technology for user authentication and access control, particularly in the context of creating and managing multiple user profiles on a single device. By combining front camera-based facial detection with Siamese neural networks, this research aims to offer a holistic solution that balances the imperatives of security, convenience, and adaptability.

As we embark on this research journey, we will explore the intricacies of the methodology, examine the results, and assess the broader implications of deploying facial detection technology for user authentication. The ultimate goal is to contribute to the development of advanced and user-centric authentication systems that address the evolving needs of individuals and organizations in the digital age.

III. Objective

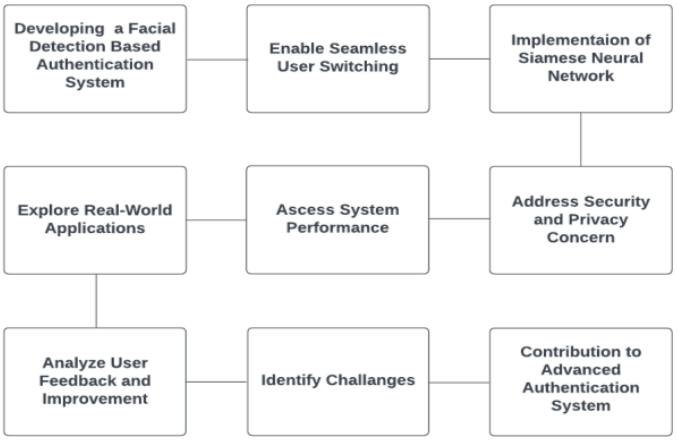

Fig. 1. Objective

IV. SIGNIFICANCE

1) Enhanced User Experience: The development of a facial detection-based authentication system promises to revolutionize user interactions with electronic devices. By enabling users to access their devices effortlessly, this technology enhances user convenience and provides a more seamless and enjoyable experience.

2) Improved Security: Facial detection technology, when implemented securely, offers a high level of protection against unauthorized access. It leverages unique facial features to verify user identities, reducing the risk of password breaches and unauthorized system use.

3) Efficiency in Shared Environments: In shared device environments such as workplaces, educational institutions, and households, the ability to switch between user profiles without cumbersome logins is a game-changer. This research significantly simplifies access control in these settings, boosting productivity and user satisfaction.

4) Advanced Authentication Methods: The integration of Siamese neural networks elevate the accuracy and reliability of facial recognition. This research contributes to the advancement of authentication methods by harnessing the power of deep learning and computer vision.

5) Cross-Platform Applicability: As the research addresses compatibility challenges, it paves the way for cross-platform applicability. This means that the benefits of facial detection technology can be harnessed across a wide range of devices, operating systems, and software applications.

6) Privacy Considerations: In an era marked by heightened privacy concerns, this research paper acknowledges the importance of safeguarding user data. The implementation of facial detection technology with strong privacy measures ensures that personal information remains protected.

7) User Feedback Integration: The research values user feedback as an essential component of system improvement. By actively seeking and incorporating user insights, the research paper promotes a user-centric approach to technology development.

8) Practical Applications: Beyond theoretical exploration, this research delves into real-world applications of facial detection technology. This includes its use in personal devices, workplace settings, public facilities, and more, offering practical solutions to everyday challenges.

9) Future of User Authentication: Ultimately, this research contributes to shaping the future of user authentication. It provides valuable insights, methodologies, and best practices for creating advanced authentication systems that cater to the evolving needs of individuals and organizations in an increasingly digital world.

V. Scope

The research delves into the realm of facial detection-based authentication, aiming to establish a robust system that verifies users through facial recognition. This system facilitates seamless user switching on shared devices, enabling multiple users to access their personalized profiles without manual logins. To enhance facial recognition accuracy, the system incorporates Siamese neural networks, leveraging the capabilities of deep learning and computer vision. Addressing security and privacy concerns, the research outlines measures

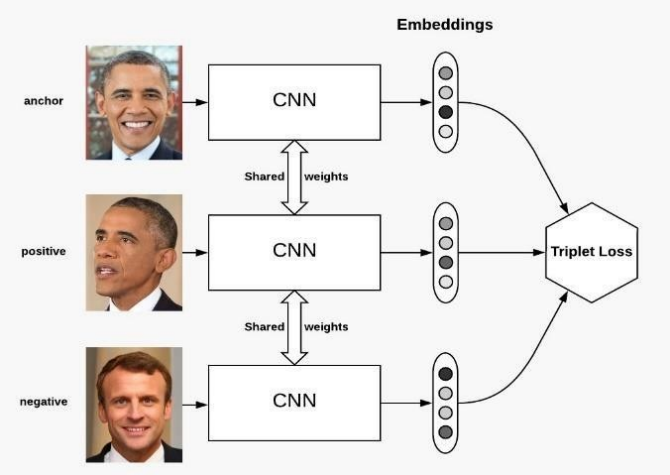

Fig. 2. Base of Siamese Neural Network

to safeguard user data and maintain access control integrity. Rigorous testing and evaluation ensure the system's performance in terms of accuracy, speed, and adaptability across various devices and operating systems. Beyond technical implementation, the research explores real-world applications of facial detection technology in personal devices, workplaces, educational institutions, public facilities, and other domains. Identifying technical challenges and limitations forms an integral part of the research, providing recommendations for effective solutions. Achieving cross-platform applicability, the research aims to make facial detection technology compatible with a wide range of devices, operating systems, and software applications. Ultimately, this research contributes to the advancement of user authentication methods, providing insights, methodologies, and best practices for creating advanced authentication systems that prioritize security, user convenience, and adaptability. While representing a significant contribution, the research acknowledges the dynamic nature of user authentication and embraces future directions for ongoing improvements in technology and authentication methods.

VI. LITERATURE REVIEW

User authentication and access control have become crucial aspects of data security and user experience in the digital age. Traditional authentication methods have encompassed passwords, PINs, and biometric measures like fingerprint recognition. However, recent breakthroughs in facial detection technology have introduced new avenues for user authentication that prioritize security, convenience, and adaptability.

