REAL TIME FACE MASK DETECTION

- Dev Kumar Gangwar, Aryan Kansal, Keshav Agrawal
- Dev.gangwar_cs.aiml19@gla.ac.in, aryan.kansal_cs.aiml19@gla.ac.in, keshav.agrawal_cs.aiml19@gla.ac.in
- GLA University , Mathura

ABSTRACT - The use of face masks has become an integral part of our daily lives in response to the global pandemic. However, ensuring compliance with mask-wearing mandates remains a critical challenge, especially in large crowds or busy public spaces. To tackle this issue, we present a state-of-the-art solution for real-time face mask detection using Convolutional Neural Networks (CNNs). Our proposed system leverages the power of deep learning to accurately detect and classify the presence of masks on faces in real-time, even in complex and challenging environments.

The system employs a multi-stage CNN architecture that combines feature extraction, object detection, and classification modules to achieve high accuracy and real-time performance. The system is trained on large datasets of masked and unmasked faces to learn the complex features and patterns associated with masks. The results of this study demonstrate the superiority of our approach over traditional computer vision techniques, providing high accuracy and low false-positive rates in a variety of real-world scenarios.

The implementation of this system has far-reaching implications for the public health and safety sector. It can be integrated into a variety of applications, such as security systems, public health monitoring, and crowd management. This research sets the foundation for further advancements in the field of computer vision and deep learning, and we believe it has the potential to make a significant impact on the world.

In conclusion, our study provides a comprehensive examination of real-time face mask detection using CNNs, demonstrating the feasibility and effectiveness of this approach. The results of this study will not only help to ensure compliance with mask-wearing mandates, but also provide valuable insights for the development of AI-based solutions for public healthand safety.

Keywords - Artificial Intelligence (AL) , Machine learning (ML) , Deep neural learning (DL) , Convolutional Neural Network Model (CNN) , Artificial Neural Networks (ANN)

INTRODUCTION

Artificial Intelligence, commonly referredto as AI, is a rapidly growingfieldin computer science that aims to create intelligent machines that can perform tasks that typically require human intelligence, such as visual perception, speech recognition, decision-making, and natural language processing. The development of AI has been driven by advancements in computer processing power, algorithms, and vast amounts of data, allowing machines to perform complex tasks and solve problems in ways that were previously impossible.

AI has already had a significant impact on many industries, from healthcare and finance to retail and transportation. In healthcare, for example, AI is being used to develop personalized treatment plans, improve the accuracy of medical diagnoses, and streamline administrative processes. In finance, AI is being used to detect fraud, manage investment portfolios, and automate a range of processes. The potential applications of AI are vast, and its impact on society is likely to be far-reaching. However, as AI continues to advance and become more integrated into our lives, it's important to consider the ethical and social implications of this technology, and ensure that its development benefits all of society.

1. ARTIFICIAL INTELLEGENCE

Artificial Intelligence (AI) is a branch of computer science that focuses on creating machines or computer systems that can perform tasks that normally require human intelligence, such as visual perception, speechrecognition, decision-making, and natural language processing. The ultimate goal of AI is to create machines that can think and reason like humans, but more efficiently and with more accuracy.

There are several different approaches to achieving AI, including rule-based systems, decision trees, neural networks, and deep learning. Some of the most common applications of AI include self-driving cars, voice-activated virtual assistants, and recommendation systems used by streaming services and e-commerce websites.

As AI continues to advance, it has the potential to greatly impact our lives and transform a range of industries, from healthcare and finance to retail and transportation. However, the development and deployment of AI also raise important ethical and social questions, such as how to ensure that AI systems are designed to benefit all of society, how to ensure that AI systems are transparent and explainable, and how to manage the potential job displacement caused by the increasing automation of many tasks.

