

# Approach to Prediction of Unmasked Face from Masked Face using Deep Learning

Apoorva Shete  
Undergraduate Student  
Thadomal Shahani Engineering  
College  
Mumbai, India

Rachael Nihalaani  
Undergraduate Student  
Thadomal Shahani Engineering  
College  
Mumbai, India

Amit Hatekar  
Professor  
Thadomal Shahani Engineering  
College  
Mumbai, India

## ABSTRACT

Due to the outbreak of COVID-19, it has become mandatory for each and every person to step outside only with a face mask on. This has raised the security and safety concerns among people as faces of criminals, burglars, etc are not recognisable through the CCTVs and security cameras. This problem can be tackled with the help of deep learning. In this paper, a model that can predict the unmasked face of a person from a masked face input image, giving an unmasked face image as the output was implemented. The accuracy achieved by the model is 91%. This paper ends with a review of the model's usefulness and its scope for further development and improved results in the future.

## General Terms

Machine Learning, Deep Learning, Autoencoders

## Keywords

Machine Learning, Deep Learning, Autoencoders, Face Mask, Unmasking

## 1. INTRODUCTION

In recent times, people have been forced to cover their faces with masks due to the COVID-19 pandemic that has hit globally. However, some people are taking undue advantage of this situation by hiding their faces behind these masks. The security concerns due to this have risen majorly in recent times and hence there was a need to predict the face of a masked person. This problem is quite difficult because as per the safety guidelines of COVID-19, people are required to cover their nose properly with a surgical mask, some people even prefer double masking for more safety from the virus. Due to this, it is becoming extremely difficult for the authorities to recognise faces through security cameras, CCTVs, etc. It is therefore essential to have a system that can predict and recognise a person's entire face from its image of a masked face.

The recent studies in image correction and object removal from an image using deep learning have achieved good results but they lack when it comes to the removal of large and complex objects, especially from facial images. In this paper, an attempt has been made to remediate this problem of unmasking a person's face using deep learning and tackle the security concerns that have risen in the past due to this issue.

The structure of our paper is as follows: Section 2 gives a detailed review of the previous research done in this field. In section 3, the complete description of the dataset used for this paper has been given along with the methodology followed for implementation of this model. Sections 4 and 5 discuss the results given by this model and the conclusion and the ways in

which this model can be further improved in the future, respectively.

## 2. LITERATURE REVIEW

There has been a lot of research done in the past in the field of image correction and completion using deep learning. However, the concept of using image correction techniques for the prediction of who is behind a face mask is fairly new.

In Paper[1], a model was proposed for interaction-free large object removal from images of faces, focusing a face mask as the object to be removed. A GAN-based image inpainting technique along with an image-to-image translation approach was used for gaining plausible results. Here, additional results were provided by retraining the model for inpainting irregular and random rectangular missing holes as well. This model was able to complete the facial images while retaining the naturalness and structure of the faces. The research done in paper [2], a Residual Convolution Neural Network (ResCNN) is designed to restore the face of an individual behind the mask. To prepare their training dataset, they used a face mask as an object and masked images of faces taken from the CelebA and MFR datasets, to test the effectiveness of the proposed model.

Paper[3] discusses the image completion and extrapolation that has been implemented using a two-stage deep learning framework. In the first stage, the complete human structure was recognized using a parsing network and in the second stage, an image completion network was used to fill in the incomplete region, using the structure recovery map obtained in the first stage. In paper[4], it is observed that the model for object removal and reconstruction of the damaged part by using a GAN setup. The limitations are that this model is restricted to lower resolutions (178×218) and also produces artifacts when the damaged part is at the margins of an image. In paper [5] an object region is automatically detected and removed in general scene-level images. The drawback of this model is that the output majorly depends on automatic object detection which more often than not fails to detect the object region due to large changes in both appearance and structure of the masks and faces. In [6] too, a two-stage model is proposed; the first stage generates coarse output and some guidance information and it is refined in the second stage i.e the complete image is generated using the guidance information from the first stage. Paper[7] uses a two-stage adversarial model with an edge generator followed by an image editing network for editing the image. Paper[8] uses a GAN-based inpainting approach to remove a microphone from the facial images given as input.

This literature study helped us in understanding the factors important for implementing an image completion model and

helped us go about with our research in this field.

### 3. PROPOSED FRAMEWORK

#### 3.1 Dataset

Data collection is a primary and integral step in building any Machine Learning model. For this unmasking model, the dataset used is obtained from Kaggle called the Face Mask Lite Dataset [9] dataset. This dataset contains 20000 image files divided into two halves. 10000 image files are of people without masks, as in only their bare faces. The other 10000 image files have the same images but with a digitally added mask covering the lower halves of their faces. This is quite useful as it enables the model to recognize what the masks look like. Having images of bare faces helps decide the accuracy of the model.

