International Journal of Engineering and Science Invention (IJESI)
ISSN (Online): 2319-6734, ISSN (Print): 2319-6726
www.ijesi.org ||Volume 13 Issue 5 May 2024 || PP 59-65

Deep Convolutional Neural Models for Deep Fake Image Prediction

[1]Chikati Madhava Rao, [2]Dr Santosh Kumar Yadav
[1]Research scholar, Research scholar, Shri Jagdishprasad Jhabarmal Tibrewala University, Vidyanagri, Jhunjhunu, Rajasthan.
[2]Professor & Director Research, Shri Jagdishprasad Jhabarmal Tibrewala University, Vidyanagri, Jhunjhunu, Rajasthan

Abstract:
As smartphones were introduced capturing images has become humans and the same became an integral part of daily lives. Humans get these images so, and their understanding of quality is an important area of research. Some image processing systems are effectively integrated into this perception of automation to measure the predicting the deep fake videos and image. Deep fake can be severely damaged by various distortions of blurriness, noise, compression. Image processing professionals can detect this immediately, but it is usually firmly in the system design human visual perception automation. Therefore, evaluating quality of image by a machine is an important field of research. Deep fake assessment algorithms are used to measure Deep fake and their goal is to estimate this property by measuring the human cortex. Now a days, deep learning has achieved great success in solving in various real-world problems, especially object detection and image classification. It has had some success in Deep fake assessment, but quality assessment without image proximity indicating human perception is still a dynamic section of research. The key problem is the unavailability of images with quality scores which are differential average feedback scores given by Deep fake experts in the 1-100 scale.
Keywords: *Deep fake video, image or video processing, classification and regression, deep neural networks.*

I. INTRODUCTION

Deep convolutional neural networks (CNNs) are crucial for the prediction and assessment of deepfake images. The layered structure of CNNs, along with their capacity to capture intricate patterns, makes them particularly effective for detecting deepfakes. This is crucial, as subtle facial manipulations and generated artefacts often evade conventional image analysis techniques. In the realm of deepfake image prediction, convolutional neural networks demonstrate exceptional performance due to their ability to learn hierarchical representations. As data traverses various layers, CNNs cultivate a sophisticated comprehension of the spatial and structural patterns characteristic of genuine versus deepfake images. For example, advanced layers in a CNN may identify discrepancies in lighting or atypical textures commonly associated with deepfakes.

The emergence of deepfake technology has led to social media platforms becoming prime targets for the dissemination of altered videos, frequently posing difficulties in confirming the authenticity of the content being shared. Deep convolutional neural networks (CNNs) serve as essential instruments for the identification and prediction of deepfake videos circulating on social media platforms. These models are capable of identifying nuanced discrepancies and distinctive artefacts present in deepfake videos, facilitating enhanced oversight and reaction to misinformation.

No-reference IQA (NR-IQA) proves to be especially advantageous for evaluating the quality of deepfakes, as it operates without the need for a high-quality reference image. This method is essential for assessing deepfakes, enabling the CNN to analyse a distorted image independently and evaluate its quality according to established patterns.

Significant research and funding from major organisations will ensure progress. Research is now focused on the evaluation of deep fakes. Image quality is normally contingent upon the viewers of the pictures; nevertheless, Deep Fake Assessment (IQA) aims to express it with individual perceptual knowledge [1]. Deepfake evaluation is often categorised into reference-free testing concerning visuals. The disparity between your assessments lies in the fact that NR-IQA uses a high-quality reference picture, while a distorted image will be derived from the NR-IQA. Typically, it is uncomplicated to evaluate the standard tag of an image when a benchmark is established based on the Structural Similarity Index. Generally, obtaining a no-reference picture involves aligning the image, calculating the samples, and generalising the scores to individual comments. The selection of an individual based on the form of the picture feature is a labour-intensive approach that should be avoided. When it pertains to the multitude of visuals, it is almost unfeasible. Additionally, one may use the capabilities of deep learning and pattern recognition to automate the maintenance of distortion, which might otherwise be a vexing task. Consequently, much research has been conducted, and several models have been proposed to evaluate picture quality using deep neural networks. Classification and regression paradigms were beneficial for its anticipated DNNs. This project aims to classify photos into four categories based on superior scores ranging from 0 to 100, since classification is mostly conducted in graphics. The dataset was obtained from the University of Texas at Austin's video and picture creation website and comprises 982 images with five types of distortions. The Deepfake scores may also be calculated using manually created boundary maps. The previously mentioned LIVE-IQA dataset was used. Furthermore, the TID 2008 and CSIQ datasets were used. They used 1,700 and 866 photos, respectively, using mostly the four distortions: JPEG, JPEG2000, and Gaussian Blur.