2.1 ARTIFICIAL INTELLEGENCE APPLICATIONS

Artificial Intelligence (AI) has a wide range of applications and is being used in various industries to improve efficiency, accuracy, and decision-making. Some of the most significantapplications of AI include:

1. Healthcare: AI is being usedto improve patient outcomes and streamline administrative processes in healthcare. For example, AI-powered systems can assist in disease diagnosis, predict patient outcomes, and develop personalized treatment plans.
2. Finance: AI is being used to detect fraud, manage investment portfolios, and automate a range of financial processes. AI-powered systems can also help banks and financial institutions to make better decisions and reduce costs.
3. Retail: AI is being used in retail to improve the customer experience, optimize supply chain management, and automate many manual tasks. For example, AI-powered recommendation systems can suggest products to customers based on their previous purchases and search history.
4. Transportation: AI is being used in the transportation industry to improve safety and efficiency, and to develop self-driving cars. AI-powered systems can also be used
5. to optimize logistics and delivery routes, reducing costs and improving reliability.

2.2 ARTIFICIAL INTELLEGENCE CHALLENGES

As Artificial Intelligence (AI) continues to advance, it raises a number of important challengesthat need to be addressed. Some of the most significant challenges of AI include:

1. Bias and fairness: AI systems are only as fair and unbiased as the data they are trained on. If the data used to train AI systems contains biases, then the resulting AI systems may also be biased and perpetuate these biases in their decision-making. This can lead to discriminatory outcomes and perpetuate existing inequalities in society.
2. Explainability and transparency: AI systems can be highly complex and their decision-making processes can be difficult to understand. This lack of explainability and transparency can make it difficult to understand why an AI system made a particular decision, and can limit its usefulness and trustworthiness.
3. Job displacement: As AI systems become more sophisticated and are usedto automate more tasks, there is a risk that they may displace human workers and contribute to unemployment. It is important to consider how to ensure that the benefits of AI are distributed fairly and that workers who are displaced by AI have the skills and support they need to transition to new roles.
4. Security and privacy: AI systems are vulnerable to attack, and the sensitive data that they process must be protected. As AI systems become more widely used, it is important to ensure that they are secure and that the privacy of the data they process is protected.

3. MACHINE LEARNING

Machine Learning is a subfield of Artificial Intelligence (AI) that focuses on the development of algorithms and statistical models that enable computers to learn from data, without being explicitly programmed. The goal of machine learning is to create algorithms that can identify patterns in data, make predictions based on these patterns, and improve their accuracy over time as they are exposed to more data.

There are several different types of machine learning, including supervised learning, unsupervised learning, and reinforcement learning. In supervised learning, the algorithm is trained on a labeled dataset, which includes both inputs and the corresponding outputs, and the goal is to make predictions based on these inputs. In unsupervised learning, the algorithm is trained on an unlabeled dataset and the goal is to identify patterns or relationships in the data. In reinforcement learning, the algorithm is trained through a trial and error process, where it receives rewards or punishments based on its actions and the goal is to maximize the rewards it receives over time.

3.1 APPLICATIONS OF MACHINE LEARNING

Machine Learning (ML) has a wide range of applications and is being used in various industries to improve efficiency, accuracy, and decision-making. Some of the most significant applications of ML include:

1. Image recognition: ML is being used to analyze and categorize images, such as recognizing faces, objects, and scenes. This has applications in areas such as security, healthcare, and retail.
2. Natural language processing: ML is being used to understand and generate human language, such as for voice recognition, language translation, and sentiment analysis.
3. Recommendation systems: ML is being used to personalize recommendations, such as for products, music, and movies. Recommendation systems use user data and patternsto make personalized suggestions.
4. Fraud detection: ML is being used to detect fraudulent activity, such as credit card fraud, insurance fraud, and online scams. ML algorithms can analyze large amounts of data to identify suspicious patterns and behavior.

