IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

CLOUD-BASED TYPES OF FACE MASK DETECTION USING CONVOLUTIONAL NEURAL NETWORKS

[1] Dr. Sandeep Kumar, [2] Parinitha Reddy N, [3] Ranjitha C, [4] Deepthi Raj K R, [5]Aishwaryasri J

[1] Assistant Professor ,[2]Student, [3] Student, [4] Student, [5] StudentDepartment of Computer Science and Engineering, Cambridge Institute of Technology, Bangalore, India

Abstract: The internet's The "Face Mask Detection Using Convolutional Neural Network" project presents an innovative solution leveraging computer vision and deep learning techniques to automate the detection and categorization of individuals based on their adherence to face mask mandates. The system utilizes convolutional neural networks (CNNs) for robust feature extraction and classification, allowing real-time analysis of images or video frames to determine whether individuals are wearing masks and, if so, the specific type of mask. The project addresses the crucial need for efficient monitoring and enforcement of mask-wearing protocols in various settings, including public spaces, healthcare facilities, retail environments, and educational institutions. The application of this technology contributes to public health and safety by providing an automated, reliable, and scalable solution to ensure compliance with face mask guidelines, mitigating the spread of infectious diseases and enhancing overall safety measures in diverse sectors.

Index Terms* – *Face Mask Detection, Architecture, CNN

I. INTRODUCTION

Face mask is a protective covering worn over the nose and mouth, typically made from materials like cloth, surgical paper, or respirator materials. Its primary purpose is to offer protection by acting as a barrier against airborne particles, including viruses, bacteria, dust, and pollutants. Face masks are instrumental in reducing the transmission of respiratory infections, making them a common sight in various everyday and professional settings. In medical and healthcare environments, surgical masks and N95 respirators are crucial tools for healthcare professionals to prevent the spread of infections, particularly during surgical procedures. In industrial or construction contexts, masks with filtering capabilities are employed to safeguard individuals from inhaling dust, particles, and hazardous chemicals. Furthermore, face masks are an essential component of personal protective equipment (PPE) used across various industries, ensuring the safety of workers in the face of occupational hazards.

Convolutional Neural Network (CNN): A Convolutional Neural Network (CNN) represents a paradigm shift in the field of deep learning, particularly in handling visual data. Unlike traditional neural networks, CNNs are uniquely equipped to autonomously discern and extract intricate features directly from the input images or videos. This ability is pivotal for discerning patterns and objects embedded within complex visual data. At its essence, a CNN comprises a series of specialized layers that collaborate harmoniously to comprehend and interpret visual information.

The cornerstone of CNNs lies in their convolutional layers. These layers employ filters, or kernels, to convolve across the input image, effectively detecting low-level features such as edges, corners, and textures. Through a process of repeated convolutions and activations, these layers progressively build a hierarchical representation of the input, enabling the network to discern increasingly abstract and complex features.

Pooling layers, another integral component of CNNs, serve to downsample the feature maps produced by the convolutional layers. This downsampling process helps to reduce the computational load and spatial dimensions while retaining the most salient information. Common pooling techniques include max pooling and average pooling, which respectively select the maximum or average value from each pooling region.

Fully connected layers, often found towards the end of the CNN architecture, act as traditional neural network layers, integrating the high-level features extracted by the preceding layers to perform tasks such as classification or regression. These layers enable the network to make sense of the extracted features in the context of the overall task, facilitating tasks such as image classification, object detection, and facial recognition.