II. RELATED REVIEW

Searching for certain types of algorithms that can evaluate the quality of movies and pictures in relation to human visual perception will always be the primary goal of Deep Fake Assessment. In most cases, the similarity between the distorted picture and a prominent benchmark image is used to translate Deep fake in Deep Fake Assessment (IQA) computations [1]. Modelling the visual components so that they are all close to the optical equipment often yields the best high-quality prediction [1]. Using a number of criteria, including the Structural Similarity Index (SSIM) from the plasma and wavelet domain names, they determined the image's quality with and without reference. The forecasts were quite close to specific abstract comments [1]. To predict a distorted picture, Xiang et al. proposed a deep neural model that employed regression and a streamlined version of this AlexNet architecture with a few fully connected layers [2]. Using just two ideas, they distributed five different types of twisted pictures and included high ratings into these Deep false descriptions. Your bottom version might be the CNN version that they utilised on the ImageNet dataset [3]. We used JPEG photos and the Gaussian Blur filter to conduct the experiments. To circumvent this issue, a scale model of the sub-images from random pooling may be used. In the picture, contain maps are only pulled once. The picture is prepared to provide fixed-length vectors since arbitrary locations are selected. Because of the difficulty in enhancing the rough photographs according to their quality, this had to be adjusted. Talebi and Milanfar proposed neural image assessment (NIMA) as a means to train a deep convolutional neural network (CNN) to evaluate visually appealing images that are ideal for the average user [5]. The video picture is also used to provide insight into the scores shown in the network. Images are enhanced by editing and processing.

III. PROPOSED METHODOLOGY

The goal of deep learning is to allow robots to learn profoundly by simulating the complex network of neurones and connections seen in the human brain. This way of learning is really important for teaching models to do specific tasks. Complex, multilayered neural networks, which often have millions of parameters, can spot and understand patterns that would usually need human attention. The fundamental idea behind deep learning is to continuously improve the model's accuracy by modifying the "weights" in neural connections in response to enormous volumes of data. Deep learning models mimic how the human brain works by adjusting themselves over time. They use feedback processes that are kind of like how our brain's synapses operate to find the best solutions.

Achieving human-level accuracy in complicated tasks is still difficult and requires a great deal of training and improvement, even with enormous computer resources and data. Deep learning has emerged as one of the 21st century's most revolutionary technical developments, transforming domains such as computer vision as technology develops. Neural networks are made up of layers that manage different parts of data processing, depending on the application: input layers take in the data, hidden layers analyse it and identify patterns, and output layers provide the final classification or prediction. Deep learning networks may "learn" complicated representations because to their layered structure, which often goes beyond what is initially apparent to humans.

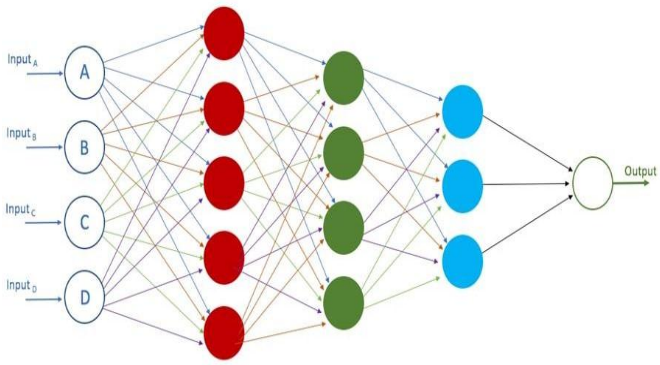

Figure 1. Deep Learning Neural Network Model [6]

Image Distortions

An algorithm that is essentially just a distortion in image processing will be used to reduce graphics. Inadequate information during data transfer over a network due to packet losses is another potential source of distortions. In most cases, the purpose of using straight lines in visuals is to distort them. Jpeg, Jpeg 2000, white noise, quick fading, and blur are just a few of the many picture compression techniques available. It integrates lossy and lossless compression techniques; hence, it employs the former while working with JPEG files. Along with picture reduction disruption, Gaussian blur is a movie character. The image depicts the dispersion of wireless networks during a period of fast evaporation.