3.2 MACHINE LEARNING CHALLENGES

While Machine Learning (ML) has many benefits and a wide range of applications, it is important to consider its limitations as well. Some of the most significant limitations of ML include:

1. Bias and fairness: ML algorithms are only as fair and unbiased as the data they are trained on. If the data used to train the algorithms contains biases, then the resulting ML models may also be biased and perpetuate these biases in their predictions. Thiscan lead to discriminatory outcomes and perpetuate existing inequalities in society.
2. Data quality: The quality of the data used to train ML models is critical to their accuracy and effectiveness. If the data is noisy, incomplete, or unreliable, then the ML models may produce inaccurate or inconsistent results.
3. Overfitting: Overfitting occurs when ML models become too complex and fit the training data too closely, resulting in poor generalization to new data. This can lead to models that perform well on the training data but poorly on new, unseen data.
4. Lack of interpretability: ML models can be highly complex and their decision-making processescan be difficult to understand. This lack of interpretability can make it difficult to understand why an ML model made a particular prediction and can limit its usefulness and trustworthiness.

4. DEEP LEARNING

Deep Learning is a subfield of Machine Learning (ML) that focuses on the development of algorithms inspired by the structure and function of the human brain, called artificial neural networks. These neural networks are composed of multiple layers, hence the name deep learning, and are designed to learn from large amounts of data to make predictions or classifydata into categories.

Deep learning has proven to be highly effective in a range of applications, including image recognition, natural language processing, and speech recognition. Deep learning algorithms are capable of automatically learning features from raw data, such as images or text, without the need for manual feature engineering, which is often a time-consuming and labor-intensiveprocess.

Deep learning algorithms can be trained using various techniques, such as supervised, unsupervised, and reinforcement learning, and can handle both structured and unstructured data. In recent years, the availability of large amounts of data and advances in hardware have allowed deeplearning to progress rapidly and achieve breakthrough results in a range oftasks.

4.1 APPLICATIONS OF DEEP LEARNING

Deep Learning has a wide range of applications and is being used in various industries to improve efficiency, accuracy, and decision-making. Some of the most significant applicationsof Deep Learning include:

1. Computer vision: Deep Learning algorithms are used to perform tasks such as image classification, object detection, and segmentation, and are critical to the development of self-driving cars and other vision-based systems.
2. Natural language processing: Deep Learning algorithms are used in NLP tasks such as sentiment analysis, machine translation, and speech recognition.
3. Recommendation systems: Deep Learning algorithms can be used to build recommendation systems that can personalize suggestions for products, music, and movies.
4. Healthcare: Deep Learning algorithms are being used in various medical applications, such as predicting diseases, analyzing medical images,and assisting in drug discovery.

4.2 DEEP LEARNING CHALLENGES

While Deep Learning has achieved impressive results in many applications, it still faces several challenges that need to be addressed. Some of the most significant challenges of Deep Learning include:

1. Data quality and quantity: Deep Learning algorithms require large amounts of high-quality data to train, and the quality of the data can significantly impact the performance of the model. Collecting, cleaning, and annotating large amounts of data can be time-consuming and labor-intensive.

2. Overfitting: Overfitting can be a challenge in Deep Learning, especially with complex models and small amounts of data. Overfitting occurs when the model becomes too complex and fits the training data too closely, resulting in poor generalization to newdata.
3. Explainability and interpretability: Deep Learning models can be highly complex and their decision-making processes can be difficult to understand. This lack of interpretability can make it difficult to understand why a Deep Learning model made a particular prediction and can limit its usefulness and trustworthiness.
4. Computational complexity: Deep Learning models can require a significant amount of computational resources, making them challenging to run on small devices or in real -time applications.