Additionally, there is considerable diversity in the dataset in terms of face shape, hair color, age, wearing glasses, and other facial attributes. Figure 1 below shows a few instances of our dataset.



Fig 1: Few instances of the Face Mask Lite Dataset

Figure 2 below shows a sample of the code used for plotting the image pairs.

```
def plot_image_pair(images = 5):
    for i in range(images):
        plt.figure(figsize = (7,7))
        plt.subplot(1,2,1)
        plt.title("No Mask", fontsize = 15)
        plt.imshow(img_array[i].reshape(SIZE, SIZE, 3))
        plt.subplot(1,2,2)
        plt.title("Mask", fontsize = 15)
        plt.imshow(mask_array[i].reshape(SIZE, SIZE, 3))

plot_image_pair(5)
```

Fig 2: Sample of the code

#### 3.2 Methodology

The main aim of this model is to ‘Unmask’ a person, i.e. to predict the face behind the mask. The model has been implemented in Python and using a special deep learning architecture called Autoencoders. This is a neural network that learns the representation of raw data in compressed form. The encoder submodel of autoencoders learns the reduced dimensional representation of input data via downsampling. On the other hand, the decoder submodel uses upsampling to regenerate data. Both these submodels use convolutional neural networks.

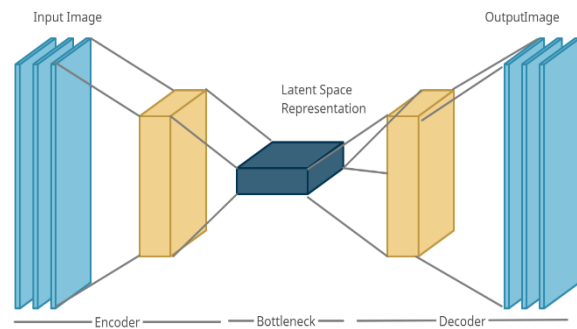


Fig 3: Autoencoder Architecture

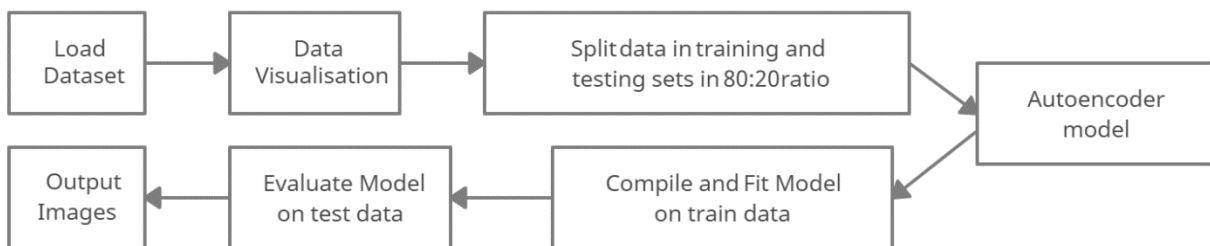


Fig 4: Proposed Framework for model

Implementation has been done using the following steps. First, all required python libraries like Keras, TensorFlow, etc. are imported. Next, the dataset is loaded. Since the Face Mask Lite dataset is huge, we have used only about 30% of the dataset, which is sufficient enough for the model. 3000 masked images and the corresponding 3000 unmasked images were used, making a total of 6000. Post this, a few instances from the dataset were displayed to get familiar with it (Figure 1).

The data is now ready to be used and is split into two sets - one used for training and one for testing the model. About 80% of the loaded data has been used for training while the remaining has been used for testing.

The autoencoder model in this framework has the encoder submodel producing a latent space vector which is then taken as input by the decoder submodel. The former compresses the images using downsampling while the latter attempts to reconstruct these images and simultaneously tries to reduce reconstruction loss, via upsampling. Once the model is defined, it is compiled and fit over the training set.

The next step is to evaluate the model over the testing set. If the accuracy of the model is not good enough, there is a need to tweak the model. Once the desired accuracy is achieved, the output images, which contain the input masked image, the original unmasked image given in the dataset, and the predicted unmasked image given by the model can be displayed.