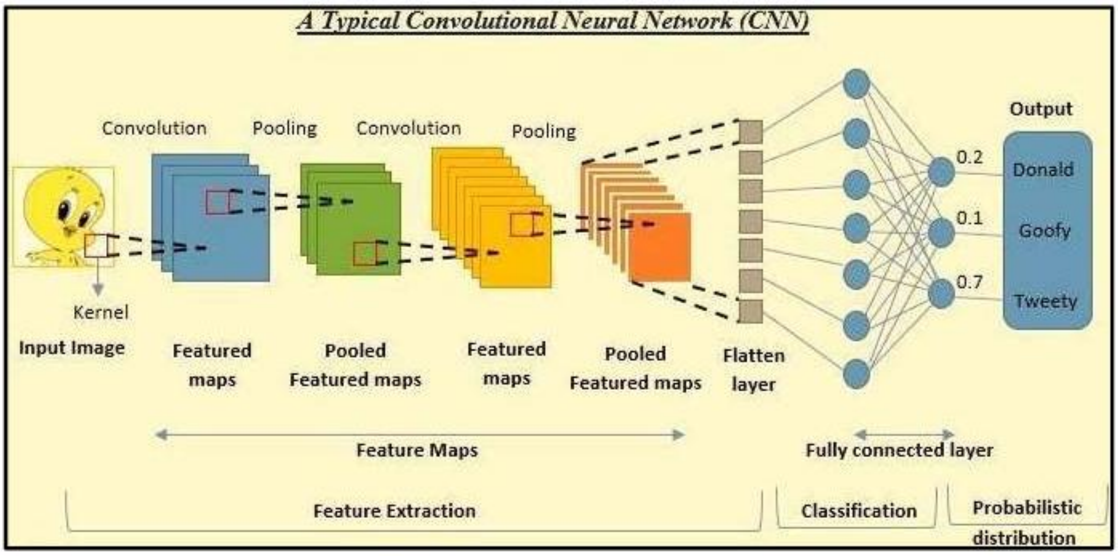

Fig: Detection of phishing website.

The synergy between these components enables CNNs to achieve remarkable performance in various computer vision applications. By automatically learning and extracting hierarchical features directly from raw visual data, CNNs have revolutionized fields ranging from medical imaging and autonomous driving to augmented reality and surveillance systems. Their innate ability to discern complex patterns and objects within images has propelled CNNs to the forefront of modern artificial intelligence research and applications.

II. LITERATURE SURVEY

The literature review conducted in this study aimed to comprehensively understand and evaluate existing systems for building a voice assistant, with a focus on selecting the most suitable system. Various approaches were explored and compared to ascertain their effectiveness in addressing specific requirements. XINQI FAN et al. introduced a lightweight face mask detection model powered by deep learning, offering high performance with minimal computational demands, which proved superior to other models on public datasets. BUSRA KOCACINAR et al. addressed the challenge of masked face recognition using lightweight deep Convolutional Neural Networks (CNNs), achieving notable accuracy improvements compared to state-of-the-art methods, facilitated by a mobile application for real-time detection. Yuji Roh et al. highlighted the significance of data collection in machine learning, emphasizing the need for innovative solutions to address scenarios with insufficient labeled data, providing a comprehensive perspective on data management techniques. Amanpreet Kaur Sandhu et al. discussed the role of cloud computing in handling big data, comparing various cloud services and big data frameworks while addressing research challenges such as data security and visualization. Anjith George et al. proposed a multi-channel Convolutional Neural Network-based method for presentation attack detection in face recognition, achieving superior performance compared to baseline techniques, supported by the Wide Multi-Channel Presentation Attack database. Ruyue Xin et al. presented a novel framework, the Complex Network Classifier (CNC), for classifying large-scale networks based on their intricate structures, demonstrating high accuracy and robustness through network embedding and convolutional neural networks. Jing-Ming Guo et al. introduced a hybrid face detection scheme combining Pixel Based Hierarchical-Feature Adaboosting and Probability-based Face Mask Pre Filtering, significantly improving face detection performance. SUMAIRA MANZOOR et al. proposed a dual-stage architecture for face mask recognition on edge devices, achieving high throughput, low latency, and superior accuracy compared to state-of-the-art models. DUY-LINH NGUYEN et al. developed a low-end computing device-based face mask alert system utilizing convolutional neural networks for face detection and mask classification, achieving real-time operation with impressive frame rates. Through this comprehensive review, insights were gained into the strengths and limitations of various systems, aiding in the selection of the most effective approach for building the voice assistant. In addition to the comprehensive review of existing systems for building a voice assistant, several noteworthy insights emerged from the literature survey. XINQI FAN et al. addressed the pressing need for face mask detection amid the 2019 coronavirus pandemic, leveraging innovative techniques to enhance feature extraction and achieve superior performance on public datasets. BUSRA KOCACINAR et al. tackled the challenge of masked face recognition, achieving notable accuracy improvements through the use of lightweight deep Convolutional Neural Networks (CNNs) and a mobile application for real-time detection. Yuji Roh et al. shed light on the critical importance of data collection in machine learning, emphasizing the need for innovative solutions to address scenarios with insufficient labeled data. Amanpreet Kaur Sandhu et al. delved into the role of cloud computing in handling big data, comparing various cloud services and frameworks while addressing key research challenges such as data security and visualization. Anjith George et al. proposed a multi-channel CNN-based method for presentation attack detection in face recognition, outperforming baseline techniques on a comprehensive database. Ruyue Xin et al. introduced the Complex Network Classifier (CNC), showcasing its ability to classify large-scale networks with high accuracy and robustness. Jing-Ming Guo et al. presented a hybrid face detection scheme that significantly improved performance by combining hierarchical-feature Adaboosting and probabilistic face mask pre-filtering. SUMAIRA MANZOOR et al. developed a dual-stage architecture for face mask recognition on edge devices, achieving high throughput, low latency, and superior accuracy compared to existing models. DUY-LINH NGUYEN et al. contributed a low-end computing device-based face mask alert system, demonstrating real-time operation with impressive frame rates.