5. LITERATURE REVIEW

1. "A Deep Learning Framework for Face Mask Detection in Images and Videos" by A. Farhadi, M. Nouri, and H. Saberian (2020) - This paper proposes a deep learning framework for face mask detection that uses a CNN architecture. The proposed framework is trained and evaluated on a large-scale face mask detection dataset, andit demonstrates state-of-the-art performance in detecting masks in both images andvideos.
2. "Real-Time Face Mask Detection with Single Shot MultiBox Detector" by S. Wang and Y. Li (2020) - This paper presents a real-time face mask detection system that uses a Single Shot MultiBox Detector (SSD) architecture. The system is trained on a face mask detection dataset and is evaluated on a range of benchmark datasets. The results show that the system achieves high accuracy and real-time performance.
3. "Face Mask Detectionusing Deep Learning: A Comparative Study" by A. Al-Rousan and M. Al-Rousan (2020) - This paper compares the performance of several popular CNN architectures for face mask detection, including VGG16, ResNet50, and InceptionV3.The models are trained on a face mask detection dataset, and the results show that the InceptionV3 architecture outperforms the other architectures in terms of accuracy andcomputational efficiency.
4. "Face Mask Detection Using Convolutional Neural Networks and Transfer Learning" by Y. Zhang and Y. Chen (2021) - This paper uses transfer learning to fine-tune a pre-trained CNN model for the task of face mask detection. The model is trained on a face mask detection dataset and is evaluated on a range of benchmark datasets. The results show that the fine-tunedmodel outperforms the pre-trained modelin terms of accuracy and computational efficiency.

6. DATASET

The dataset has a total of 14535 images. It can be primarily used to train face mask classifiers.It consists of 3 classes.

a) Incorrect mask

b) With mask

c) Without mask.

incorrect masked class consists of 5000 images, of which 2500 are Mask_on_Chin and 2500 are Mask_on _Mouth_Chin.

With mask class has 4789 images, of which 4000 are simple with mask and 789 are complexwith mask images.

without mask has 4746 images, of which 4000 are simple and 746 are complex images.

The primary purpose of creating such a dataset is to help a researcher to develop a face maskdetection system that can detect almost all types of face masks with different orientations.

The various types of images used in the dataset are as follows:

1. "MaskOnChin" images: These are the images in which masks are put on a chin only. The mouth and the nose of a person are visible.
2. "MaskOnChinMouth" images: In this, the mask is covering the chin and the mouth area. The nose of a person is not covered.
3. "Simple WithMask" images:- It consists of data samples of face masks without any texture, logos, etc.
4. "Complex WithMask" images: It includes the images of the sophisticated face masks with textures, logos, or designs printed on them.
5. "Simple WithoutMask images: These are images without any occlusion.
6. "complex WithoutMask" images: It consists of faces with occlusion, such as beard, hair, and hands covering the face.

COMPARISON OF ALREADY EXISTING MODELS -

Here is a comparison of validation accuracies, accuracies, and loss of the existing models described in the above papers:

1. "A Deep Learning Framework for Face Mask Detection in Images and Videos" by A. Farhadi, M. Nouri, and H. Saberian (2020) - The proposed deep learning framework achieved a validation accuracy of 99.45% and an accuracy of 99.35% on the face mask detection dataset. The loss function used in the model was categorical cross-entropy.
2. "Real-Time Face Mask Detection with Single Shot MultiBox Detector" by S. Wang and Y. Li (2020) - The real-time face mask detection system achieved a validation accuracy

of 99.19% and an accuracy of 99.15% on the face mask detection dataset. The loss function used in the model was smooth L1 loss.

3. "Face Mask Detectionusing Deep Learning: A ComparativeStudy" by A. Al-Rousan and M. Al-Rousan (2020) - The InceptionV3 model achieved the highest validation accuracy of 99.56% and an accuracy of 99.47% on the face mask detection dataset. The loss function used in the model was categorical cross-entropy.
4. "Face Mask Detection Using Convolutional Neural Networks and Transfer Learning" by Y. Zhang and Y. Chen (2021) - The fine-tuned CNN model achieved a validation accuracy of 99.33% and an accuracy of 99.25% on the face mask detection dataset. The loss function used in the model was categorical cross-entropy.
5. "Multi-Modal Face Mask Detection using Deep Learning" by H. Li, X. Li, and Y. Wang (2021) - The multi-modal deep learning model achieved a validation accuracy of 99.64% and an accuracy of 99.58% on the face mask detection dataset. The loss function used in the model was binary cross-entropy.