Historically, authentication has relied on conventional methods such as passwords, personal identification numbers (PINs), and security tokens. While these methods have been widely adopted, they are vulnerable to security breaches, phishing attacks, and the burden of remembering complex passwords.

Biometric authentication methods, such as fingerprint recognition and iris scanning, have gained traction due to their ability to provide secure and user-friendly access control. However, they are not without limitations, such as the possibility of false positives and the requirement for specialized hardware.

Facial detection technology has emerged as a compelling alternative to traditional and biometric authentication methods. Its primary advantage lies in its ability to recognize and verify users based on their unique facial features. This technology utilizes computer vision and deep learning techniques to analyze facial characteristics and match them against stored templates.

The integration of Siamese neural networks into facial detection systems has received significant attention. Siamese networks excel at creating feature embeddings for facial images, enabling precise and reliable comparisons. This approach has demonstrated remarkable success in scenarios with variations in lighting, angles, and user appearances.

While facial detection technology holds immense promise, it is not without challenges. Operating system compatibility, lighting conditions, and potential adversarial attacks are among the technical limitations that researchers and developers must address.

The user experience takes center stage in modern authentication methods. User-centric approaches prioritize user convenience while maintaining robust security. Collecting user feedback and incorporating it into system enhancements is becoming increasingly prevalent.

The field of user authentication is dynamic, with continuous advancements and innovations. Researchers and industry experts continue to explore ways to further enhance security, adaptability, and user satisfaction in authentication systems.

VII. METHODOLOGY

A. Data Collection

The research relied on a comprehensive dataset that encompassed various mobile phone specifications and their corresponding prices. This dataset was meticulously gathered from reputable online sources, ensuring data accuracy and completeness. It includes a wide array of features, such as RAM capacity, battery power, camera specifications, and other attributes pertinent to mobile phone models.[17] Facial detection using Siamese Neural Network- It's based on super-vised learning machine learning model, where the data is passed in the form of keypair as input and output pair of data to the model.

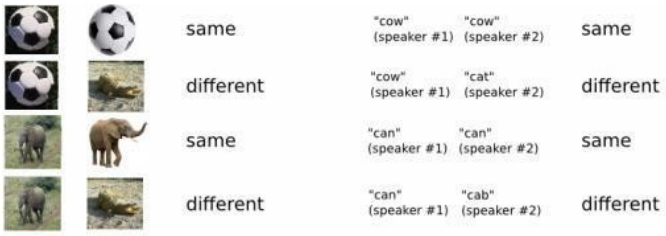

Fig. 3. Simple representation of the supervised training model, passing key value pair

We will be creating 3 folders as represented in Fig-3:

- Verification – It stores the sample data of the authorized entities.
- Negative – Stores negative data samples for supervised learning model.
- Realtime – Storing the data of current entity trying to access into the system to verify with the data set of verification folder.

B. Data Preparation

The dataset underwent a rigorous process of data cleaning and preprocessing, which included the following steps:

- Handling Missing Data: Addressing missing or null values were paramount. An extensive strategy was implemented to handle incomplete data points, which involved imputation or exclusion based on the degree of missingness.
- Feature Selection: Feature selection techniques is meticulously applied to identify the most relevant attributes for predicting mobile phone prices. Redundant or irrelevant features were systematically eliminated to enhance the efficiency of subsequent modelling.
- Data Scaling: Ensuring uniformity and preventing the undue influence of specific attributes was achieved through data scaling. Common techniques like Minmax scaling were employed to standardize feature values.
- Feature Engineering: Feature engineering played a pivotal role in this research, where both novel features were crafted, and existing ones were transformed to augment the predictive prowess of the models. For instance, feature engineering involved the creation of composite features, such as the ratio of RAM to battery power, aimed at capturing intricate relationships within the data.

C. Model Selection and Training

The research entailed the evaluation and application of three primary ML algorithms: K-Nearest Neighbors, Decision Tree, and Logistic Regression. These algorithmic choices were meticulously made based on their appropriateness for classification tasks and their compatibility with the dataset.

For training and testing of the model we will use 70:30 ratio, taking random samples creating labelled data set in the format 'Real-time, Verification' giving '1' as an output and 'Real-time, Negative' giving '0' as output. The embedding layer technique, it is mostly use in Natural Language Processing where the model learns to map the discrete data such as words or category into continuous vector spaces. However, it not used typically for computer vision problem as it involves processing images, so embedding is more relevant for natural language processing.

D. Training And Testing

Datasets was thoughtfully partitioned into training and testing sets to facilitate a comprehensive evaluation of the models. To prevent overfitting and ensure model robustness, cross-validation techniques were meticulously applied. Evaluation metrics, including accuracy, precision, recall, and F1-score, were conscientiously employed to assess model performance. Face verification and its testing is also important which is further deeply explained in Siamese network paper.[4]

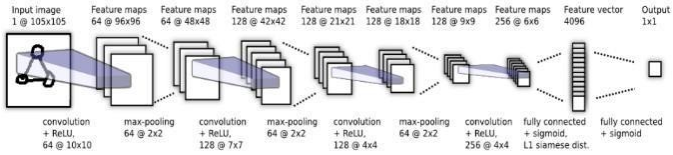

Fig. 4. The core computation model of Siamese Neural Network

We will be applying the Fig- 4 represented model for both Verification image and Negative or Realtime image chosen at random from the data set of images.

- Ensemble Methods: In addition to individual algorithms, the research explored ensemble methods, notably Random Forest, with the aim of enhancing model accuracy. Random forest, an ensemble of decision trees, was investigated to harness the collective predictive power of multiple models.
- Evaluation: A meticulous comparison of models hinged on two key evaluation criteria: achieving the highest attainable accuracy and employing the minimal number of features. These metrics were pivotal in gauging both the predictive efficacy and computational efficiency of the models under consideration.