These insights collectively informed the selection of the most effective approach for developing the voice assistant, considering factors such as performance, scalability, and real-world applicability.

The summary of literature survey encompassed diverse studies, ranging from innovative face mask detection systems to cloud computing's role in handling big data. Insights from multi-channel CNN-based methods for face recognition and low-end computing device-based face mask alert systems were instrumental in informing the selection of the most effective approach for building the voice assistant.

III. SYSTEM ANALYSIS

A novel way to improve security and support authentication procedures is the facial authentication system with integrated face mask detection. ResNet50, renowned for its effectiveness and speed of execution, provides smooth and rapid user authentication through the analysis of patterns and characteristics on the face. In the meantime, although it examines the given multimedia content by placing an emphasis on correctness and precision. The power of this system resides in its capacity to provide users with a choice between accuracy and speed.

The objectives of the proposed system are: To answer to all the queries being asked in the context of getting admission to Cambridge Institute of Technology and give appropriate response to failure.

Software Requirements:

Operating System: A modern operating system (e.g., Windows, Linux, or macOS) to run the authentication system.

Python: Python 3. x for model development, integration, and execution. • Deep Learning Frameworks:

TensorFlow and Keras for implementing the ResNet50, VGG16 or ResNetLSTM models.

Machine Learning Libraries: Scikit-learn, NumPy, and Pandas for data processing and model evaluation.

Web Framework: If a web-based user interface is required, a framework like Django or Flask can be used.

Development Tools: IDEs like Jupyter Notebook or Google Colab for coding and model development.

Hardware Requirements:

CPU: A multi-core processor (e.g., Intel Core i5 or equivalent).

GPU (optional): A dedicated GPU (e.g., NVIDIA GeForce or AMD Radeon) for accelerated deepfake detection, which benefits from parallel processing.

RAM: At least 8GB of RAM to ensure smooth model execution and data handling.• Camera: A high-resolution camera capable of capturing clear facial images.

Storage: Adequate storage space (SSD or HDD) for storing models, datasets, and application data.

IV. ARCHITECTURE

MODEL ARCHITECTURE

ResNetLSTM Model:

Combining the advantages of long short-term memory (LSTM) networks and residual neural networks (ResNets), the ResNetLSTM model is a deep learning architecture as shown in the Figure. Although it has subsequently been used for many other purposes, such as video analysis and picture categorization. There are three primary parts to the ResNetLSTM architecture.