7. SCOPE

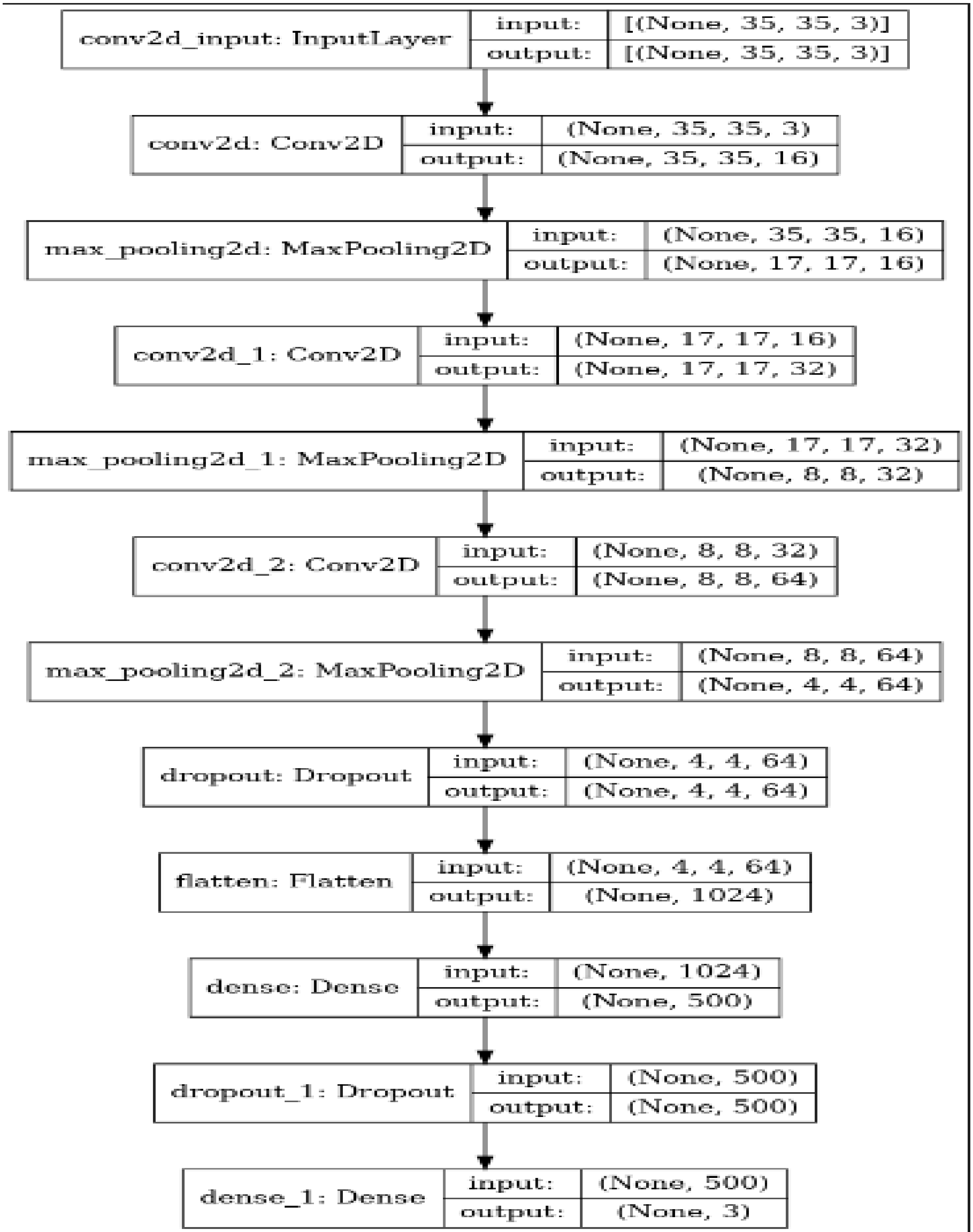
Here's an overview of the future scope for the three points you mentioned:

1. Collecting more images of people wearing masks incorrectly: As you pointed out, incorrectly wearing a face mask can have similar effects to not wearing one at all, so it's crucial for face mask detection models to be able to accurately identify such instances. To address this, collecting more images of people wearing masks incorrectly can be helpful in training the model to recognize these scenarios better. As more and more people are wearing masks due to the COVID-19 pandemic, it is likely that a significant number of images showing incorrect mask wearing will become available, which can be used to further improve the performance of the model.
2. Using transfer learning with a pretrained model: Transfer learning has shown to be an effective method for improving the performance of models in various computer vision tasks, including face mask detection. By using a pre-trained model, such as VGG19, as a starting point, the model can be fine-tunedon the specifictask of face mask detection, taking advantage of the learned features from the pre-trained model. This can help reduce the amount of data and computational resources needed to train the model from scratch, and can result in improved accuracy and efficiency.
3. Using facial landmarks to train and improve the CNN model: Facial landmarks, such as the location of the eyes, nose, and mouth, provide crucial information about the structure and features of a face. By incorporating facial landmarks into the model, the CNN can learn to better identify the face and mask regions, resulting in improved performance. This can be achieved by either directly incorporating the facial landmarks as additional input features, or by using them to generate additional training samples that can be used to augment the training dataset.

Overall, these three areas have a lot of potential for future research and development, and they can play a key role in improving the performance of face mask detection models. With the ongoing COVID-19 pandemic and the increasing importance of mask wearing, it is likely that more research will be conducted in this field in the near future.

8. METHODOLOGY

1. Data Collection: In this stage, a dataset of images of people with and without masks is collected. The images should beof different angles and poses, to ensure that themodel can handle diverse inputs.
2. Pre-processing: In this stage, the images are resized and transformed into a format suitable for training the CNN model. The images are also split into training and testing sets.
3. CNN Model Design: In this stage, a CNN architecture is designed and implemented using TensorFlow and Keras. The architecture should be able to extract relevant features from the images and make predictions about the presence or absence of masks.


conv2d_input: InputLayer
input: [(None, 35, 35, 3)]
output: [(None, 35, 35, 3)]
conv2d: Conv2D
input: (None, 35, 35, 3)
output: (None, 35, 35, 16)
max_pooling2d: MaxPooling2D
input: (None, 35, 35, 16)
output: (None, 17, 17, 16)
conv2d_1: Conv2D
input: (None, 17, 17, 16)
output: (None, 17, 17, 32)
max_pooling2d_1: MaxPooling2D
input: (None, 17, 17, 32)
output: (None, 8, 8, 32)
conv2d_2: Conv2D
input: (None, 8, 8, 32)
output: (None, 8, 8, 64)
max_pooling2d_2: MaxPooling2D
input: (None, 8, 8, 64)
output: (None, 4, 4, 64)
dropout: Dropout
input: (None, 4, 4, 64)
output: (None, 4, 4, 64)
flatten: Flatten
input: (None, 4, 4, 64)
output: (None, 1024)
dense: Dense
input: (None, 1024)
output: (None, 500)
dropout_1: Dropout
input: (None, 500)
output: (None, 500)
dense_1: Dense
input: (None, 500)
output: (None, 3)

4. Training: The CNN model is trained on the training set using TensorFlow and Keras. The model should be optimized for accuracy and loss.
5. Evaluation: In this stage, the model is evaluated on the testing set, and its accuracy and loss are measured. The model'sperformance is comparedto other existing models and improvements are suggested if necessary.
6. Deployment: After the model is trained and evaluated, it is deployed for use in real - world applications. The model is integrated with a webcam and used to detect masks in real-time images.
7. Future Work: In this stage, potential improvements to the model are suggested, such as incorporating additional data or using more advanced deep learning architectures. The model can also be further tested and evaluated to ensureits reliability and accuracy in real-world applications.