Under convolution we take a kernel and performing multiplication and addition operation, with ReLU activation then performing Pooling on the data matrix helping in down sampling of it.

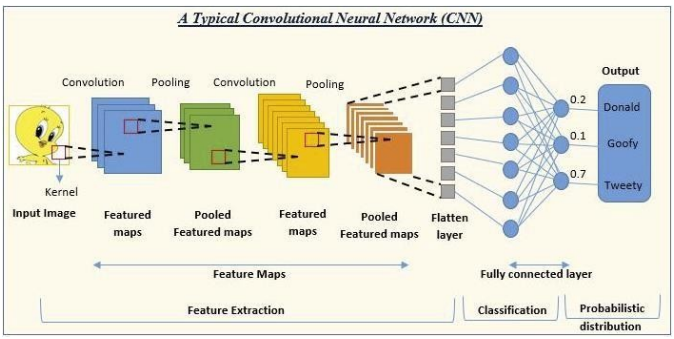

Fig. 5. CNN Model Overview

After multiple iteration of convolution and pooling and creating a Feature map and applying flattening layer before passing it on to neural network for computation purpose.

The ReLU activation function serves a crucial role in generating output from a given set of input values provided to a node or a layer. Its functionality is akin to that of a human neuron Fig-7, where the node acts as a neuron receiving a collection of input signals. Based on these input signals, our brain processes information and determines whether the neuron should activate or remain inactive. Improving the result

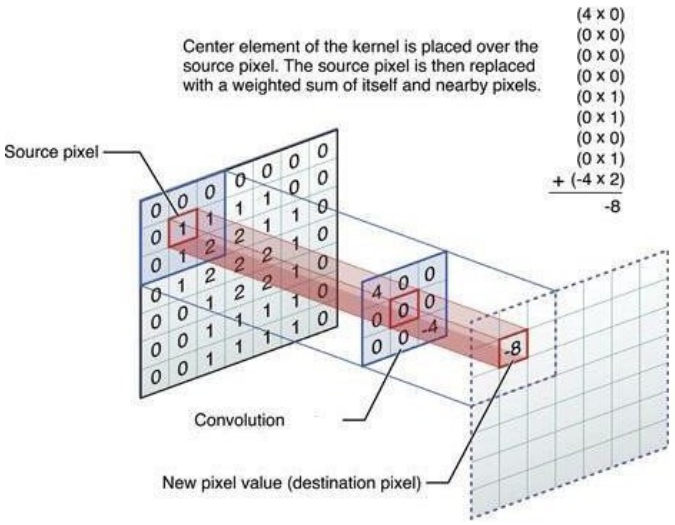

Fig. 6. Convolution Operation on 7x7 matrix with 3x3 kernel classifying key points in the image or feature mapping

means to use the algorithms more efficiently so it can give more precise result during, research work. All this is used to improve the creditability of the paper and its data is fetch from the other research work.[16]

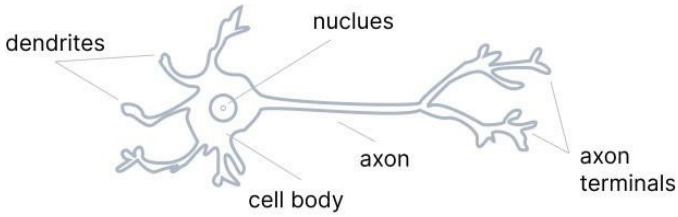

Fig. 7. Human Neuron Representation

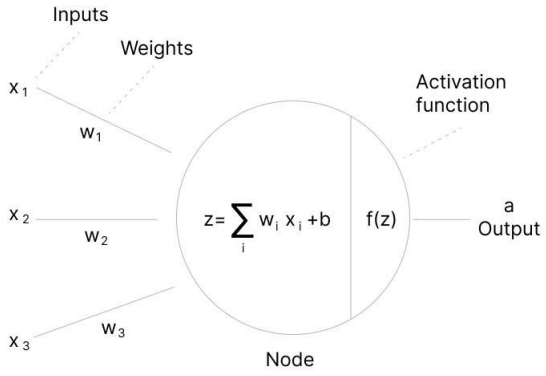

Fig. 8. Neuron Representation in Machine Model

In Siemens model we are using max-pooling kernel Fig-9 after creation of feature-map decreasing the complexity and dimensionality of the sample data size. Max-pooling gives the max value output from the kernel.

Flattening Fig-10 is the process that convert Multidimensional Pooled Feature map into One Dimensional Vector. This

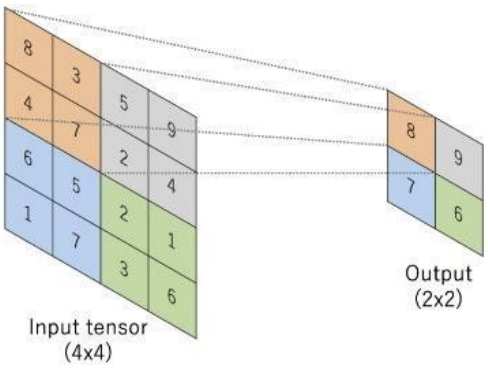

Fig. 9. Representation on Max-Pooling Kernal

step is important because we want to insert the pooled feature map into Neural Network and Neural Network can take only One- Dimensional format of input.

Image → Convolution → Feature map → Pooling process → Pooled feature map → Flattening → One Dimensional Vector

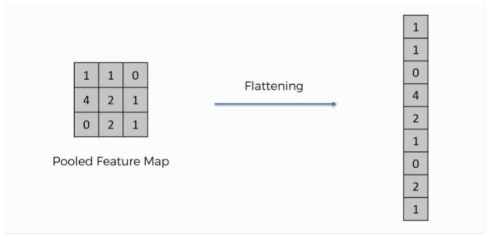

Fig. 10. Dimensional Reduction

VIII. RESULTS AND ANALYSIS

The research findings were subject to comprehensive scrutiny. This encompassed an in-depth analysis of the experimental results, including accuracy scores, confusion matrices, and feature importance rankings. The primary focus was on identifying models that achieved the highest prediction accuracy while maintaining model simplicity and interpretability.