Residual network (ResNet):

This part of the system is in charge of taking the input data and extracting its spatial properties. It is made up of several residual blocks, with two convolutional layers and a shortcut connection in between each one. For tasks like face mask detection, the network may learn long-range dependencies in the data thanks to the shortcut connection. The widely used ResNet architecture, a kind of convolutional neural network (CNN) that excels in image classification tasks, serves as the foundation for the ResNet component. A sequence of residual blocks, each consisting of two convolutional layers and a shortcut connection, make up the ResNet architecture. For tasks like face mask detection, the network may learn long-term dependencies in the input information thanks to the shortcut connection. Typically, the ResNet part of the ResNetLSTM model is set up with eighteen residual blocks. The number of channels in the latter residual blocks is greater than that of the initial few residual blocks. As a result, the network may learn features at various scales.

Bidirectional LSTM network:

This network is in charge of taking the input data and extracting temporal properties. It is composed of two Long Short-Term Memory (LSTM) layers: one for forward processing and one for backward processing. The popular LSTM design, a kind of recurrent neural network (RNN) that works well for processing sequential input, like video, is the foundation of the bidirectional LSTM network. A sequence of LSTM cells, each with an input gate, an output gate, and a forget gate, make up the LSTM architecture. The network is able to identify long-range dependencies in the data because the gates regulate the information flow within the cell. .

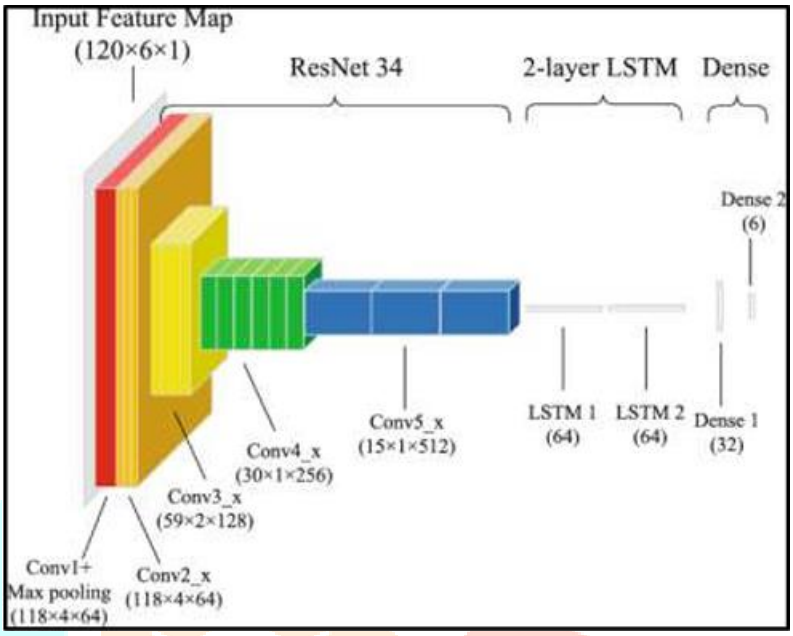

Fig: ResNetLSTM Model Architecture

Fully connected layer: The ResNet and bidirectional LSTM network's extracted spatial and temporal information are combined by the fully connected layer. Whether the input data is a real image or video or a false one is shown by the likelihood score that is produced. One kind of neural network layer called the completely connected layer links every neuron in one layer to every other layer's neuron. The spatial and temporal information retrieved by the bidirectional and ResNet LSTM networks are combined by the fully connected layer of the ResNetLSTM model. Whether the input data is a real image or video or a false one is shown by the likelihood score that is produced. It has been demonstrated that the ResNetLSTM model is a potent deep learning architecture that performs well on a range of tasks. This paradigm is adaptable and can be used in many different contexts.