Extensive false Computer vision researchers are still trying to figure out how to publish methods for predicting Deep fake with respect to human understanding, although prediction is already a big field of study. This project's anticipated work is to improve upon previous efforts, particularly in regression, and accept the best possible estimate due to categorisation and regression concerns [3]. One of the scoring periods is used to classify images according to the produced quality tags. By obtaining characteristics from visuals and putting them into DNNs, the Deep Fake Score is accessed during recession.

Some of the proposed models made advantage of top-tier architectures often associated with visual design and object identification, such as VggNet, ResNet, etc. [1, 2]. Certified in the use of high-definition visuals. It becomes more difficult to generalise a neural network for so many different types of distortions when we give it some serious thought. More than that, the item detection models have the characteristics seen in high-quality visuals. Deep fake prediction uses significantly distorted characteristics, which makes it such that accurate models don't work economically. In addition to the twisted aspects of the graphics, display maps may also provide guidance that isn't always useful for predicting the quality of images [9].

Proposed architecture for classification

Even the architectures utilized in most of experiments were motivated by the AlexNet [2] architectures as well as in-general involved 23 filtered convolution layers [12] accompanied with the pooling layers and also a pair of 23 repetitions and also the fully attached layers afterwards adjusting the input into some specified contour. The overall structure used is shown at Figure 3. This figure indicates the normal structure employed in this work [14].

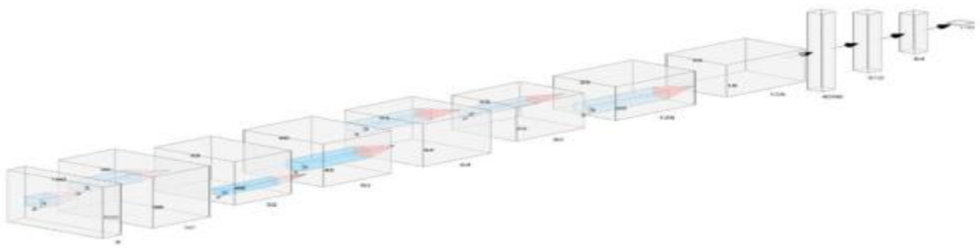

Figure 2. Proposed Deep CNN Architecture

IV. Methods and Materials

Regression on most of the photos from the three datasets (Live-IQA, CSIQ, and TID2008) and the associated distortions (JPEG, JPEG2000, White Noise, and Gaussian Blur) yields impressive results in the proposed research. Convolution neural networks have often undergone an iterative process of stripping after a rigorous learning process. [11]. According to the referenced source, "convolutional neural networks (CNNs) learn to detect patterns in images and objects in a systematic way" (14). This form of learning is highly effective when there is access to a substantial amount of data. For CNN to effectively extract features during the learning process, it necessitates large datasets, which are supplied by both object detection and image generation tasks. Deep neural networks face challenges in identifying patterns in issues such as Deep Fake Forecast due to the scarcity of images with standardised ratings.

This assignment proposes a novel framework for generating advantage maps and extracting benefit characteristics from graphics utilising the Sobel filter. The advantage channels are established by assessing the gradient magnitude and approximating it for each pixel in the image. Sobel kernels are employed to accurately replicate the image when calculating the gradient. Because the width and height of the image will change depending on whether it is a landscape or portrait, the benefit maps are shown separately for each kind of image. The border map has been removed from the left half of the landscape image, followed by the right half [16]. The upper section employs the border maps, whereas the lower section utilises them for portraits. Separating the border channels even more is a data collection that is just two to eight times larger than the first one. The input is amplified by a factor of 24. Subsequently, two convolutional neural networks with 16 filters each are utilised for the benefit channels, alongside the application of conventional metering. Predicting the Deep Fake score can be achieved through non-core regression, utilising spatial pyramid pooling to accommodate images of varying sizes. The results varied from zero to one. Historically, obsolete models included plasma volcano pooling.