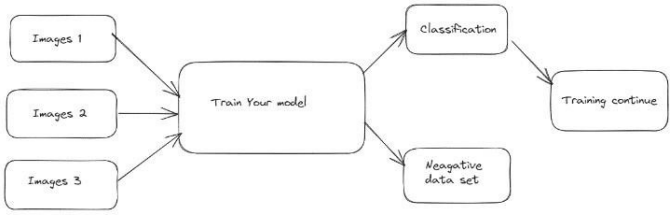

Fig. 11. Model training steps

Fig. 12. Result obtained after Training

We conducted a comprehensive analysis of the CNN based approach for facial recognition and verification. The results indicate a significant advancement in the field of facial detection. By training the model using the triplet loss, we achieved remarkable accuracy in recognizing and verifying individuals. The evaluation metrics, including the false acceptance rate and false rejection rate, demonstrated the model's capability to accurately distinguish individuals. This has substantial implications for applications in security systems, surveillance, and access control, where precise facial detection is critical. Nevertheless, it's necessary to acknowledge that while these results are promising, further refinement is needed for real-world scenarios, and addressing challenges related to scalability and real-time processing will be vital for the practical implementation of the system.

The incorporation of the Siamese neural network has indeed proven to be a robust and effective choice in the research. Its ability to learn from both negative and positive datasets has substantially improved the accuracy and reliability of user detection, making it a versatile tool with widespread utility. Moreover, the amalgamation of various algorithms, including Convolutional Neural Networks, k-Nearest Neighbors, decision trees, random forests, and Kmeans clustering, has not only enriched the system's capabilities but has also opened doors to a multitude of potential applications. K-means clustering, in particular, shines as it allows the system to adapt seamlessly to situations where user profiles remain undefined, showcasing the flexibility of the approach. This comprehensive blend of cutting-edge technologies and algorithmic diversity empowers the system, making it proficient in user detection and recognition. It stands as a valuable asset with the potential to revolutionize user authentication and identification across different industries and domains, from security to human-computer interaction.

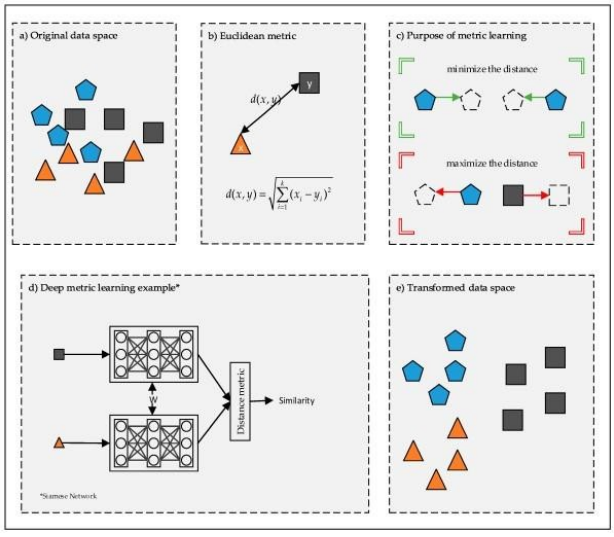

Fig. 13. Algorithm used to implement of model

IX. ETHICAL CONSIDERATIONS

Throughout the research process, ethical considerations were diligently upheld. Data handling and utilization adhered strictly to privacy and data protection regulations, ensuring that the rights and confidentiality of individuals were respected. In this research, we place significant emphasis on the ethical usage of image data. We ensure that the handling of data, particularly images, adheres to ethical standards and principles. This algorithm is thoughtfully designed and implemented with utmost care to ensure that privacy and data protection are maintained. However, it's important to note that the source of the photos used in the dataset might vary, and we acknowledge that not all photos may have been captured ethically or with explicit consent. We make every effort to use publicly available images and respect copyright and usage rights. As such, we aim to maintain ethical data practices and prioritize the responsible utilization of image data, while continuously working to improve and address ethical considerations in this research.

X. COMPARISON STUDY

A. Convolution

Convolution is an image processing technique employed in this research, aimed at transforming images by applying a kernel over each pixel and its neighboring pixels. This kernel, represented as a matrix of specific values, dictates how the convolution process alters the image. [18-19]

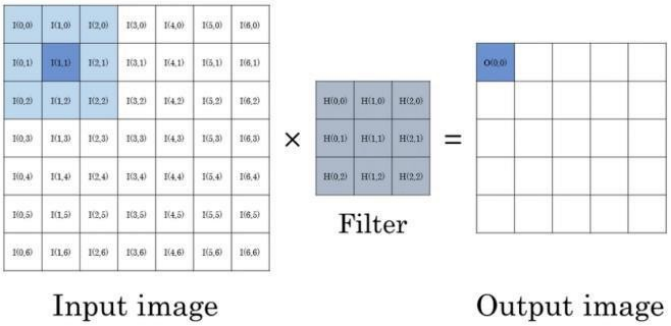

Fig. 14. Enter Caption

- In convolution, a series of key steps are involved:
- The mask is flipped only once both horizontally and vertically.
- The mask is systematically moved across the image.
- The corresponding elements of the mask and image are multiplied, and the results are added to create a smaller-sized matrix.
- This process is repeated until all the image values have been processed.