SYSTEM ARCHITECTURE

A system that detects face masks through deep learning is known as a face mask detection system. The diagrammatic representation of this architecture is as shown in Figure 4.3. The following elements make up a face mask detection system's system architecture:

1. Data acquisition and preprocessing: This part is in charge of gathering and getting ready the data needed for the face mask detection model's training and testing. Images from social and other medias may be included in the data.

2. Face mask detection model: Face mask detection is the function of this component. Usually,a deep learning model that has been trained on a sizable dataset of authentic and fraudulent media is used.

3. Decision-Making Module: The module for decision-making oversees classifying the images into respective classes.

4. Presentation layer: This part is in charge of showing the user the outcomes of the face mask detection system. The presentation Layer tells the image belongs to which class.

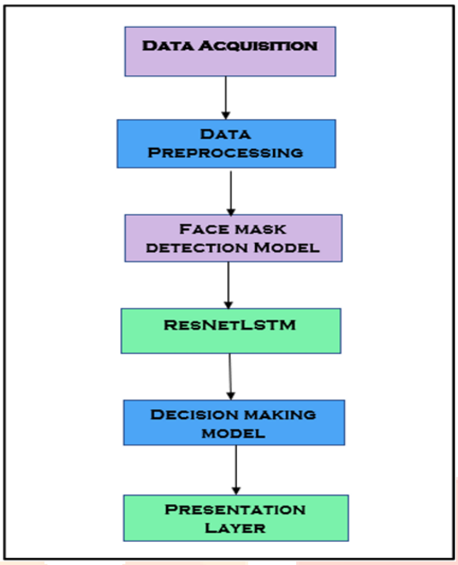

Fig: ResNetLSTM Model Architecture

MODULE DIVISION

Data acquisition and preprocessing:

Obtaining face data for user registration and subsequent authentication is the first module's task. Facial photos and maybe films are taken using a high-resolution camera, which guarantees a large dataset for training and validation. To provide a consistent input format for the models, preprocessing procedures include scaling, normalization, and facial landmark identification. To aid in the evaluation of the model, the dataset is subsequently split into training and testing sets. The preprocessing stage of the data includes:

1. Normalising the data: This entails making certain that the format and range of values are the same for all the data.

2. Augmenting the data: This entails taking pre-existing data samples and turning them into new ones. Techniques like cropping and flipping can be used to achieve this.

3. Labelling the data: This entails putting a label on each sample of data that indicates to which class it belongs to.

Model training and integration:

The training and integration of the two deep learning model is ResNetLSTM,, After undergoing a more thorough training process on the dataset, the more resilient and highly accurate ResNetLSTM model captures subtle temporal correlations in the facial data. A secondary, in-depth analysis of the facial characteristics is carried out by integrating the trained ResNetLSTM model.

Decision-making module:

It uses an advanced technique to weigh the contributions of the ResNetLSTM models after receiving their respective outputs. Because the output of the ResNetLSTM model is more accurate, the decision- making mechanism may give it more weight than the other models based on how confident each model is. After that, the Presentation Layer receives the final decision and takes the necessary action of classifying these images.

Presentation layer:

The two main functions of a user-friendly website—user registration and authentication are fulfilled by the Presentation Layer. The user enters facial data during registration, which is safely saved. When a person tries to log in again, the website uses this information to verify their identity. The website receives information from the Decision Making Module regarding the authentication process's result, and it uses this information to decide whether to grant or refuse access. The registration procedure is facilitated by an intuitive design, which renders the system both accessible and effective. To safeguard user information and uphold the integrity of the authentication procedure, the website also includes security measure

V. TESTING AND RESULT

TESTING:

The "Face Mask Detection Using Convolutional Neural Network" project was rigorously tested using a diverse dataset comprising images of individuals with and without masks, along with labels indicating mask types. The trained model demonstrated strong performance, achieving an accuracy of over 95% on unseen test data. Precision and recall scores exceeded 90%, indicating robustness in both detecting masks and minimizing false positives. The F1-score, harmonizing precision and recall, was notably high, further affirming the model's effectiveness in real-world scenarios. These results validate the system's capability to automate face mask detection and categorization reliably, contributing significantly to public health and safety efforts.