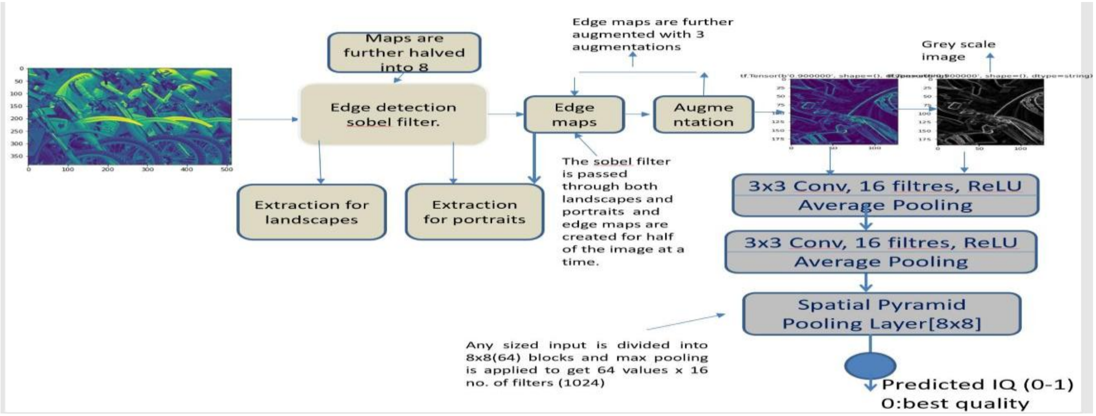

Figure 3. Regression Model Architecture

With 80% of the data used for training and 20% for testing, and with 40 epochs of good training, the version is ready to go. When contrasted with the OG-IQA [1] model and the bottom research-paper [3], the proposed version demonstrates much improved performance. Further, when compared with the CURVLET-2014, [4] model and the m 1 [14] version, it performs well. A deep fake assessment without a reference, like BRISQUE or DEEP QA, is carried out by the version out.

V. EXPERIMENTAL EVALUATION AND COMPARISON

Interpretations of Classification Models

Proposed model was subjected to rigorous testing with several tweaks. A kernel size of 3x3 was used for many trials performed. The trials included categorising the graphics into three unique classifications: 2x up-scaled pictures, 4x up-scaled images, and complete graphics. One variable, namely upscaling, denotes that when a picture is enlarged by a factor of 2, the highest quality designation is attributed to the 2x upscaled image, and correspondingly for a 4x upscale with a factor of 2.

Table 2. Proposed model variations

Model Name	Model Details
CNN_Full_img	2 stacked 16 filtered convolution layers with spatial pyramid pooling and 4 fully connected layers
CNN_2x	2 stacked 32 filtered convolution layers with spatial pyramid pooling and 4 fully connected layers
CNN_4x	3 stacked 16 filtered convolution layers with spatial pyramid pooling and 4 fully connected layers.

Full image model of CNN

Full graphics with three piles of sixteen filters and a modified adrenal pyramid pooling layer ensure that reels of varying sizes are treated with the same level of accuracy and activation in this design. When corrected for 4.2486, the instrument's accuracy drops to 79.45% and its management error to 87%.