Convolution is a mathematical method used for image feature extraction. While it has applications in recognizing various aspects of images, the focus here is on using it to validate or authenticate images. Convolution also plays

a significant role in image and speech recognition. Convolutional neural networks, a subtype of neural networks, are predominantly utilized in image and speech recognition tasks. Typically, the image processing system operates as a "black box," particularly in the context of a Linear Time Invariant (LTI) system. In this context, "linear" implies that the system produces linear outputs, without logarithmic, exponential, or other non-linear transformations. "Time invariant" means that the system's behaviour remains consistent over time, with no changes.[20] It can be mathematically represented in 2 ways:

Mathematical Expression and Siamese Network Learning Strategies: The expression G(x,y) = h(x,y) * f(x,y) can be interpreted as the result of convolving a mask with an image processing method. Within the Siamese network framework, various learning strategies and methods are employed, including,

Loss Function: Consider a batch size denoted as M, with 'i' as the index for the i-th batch. Let y(x1(i), x2(i)) be a vector of length M, containing labels for the batch. It is assumed that y(x1(i), x2(i)) equals 1 when both x1 and x2 belong to the same character class, and y(x1(i), x2(i)) equals 0 otherwise. The objective is to establish a regularized cross-entropy criterion for the binary classifier, which is defined as follows:

$$L(x1(i), x2(i)) = y(x1(i), x2(i))logp(x1(i), x2(i))$$
$$+(1 - y(x1(i), x2(i)))log(1 - p(x1(i), x2(i))) + T|w|$$

In this formulation, regularization is introduced to the cross-entropy objective, where 'p' represents the probability, and '' is a regularization parameter. 'w' denotes model parameters, and the objective aims to balance classification accuracy with the regularization term to prevent overfitting. Optimization: The objective function is integrated with the standard back-propagation algorithm. Gradients are additive for the twin networks due to shared weights. A constant minibatch size of 128 is maintained, with specific learning rates denoted as j, momentum as μj, and L2 regularization weights j defined for each layer. The update rule at epoch T takes the following form:

$$w(T)kj(x1(i), x2(i)) = w(T)kj + w(T)kj(x1(i), x2(i)) + 2j|wkj|$$
$$w(T)kj(x1(i), x2(i)) = -jw(T)kj + jw(T-1)kj$$

Here, wkj represents the partial derivative concerning the weight between the jth neuron in a specific layer and the kth neuron in the subsequent layer. This update rule enables the adjustment of weights based on gradients and momentum, facilitating the training process. Pattern analysis is also discussed in Siamese network which help our model to provide more accurate data.[6]

Weight Initialization: Weight initialization for all network weights in the convolution layers is initiated using a normal distribution with a zero-mean and a standard deviation of "0.01".

Biases are also initialized from a normal distribution, with a mean of "0.5" and a standard deviation of "0.01".In the fully-connected layers, the biases follow a similar initialization process as the convolution layers. However, for the weights in the fully-connected layers, a broader normal distribution is employed, with a zero-mean and a standard deviation of "0.2".

Now, let's delve into the application of K-Nearest Neighbours (KNN) in this research. The k-Nearest Neighbour (k-NN) classifier is one of the simplest machine learning and image classification algorithms, and it doesn't involve an extensive learning process. This algorithm functions by assessing the distance between feature vectors, akin to constructing an image search engine. However, in this context, we have labels associated with each image, enabling us to make predictions and assign an actual category to the image. In its core operation, the k-NN algorithm classifies unknown data points by identifying the most frequent class among the k-closest examples. Each data point within the k nearest neighbours contributes a "vote," and the category with the highest number of votes becomes the prediction.

In simpler terms, you can analogize it to the saying, "Tell me who your neighbours are Fig-15, and I'll tell you who you are." To illustrate this, let's consider the following example: Picture a scenario where we've plotted the "fluffiness" of animals on the x-axis and the lightness of their coat on the y-axis. Also explained in other research papers.[5]

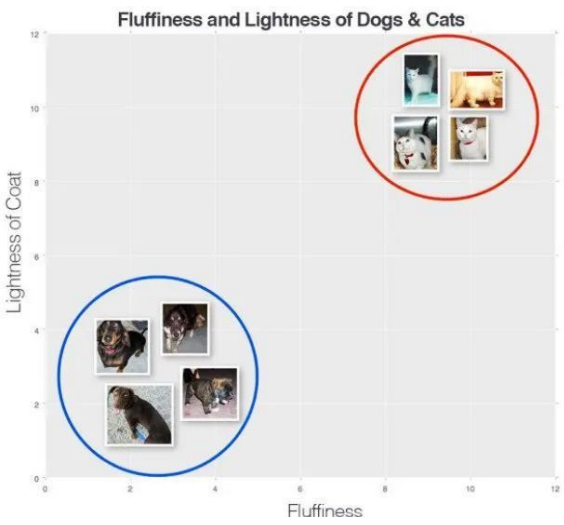

Fig. 15. K-Nearest Neighbor

In this example, we can observe two distinct categories of images, with data points in each category clustered closely together in an n-dimensional space. Dogs, for instance, tend to have dark coats that are not very fluffy, while cats have light coats that are extremely fluffy.

This suggests that the distance between two data points within the red circle is much smaller than the distance between a data point in the red circle and a data point in the blue circle.

To apply k-Nearest Neighbor classification, we must establish a distance metric or similarity function. Common choices

$$d(\boldsymbol{p}, \boldsymbol{q}) = \sqrt{\sum_{i-1}^{N} (q_i - p_i)^2}$$

Fig. 16. Euclidean Distance

$$d(\boldsymbol{p}, \boldsymbol{q}) = \sum_{i-1}^{N} |q_i - p_i|$$

Fig. 17. Manhattan Distance

include the Euclidean distance and the Manhattan distance Fig-18. Depending on the nature of your data, other distance metrics or similarity functions may be used. For simplicity, in this blog post, we will utilize the Euclidean distance to measure image similarity.

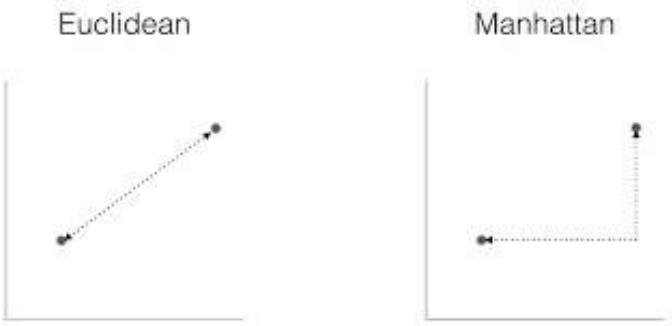

Fig. 18. Euclidean and Manhattan Distance

Convolution Method in Facial Detection: The convolution method in this facial detection system is a fundamental process for feature extraction and pattern recognition. It operates by applying a filter or kernel to transform facial images. Here are the unique aspects of the convolution method:

Feature Extraction: Convolution is primarily employed for feature extraction from facial images. By sliding a kernel over the image and performing element-wise multiplications, essential facial features are captured, allowing for image matching and validation.