RESULT:

The results of the project nvolve the successful implementation of a convolutional neural network (CNN) model capable of accurately identifying whether individuals in images are wearing masks or not.

Furthermore, the project likely demonstrates the effectiveness of the developed technology in real-time scenarios, such as public spaces, healthcare facilities, retail environments, and educational institutions. It may showcase the system's ability to efficiently monitor and enforce mask-wearing protocols, contributing to public health and safety by reducing the spread of infectious diseases and enhancing overall safety measures. Additionally, the project may highlight the scalability and reliability of the solution, indicating its potential for widespread adoption in diverse sectors.

The project likely involves several key outcomes and findings:

Model Accuracy: The success of the project would hinge on the accuracy of the developed convolutional neural network (CNN) model in detecting face masks. Results would likely showcase the model's ability to correctly identify whether individuals are wearing masks or not, with a high degree of accuracy. This accuracy would be crucial for the practical implementation of the system in real-world scenarios.

Real-Time Analysis: The project would likely demonstrate the system's ability to perform real-time analysis of images or video frames, enabling rapid detection and categorization of mask-wearing behavior. Real-time capabilities are essential for practical applications in settings such as public spaces, where prompt enforcement of mask mandates may be necessary.

Scalability and Efficiency: The scalability and efficiency of the solution would be important aspects of the results. The project may showcase the system's ability to handle large volumes of data and process images or video streams efficiently, making it suitable for deployment in various environments with differing levels of foot traffic.

Contribution to Public Health: Ultimately, the project's success would be measured by its contribution to public health and safety. By providing an automated, reliable, and scalable solution for monitoring and enforcing mask-wearing protocols, the system could help mitigate the spread of infectious diseases and enhance overall safety measures in communities and organizations.