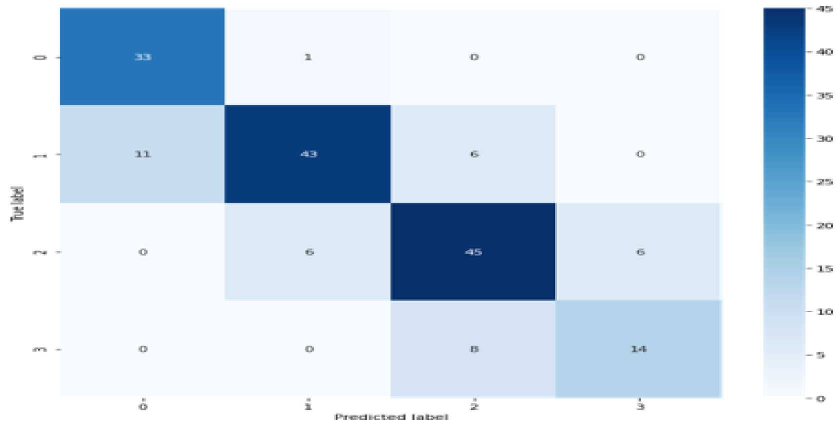

Figure 4: CNN_Full_img Confusion Matrix Model

The results are displayed in relation to the confusion matrix. Weighted accuracy, recall and F1 scores were 0.785 and 0,780 and 0,778 respectively. The above figure shows the confusion matrix which represents the true positives, true negatives, false positives, and false negative values for each label 0,1,2,3.

CNN_1_Mvote

The design cited in Table 4 has been used to extract more features by having a growing quantity of filters. Since most of the distortions are obtained simultaneously, a significant selection of features is going to soon be pulled. Drop-out has been inserted to lose a number of those features from the feature matrix. The evaluation set was subsequently split into spots, and each patch tag was called, along with also a vast majority vote is extracted from all of the patch forecasts to acquire the superior tag to get a graphic. A precision of 82.4percent has been obtained. The weighted accuracy, recall, and F1 scores to this particular version were all 0.81182, 0.8292, and 0.8195.

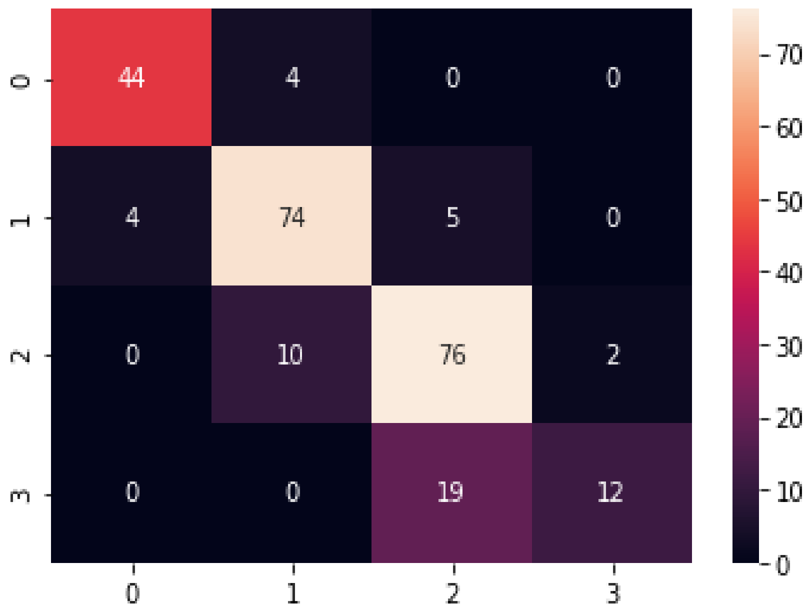

Figure 5: Confusion Matrix of CNN_1_Mvote

Figure 5 shows the confusion matrix results for CNN_Mvote_1. The results provide: 1) the number of correct and wrong predictions. 2) The total number of predictions.

Dataset Results Comparisons

The operation comparisons have been in relation to SROCC, KROCC, also PLCC.

Table 3. Comparison table of SROCC, KROCC, and PLCC for the LIVE IQA

	Model Name	BLUR	JPEG	WN	JP2K	ALL
Spearman Rank Order Correlation Coefficient (SROCC)	Deep QA	0.950	0.940	0.890	0.958	0.940
	BLISS-S	0.869	0.922	0.779	0.919	0.898
	Proposed	**0.972**	**0.979**	**0.944**	**0.984**	**0.947**
Kendall Tau Rank Order Correlation Coefficient (KROCC)	CURVLET-2014	0.665	0.635	0.696	0.574	-
	M1	0.659	0.633	0.728	0.646	-
	Proposed	**0.896**	**0.833**	**0.797**	**0.868**	**0.797**
Pearson Linear Correlation Coefficient (PLCC)	BLISS-S	0.875	0.955	0.748	0.945	0.910
	SSIM	0.954	0.964	0.816	0.971	0.902
	Proposed	**0.977**	**0.989**	**0.947**	**0.984**	**0.947**
	Proposed	**0.977**	**0.989**	**0.947**	**0.984**	**0.947**

Table 3 shows the The proposed deepfake video detection model surpasses existing IQA models in regression performance, achieving higher SROCC, KROCC, and PLCC metric scores on distortions from the TID 2008 database.