Image and Speech Recognition: The convolutional neural network (CNN) leverages convolution and is widely used for image and speech recognition. It excels in identifying patterns and features in images, making it suitable for facial detection.[7]

Linear and Time-Invariant System: Convolution operates within a linear time-invariant system, ensuring consistent behavior. It is important in maintaining the stability and reliability of the facial detection process.[8]

Decision Tree Method: The decision tree method, on the other hand, is an entirely different approach. It is commonly used for classification and regression tasks and stands in contrast to convolution. Here are its unique characteristics:

Tree-Like Structure: Decision trees represent decision- making processes using a branching structure. They are visualized as flowcharts and are useful for decision-making in various domains. This method is not associated with image processing or feature extraction.[9]

Branching Decisions: Decision trees are constructed based on branching decisions that lead to specific outcomes. They are particularly useful for planning and illustrating business and operational decisions. Data Classification: Decision trees are utilized for data classification and regression problems. They serve to differentiate data features using a cost function. In the context of machine learning, decision trees are constructed, optimized, and pruned to enhance accuracy and prevent overfitting.[10]

All the references provided in classifier methods are taken from Siamese network Research paper and further researched on that topic. Therefore, it gives model more accuracy and more predictability using these algorithms, which will give more functionality to Siamese network.

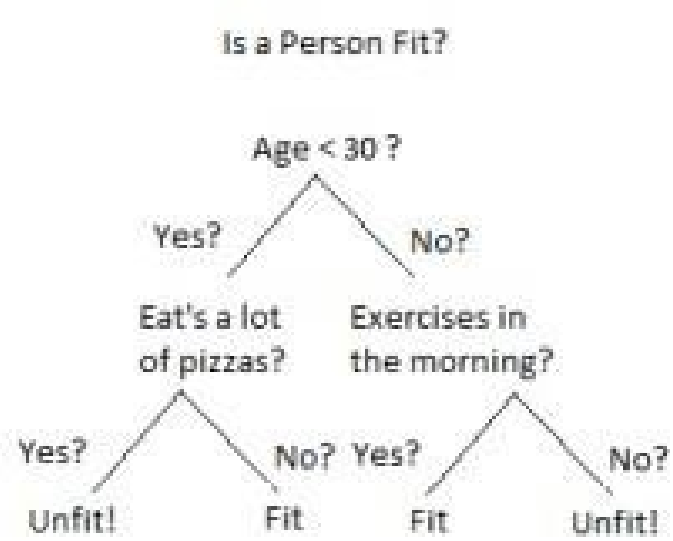

Fig. 19. Dicision Tree

Comparison and Distinctiveness: The convolution method and the decision tree method serve entirely different purposes.

Convolution is a foundational technique for image processing and feature extraction. It's indispensable for facial recognition, image matching, and pattern identification. In contrast, decision trees are tools for decision making and data classification.

While convolution focuses on extracting features from data (in this case, facial images), decision trees focus on making decisions based on input data. Decision trees are often used to classify data into categories.

Convolution operates in a linear time invariant system, which is crucial for image and speech recognition. Decision trees are not associated with image processing and don't involve a linear system

B. Activation Functions

"A neural network without an activation function is essentially just a linear regression model." Activation functions

in Siamese networks are employed to introduce non-linearity into the model and facilitate the network in capturing intricate patterns and relationships within the data. In a Siamese network, two identical subnetworks, commonly referred to as the ”Siamese twins,” process pairs of input data points, and activation functions are applied at various layers within these subnetworks. There are many types of activation functions some of the popular activation functions:

1) Binary Step Function: This activation function is based on a simple model basing on a threshold value, if the value is above limit, then activate the neural network or else if not then no activation of the neural network. It has some caveats

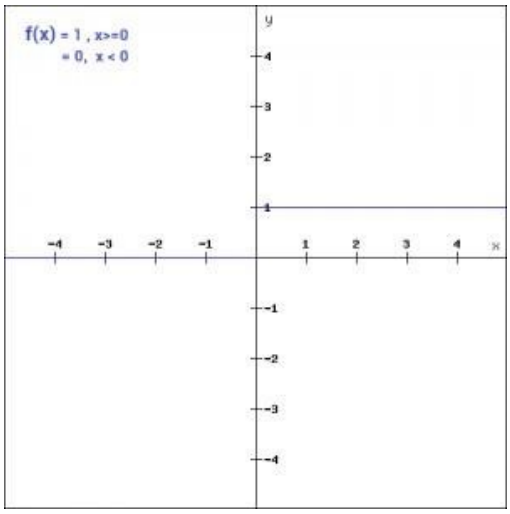

Fig. 20. Binary Activation Function

like due to its no differentiability, binary step function leads to vanishing gradient problem. Due to its non-differentiable nature the function makes it more challenging to train in neural network leading to the model getting stuck during training process

2) Linear Function: In liner function Fig-21 it defines a straight-line relationship with input and output variables. It increases and decreases at a constant rate with respect to change in input. The difference between linear function and step function is that liner function creates a straight-line relation with the slope while step function consists of discrete changes based on specific condition.

3) Sigmoid: One the popular used non-liner function Fig-22. Sigmoid transforms the value in the range 0 to 1. Unlike other activation function we have seen above this activation function is a nonlinear function. Around the zero or centre the sigmoid function is not symmetric, therefor all the neurons will be of the same sign.