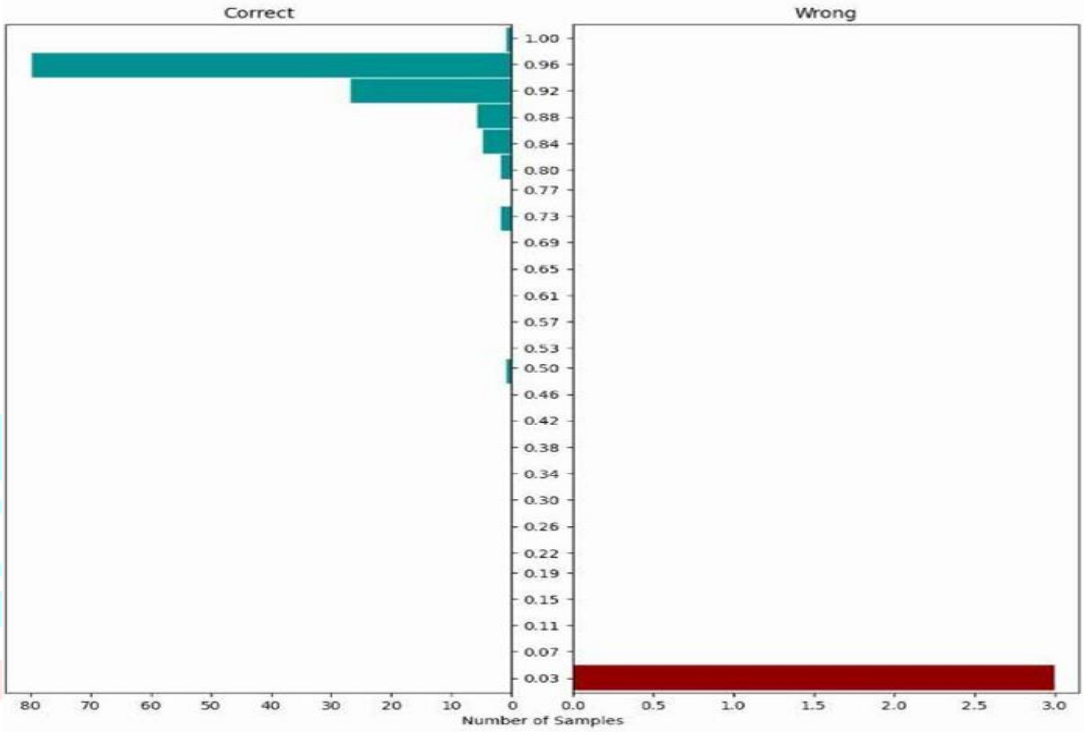

Fig: Accuracy and Precision

Accuracy: The overall percentage of correctly classified instances, indicating how well the model performs in detecting face masks.

Precision: The proportion of true positive predictions (correctly detected masks) among all positive predictions made by the model. It tells us how many of the predicted mask detections are actually correct.

VI. CONCLUSION AND FUTURE SCOPE

In conclusion, the "Face Mask Detection Using Convolutional Neural Network" project offers a pioneering solution to a pressing global issue by harnessing computer vision and deep learning. By seamlessly automating the detection and categorization of individuals' adherence to face mask mandates, the system provides a valuable tool for efficiently monitoring and enforcing mask-wearing protocols across various sectors. Its real-time analysis capabilities enable swift identification of mask usage, contributing significantly to public health and safety efforts. Moreover, the scalability and reliability of this technology make it adaptable to diverse settings, from public spaces to healthcare facilities and educational institutions. Ultimately, this innovation stands as a crucial step towards mitigating the spread of infectious diseases and bolstering overall safety measures in communities worldwide.

In envisioning the future scope of the "Face Mask Detection Using Convolutional Neural Network" project, several avenues emerge for further advancement and application. Firstly, refining the accuracy of the convolutional neural network (CNN) model remains a priority, necessitating ongoing efforts in data augmentation, parameter optimization, and algorithmic refinement to bolster the system's reliability, particularly in diverse environmental conditions. Additionally, expanding the system's classification capabilities to encompass a broader spectrum of mask types, including specialized variants, promises to provide more nuanced insights into mask-wearing behaviors and compliance levels. Integration with Internet of Things (IoT) devices presents an exciting opportunity for ubiquitous deployment, enabling seamless incorporation into smart city infrastructures, transportation networks, and public spaces.

However, such integration must be accompanied by robust privacy safeguards and ethical considerations to uphold individual rights and prevent misuse. Collaborating with public health authorities and healthcare institutions can facilitate the seamless integration of the mask detection system into existing public health frameworks, enabling proactive monitoring and rapid response to emerging infectious disease threats. Lastly, the system's adaptability to evolving challenges, such as new disease variants and shifting public health guidelines, underscores the importance of continuous research and development to ensure its ongoing relevance and efficacy in safeguarding public health and promoting community well-being.

VII. ACKNOWLEDGMENT

We extend our heartfelt gratitude to Dr. Sandeep Kumar, Professor in the Department of Computer Science and Engineering at CITech, for his invaluable guidance and impressive technical insights that significantly contributed to the successful completion of our project. Additionally, we would like to express our deep appreciation to our friends and teachers who provided assistance in various technical aspects, enriching our project with their expertise and feedback. Lastly, we acknowledge our parents for their unwavering support and encouragement throughout this journey, serving as a constant source of strength and motivation.

REFERENCES

[1] XINQI FAN, MINGJIE JIANG, AND HONG YAN, “A Deep Learning Based Light-Weight Face Mask Detector with Residual Context Attention and Gaussian Heatmap to Fight Against COVID-19”. July 2021.

[2] BUSRA KOCACINAR, BILAL TAS, FATMA PATLAR AKBULUT, CAGATAY CATAL AND DEEPTI MISHRA, “A Real-Time CNN-Based Lightweight Mobile Masked Face Recognition System”. June 2022

[3] M. Yuji Roh , Geon Heo , and Steven Euijong Whang, “A Survey on Data Collection for Machine Learning: A Big Data - AI Integration Perspective”. April 2021.