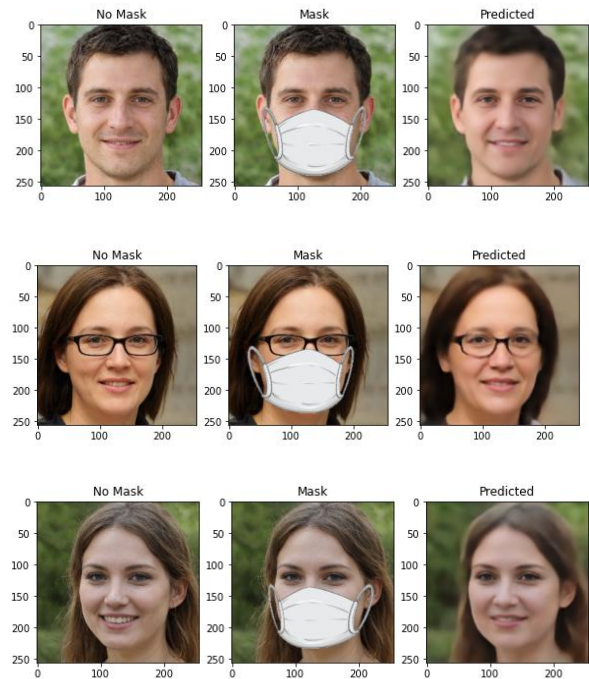
#### 4. RESULTS AND ANALYSIS

As a result of the COVID-19 pandemic, face masks have become an integral part of daily life. The mask In deep learning, the loss can be understood as a penalty for bad predictions. Loss indicates how bad the model's prediction was. If the prediction is perfect, the loss is 0, else it is higher. Since the model cannot make 'right' or 'wrong' predictions, due to a lack of concrete verification by its side, the predictions can be classified as 'good' or 'bad'. In the case of this unmasking model, if a predicted unmasked image is extremely similar to the original unmasked image provided in the dataset, as shown in Figure 5 , the prediction is considered to be 'good'. The other evaluation metric that has been used is accuracy, which is defined as the percentage of correct predictions for test data. The accuracy metric measures the performance of the model in a way that is interpretable. This unmasking model has shown an accuracy of 91% with a loss of 3%.

**Table 1. Results obtained using this model**

| Evaluation Metric | Top |
|-------------------|-----|
| Accuracy          | 91% |
| Loss              | 3%  |

The model output has been shown in Figure 5. Three images from the oupt are displayed here.. The input masked image, the original unmasked image given in the dataset, and the predicted unmasked image given by the model are all displayed side-by-side for visual comparisons.



**Fig 5: Model Output**

#### 5. CONCLUSION AND FUTURE SCOPE

As a result of the COVID-19 pandemic, face masks have become an integral part of daily life. The mask mandate is still in effect in many countries such as India. Although people have begun to return to normal life, they still wear masks. The aim of this model is to find the face under these masks. This can find its application in security as well as safety, for example, people returning to work and public places are usually asked to lower their masks for verification purposes, which can be unsafe in contaminated environments. This model can help in such cases. Other applications can be found in hospitals, where surgeons and certain doctors are always required to wear masks.

The model has shown an accuracy of 91% which is a commendable starting point. It can be expanded on and improved further to make it more reliable to be integrated with CCTV cameras to detect and identify people without masks. This improved model can then be used in real-time applications such as those requiring face-mask detection, in general, or for COVID-19 specifically. The model can also be used to ensure that public safety guidelines are being followed, by integrating with security systems of schools, offices, airports, railway systems, and other public places.

#### 6. ACKNOWLEDGMENTS

Our thanks to Professor Amit Hatekar for his immense support and guidance, and to the Principal, Dr. G.T. Thampi.

#### 7. REFERENCES

- [1] N. Ud Din, K. Javed, S. Bae and J. Yi, "A Novel GAN-Based Network for Unmasking of Masked Face," in IEEE Access, vol. 8, pp. 44276-44287, 2020, doi: 10.1109/ACCESS.2020.2977386.
- [2] F. Bhusra, A. Jhapate.(2021) RESCNN: A Deep Learning Approach for Unmasking Face Masks. Smart Moves Journal, ISSN NO: 2582-4600. [Online]. [Available]:
- [3] <https://ijoscience.com/ojsscience/index.php/ojsscience/ar>

ticle/view/350/815

- [4] X. Wu, F.L. Zhang, J. Wang, et al. (2019). Deep Portrait Image Completion and Extrapolation. [Available]: <https://arxiv.org/pdf/1808.07757.pdf>
- [5] S. Iizuka, E. Simo-Serra and H. Ishikawa, "Globally and locally consistent image completion", *ACM Trans. Graph.*, vol. 36, no. 4, pp. 1-14, Jul. 2017.
- [6] R. R. Shetty, M. Fritz and B. Schiele, "Adversarial scene editing: Automatic object removal from weak supervision", *Proc. Adv. Neural Inf. Process. Syst.*, pp. 7706-7716, 2018.
- [7] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu and T. S. Huang, "Generative image inpainting with contextual attention", *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, pp. 5505-5514, Jun. 2018.
- [8] K. Nazeri, E. Ng, T. Joseph, F. Z. Qureshi and M. Ebrahimi, "EdgeConnect: Generative image inpainting with adversarial edge learning", arXiv:1901.00212, 2019, [online] Available: <http://arxiv.org/abs/1901.00212>.
- [9] M. K. J. Khan, N. Ud Din, S. Bae and J. Yi, "Interactive removal of microphone objects in facial images", *Electronics*, vol. 8, no. 10, pp. 1115, Oct. 2019.
- [10] "Dataset", <https://www.kaggle.com/prasoonkottarathil/face-mask-lite-dataset>