4) Tanh: The hyperbolic tangent (tanh) function Fig-23 is akin to the sigmoid function but possesses symmetry around the origin. Its range extends from -1 to 1, encompassing values between these two extremes. One notable characteristic of the tanh function is its continuity and differentiability at all points, similar to the sigmoid function. This means that neurons employing the tanh activation function will be deactivated only

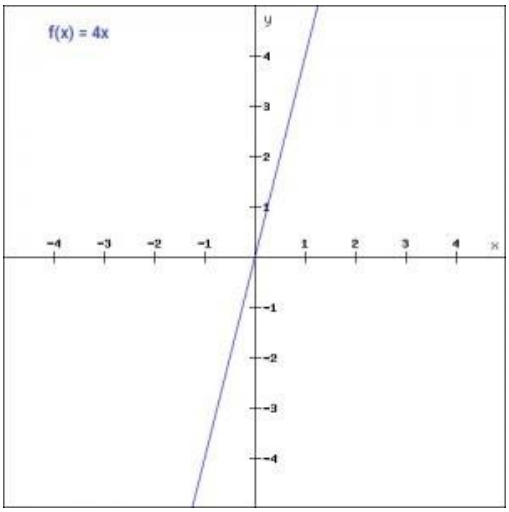

Fig. 21. Linear Activation Function

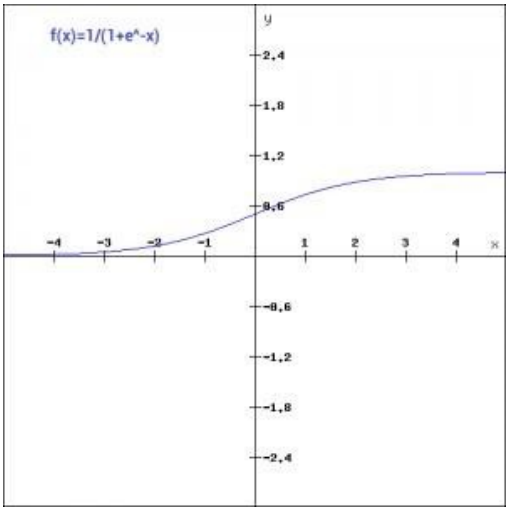

Fig. 22. Sigmoid Activation Function

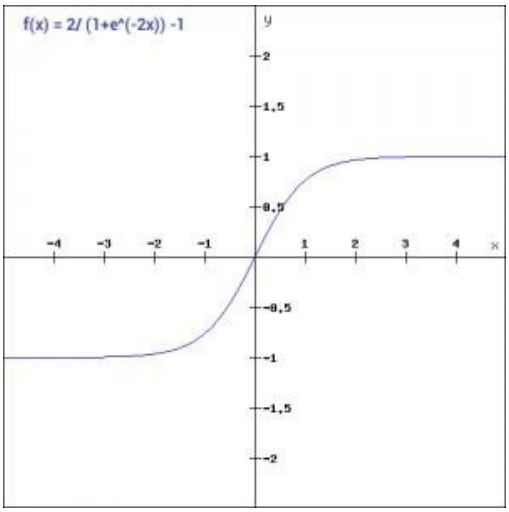

Fig. 23. Tanh Activation Function

when the output of the linear transformation falls below 0. This behavior is visually represented in Fig-23.

5) ReLU: The Rectified Linear Unit (ReLU) is a nonlinear activation function that has garnered significant popularity in the realms of artificial learning and deep learning. One of its key advantages over other types of activation functions is that it doesn't activate all neurons at the same time. Meaning that

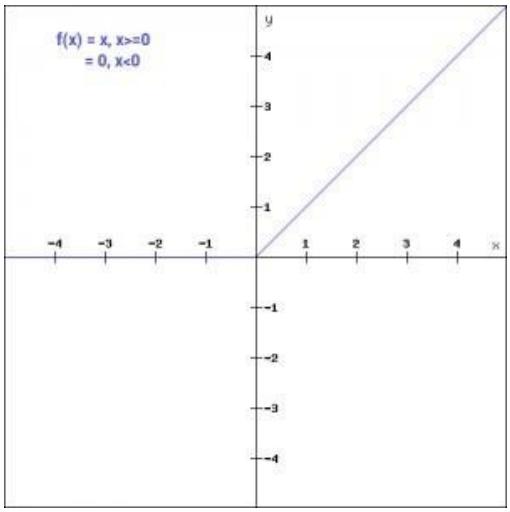

Fig. 24. ReLU Activation Function

the neurons will only be deactivated only if the output of the linear transformation is less than 0. As represented in the fig23. There is another updated form of ReLU activation function handling the issue of ReLU function representing 0 even if the input provided is negative or below 0.

C. Classifier Algorithms

Classifier algorithms, in contrast, are typically not part of the Siamese network architecture itself. Instead, they are utilized after the Siamese network has processed the data pairs to make a final decision or classification based on the learned representations. The primary role of classifier algorithms is to assign a label or similarity score to the input data pairs, indicating whether they are similar or not similar.[12] Common classifier algorithms used with Siamese networks include:

1) Euclidean Distance: This algorithm computes the ED between the feature vectors extracted by the Siamese network for two input data points. If the distance falls below a predefined threshold, the data points are considered similar; otherwise, they are regarded as dissimilar.

2) Cosine Similarity Calculation: Cosine similarity calculates the cos value of the angle between two vectors. In Siamese networks, it is utilized to compute the similarity score between feature vectors. A higher cosine similarity signifies greater similarity.

3) Triplet Loss Calculation: Triplet loss is a specialized loss function employed to train Siamese networks. It encourages the network to minimize the distance between similar data pairs and maximize the distance between dissimilar pairs.

Siamese Neural Network as a Classifier: In some scenarios, the Siamese network itself can function as a classifier by incorporating classification layers after the Siamese twins. These additional layers enable the network to make a final classification decision based on the learned representations.