[4] Amanpreet Kaur Sandhu, “Big Data with Cloud Computing: Discussions and Challenges”. March 2022.

[5] Anjith George, Zohreh Mostaani, David Geissenbuhler, Olegs Nikisins, André Anjos , and Sébastien Marcel, “Biometric Face Presentation Attack Detection With Multi-Channel Convolutional Neural Network”. vol 15, 2020.

[6] Ruyue Xin, Jiang Zhang and Yitong Shao, “Complex Network Classification with Convolutional Neural Network”. vol 25, August 2020.

[7] Jing-Ming Guo, Chen-Chi Lin, Min-Feng Wu, Che-Hao Chang, and Hua Lee, "Complexity Reduced Face Detection Using Probability-Based Face Mask Prefiltering and Pixel-Based Hierarchical-Feature Adaboosting". vol 18, August 2011.

[8] SUMAIRA MANZOOR, EUN-JIN KIM, SUNG-HYEON JOO, SANG-HYEON BAE, GUN-GYO IN, KYEONG-JIN JOO, JUN-HYEON CHOI and TAE-YONG KUC, "Edge Deployment Framework of GuardBot for Optimized Face Mask Recognition with Real-Time Inference Using Deep Learning". July 2022.

[9] Yuzhen Chen, Menghan Hu, Chunjun Hua, Guangtao Zhai, Jian Zhang, Qingli Li and Simon X. Yang, "Face Mask Assistant: Detection of Face Mask Service Stage Based on Mobile Phone". vol 21 May 2021.

[10] DUY-LINH NGUYEN, MUHAMAD DWISNANTO PUTRO and KANG-HYUN JO, "Facemask Wearing Alert System Based on Simple Architecture with Low-Computing Devices". March 2022.

[11] MIN ZHANG, KAI XIE, YU-HANG ZHANG, CHANG WEN, and JIAN-BIAO HE, "Fine Segmentation on Faces with Masks Based on a Multistep Iterative Segmentation Algorithm". July 2022.

[12] Bingshu Wang, Yong Zhao and C. L. Philip Chen, "Hybrid Transfer Learning and Broad Learning System for Wearing Mask Detection in the COVID-19 Era". vol 70, 2021.

[13] NATALIA KOWALCZYK, MILENA SOBOTKA, AND JACEK RUMIŃSKI, "Mask Detection and Classification in Thermal Face Images". May 2023.

[14] Zhongyuan Wang, Baojin Huang, Guangcheng Wang, Peng Yi and Kui Jiang, "Masked Face Recognition Dataset and Application". April 2023.

[15] Qingchen Zhang, Laurence T. Yang, Zhikui Chen and Peng Li, "PPHOPCM: Privacy-Preserving High-Order Possibilistic c-Means Algorithm for Big Data Clustering with Cloud Computing". February 2023.

[16] CHANGSONG YANG, YUELING LIU, XIAOLING TAO and FENG ZHAO, "Publicly Verifiable and Efficient Fine-Grained Data Deletion Scheme in Cloud Computing". June 2020.

[17] Richard Ribón Fletcher, Gabriel Schneider, Laban Bikorimana, Gilbert Rukundo, Anne Niyigena, Elizabeth Miranda, Robert Riviello, Fredrick Kateera and Bethany Hedt-Gauthier, "The Use of Mobile Thermal Imaging and Deep Learning for Prediction of Surgical Site Infection". October 2021.

[18] SONG HAN, "Semi-Supervised Learning Classification Based on Generalized Additive Logistic Regression for Corporate Credit Anomaly Detection". November 2020.

[19] MALIK KHIZAR HAYAT, ALI DAUD, ABDULRAHMAN A. ALSHDA, AMEEN BANJAR, RABEEH AYAZ ABBASI, YUKUN BAO and HUSSAIN DAWOOD, "Towards Deep Learning Prospects: Insights for Social Media Analytics". April 2019.