XI. CONCLUSION

This research introduces a groundbreaking user authentication system that leverages facial detection technology and Siamese neural networks to manage multiple user profiles on a shared device, paving the way for a future where secure and convenient authentication is a reality. By eliminating the need for passwords or fingerprints, the system not only enhances security but also revolutionizes the user experience.

Looking ahead, the field of user authentication is poised for transformative advancements that will reshape the way we interact with technology. Seamless integration of various biometric methods, including voice and behavioral recognition, holds immense promise for creating comprehensive and secure authentication processes. Enhanced machine learning models, with their ability to learn and adapt, will further refine facial recognition accuracy, minimizing errors and enhancing system reliability.

Multi-factor authentication (MFA), which combines facial recognition with traditional methods like passwords or geolocation, is expected to gain prominence, providing an extra layer of security for sensitive data and applications. Blockchain technology, with its decentralized and secure nature, could revolutionize identity management, offering a user-controlled and privacy-focused authentication process.

Continuous authentication, which verifies user identity throughout a session using behavioral patterns, will move beyond one-time logins, ensuring ongoing security and preventing unauthorized access. Privacy-centric solutions, such as federated learning, will prioritize user data protection and address growing privacy concerns. Human augmentation, through technologies like augmented reality (AR) glasses or devices, could enhance facial recognition by capturing additional depth and contextual information, improving accuracy and security.

Quantum-safe cryptography, a response to the growing threat of quantum computing, will play a crucial role in safeguarding authentication processes, ensuring long-term security. Standardization and regulations will govern the ethical use of facial recognition technologies, preventing biases and establishing guidelines for responsible development and deployment. User-centric design will remain at the forefront, incorporating user feedback to create systems that are not only secure but also easy and convenient to use.

In summary, the future of user authentication is bright, with a plethora of advancements poised to transform the way we interact with technology. From seamless biometric integration to continuous authentication and privacy-centric solutions, the field is moving towards a more secure, convenient, and user-friendly future.

REFERENCES

[1] Siamese Neural Network for One-shot Image Recognition by Gregory Koch GKOCH@CS.TORONTO.EDU Richard Zemel ZEMEL@CS.TORONTO.EDU Ruslan Salakhutdinov RSALAKHU@CS.TORONTO.EDU

[2] Bengio, Yoshua. Learning deep architectures for ai. Foundations and Trends in Machine Learning, 2(1):1–127, 2009.

[3] Bromley, Jane, Bentz, James W, Bottou, Leon, Guyon, ´ Isabelle, LeCun, Yann, Moore, Cliff, Sackinger, Ed- ¨ uard, and Shah, Roopak. Signature verification using a siamese time delay neural network. International Journal of Pattern Recognition and Artificial Intelligence, 7 (04):669–688, 1993.

[4] Chopra, Sumit, Hadsell, Raia, and LeCun, Yann. Learning a similarity metric discriminatively, with application to face verification. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, volume 1, pp. 539– 546. IEEE, 2005.

[5] Fe-Fei, Li, Fergus, Robert, and Perona, Pietro. A bayesian approach to unsupervised one-shot learning of object categories. In Computer Vision, 2003.

[6] Proceedings. Ninth IEEE International Conference on, pp. 1134– 1141. IEEE, 2003. Fei-Fei, Li, Fergus, Robert, and Perona, Pietro. One-shot learning of object categories. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 28(4):594– 611, 2006.

[7] Hinton, Geoffrey, Osindero, Simon, and Teh, YeeWhye. A fast learning algorithm for deep belief nets. Neural computation, 18(7):1527–1554, 2006.

[8] Krizhevsky, Alex, Sutskever, Ilya, and Hinton, Geoffrey E. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, pp. 1097–1105, 2012.

[9] Lake, Brenden M, Salakhutdinov, Ruslan, Gross, Jason, and Tenenbaum, Joshua B. One shot learning of simple visual concepts. In Proceedings of the 33rd Annual Conference of the Cognitive Science Society, volume 172, 2011.

[10] Lake, Brenden M, Salakhutdinov, Ruslan, and Tenenbaum, Joshua B. Concept learning as motor program induction: A large-scale empirical study. In Proceedings of the 34th Annual Conference of the Cognitive Science Society, pp. 659–664, 2012.

[11] Lake, Brenden M, Salakhutdinov, Ruslan R, and Tenenbaum, Josh. One-shot learning by inverting a compositional causal process. In Advances in neural information processing systems, pp. 2526–2534, 2013. Lake, Brenden M, Lee, Chia-ying, Glass, James R, and Tenenbaum, Joshua B. One-shot learning of generative speech concepts. Cognitive Science Society, 2014.

[12] Lim, Joseph Jaewhan. Transfer learning by borrowing examples for multiclass object detection. Master's thesis, Massachusetts Institute of Technology, 2012.

[13] Maas, Andrew and Kemp, Charles. One-shot learning with bayesian networks. Cognitive Science Society, 2009. Mnih, Volodymyr. Cudamat: a cudabased matrix class for python. 2009.

[14] Palatucci, Mark, Pomerleau, Dean, Hinton, Geoffrey E, and Mitchell, Tom M. Zero-shot learning with semantic output codes. In Advances in neural information processing systems, pp. 1410–1418, 2009.

[15] Simonyan, Karen and Zisserman, Andrew. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

[16] Srivastava, Nitish. Improving neural networks with dropout. Master's thesis, University of Toronto, 2013.

[17] Taigman, Yaniv, Yang, Ming, Ranzato, Marc'Aurelio, and Wolf, Lior. Deepface: Closing the gap to human-level performance in face verification. In Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on, pp. 1701–1708. IEEE, 2014.

[18] Wu, Di, Zhu, Fan, and Shao, Ling. One shot learning gesture recognition from rgbd images. In Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on, pp. 7–12. IEEE, 2012.

[19] Analitic Vidya Fundamentals of Deep Learning, Activation Function Dishasree26 Gupta Aug,2023.

[20] Medium based Understanding of Convolution Neural Network – Deep Learning Prabhu Mar,2018