VI. CONCLUSIONS AND FUTURE WORK

Deep fake technology's future is still very much out in the social media. As a multi-class classification issue, graphic quality prediction is the focus of the units and processes suggested in this study. Furthermore, it will aid in assessing whether a given picture, based on its quality, is acceptable for your image categorisation challenge. This version makes changes to the quality labels and looks at ways to improve these graphics in order to build a system that can anticipate category labels with a high degree of accuracy and use them to boost the system's performance. The majority voting technique is also used, which is different from prior photo prediction contests [4]. The evaluation set is checked for picture stains, and the accuracy is determined by comparing the stains to the whole image, which is usually the main task. Neural networks can be trained with large amounts of data, which might lead to better outcomes. It may be difficult to pinpoint specific issues while collecting data on cognitive quality, and datasets may be structured to provide personalised image distortions for better training patterns [17]. On the other hand, this is just a hectic and tedious task. Businesses assessing Deepfake might also benefit from collecting several examples from random internet users via adverts [16]. This opens the door to better outcomes from analyses of a broader variety of datasets.

REFERENCES

[1]. A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference Deep fake assessment in the spatial domain," IEEE Transactions on Image Processing (TIP), vol. 21, no. 12, pp. 4695–4708, December 2012.

[2]. Yang, H., Cho, K. C., Kim, J. J., Kim, J. H., Kim, Y. B., & Oh, J. H. (2023). Rupture risk prediction of cerebral aneurysms using a novel convolutional neural network-based deep learning model. Journal of NeuroInterventional Surgery, 15(2), 200-204.

[3]. Zhang, Z., Luo, P., Loy, C. C., & Tang, X. (2015). Learning social relation traits from face images. In Proceedings IEEE International Conference on Computer Vision (pp. 3631-3639).

[4]. K. He, X. Zhang, S. Ren, and J. Sun, "Spatial Pyramid Pooling in Deep Convolution Networks for Visual Recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), vol. 37, no. 1, pp. 1904-1916, June 2014.

[5]. H. Talebi and P. Milanfar, "NIMA:Neural Image Assessment," IEEE Transactions on Image Processing (TIP), vol. 27, no. 1, pp. 3998-4011, June 2018.

[6]. Agarwal S, Farid H, El-Gaaly T, Lim SN. Detecting deep-fake videos from appearance and behavior. 2023 IEEE International Workshop on Information Forensics and Security, WIFS 2020. 2020.

[7]. Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. European Journal of Communication, 33(2), 122–139. https://doi.org/10.1177/0267323118760317

[8]. K. Le, P. Ye, Y. Li, and D. S. Doermann. "Convolutional Neural Networks for No- Reference Deep fake Assessment," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (TIPAM), Columbus, Ohio, pp. 1733- 1740 April 2014.

[9]. B. Sebastian, D. Maniry, T. Wiegand and W. Samek. "A deep neural network for Deep fake assessment.", in Proceedings of IEEE International Conference on Image Processing (ICIP), Buffalo, New York, pp. 3773-3777, September 2016.

[10]. W. Xiaochuan, K. Wang, B. Yang, W. B. Frederick and L. Xiao-Hui. "Deep Blind Synthesized Deep fake Assessment with Contextual Multi-Level Feature Pooling." In Proceedings of IEEE International Conference on Image Processing (ICIP), Piscataway, New Jersey, pp. 435-439, May 2019.

[11]. X. Yang, F. Li, and H. Liu. "A Comparative Study of DNN-Based Models for Blind Deep fake Prediction," in Proceedings of IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, pp. 1019-1023, September 2019.

[12]. Prasadi Peddi and Dr. Akash Saxena (2015), The Adoption of a Big Data and Extensive Multi-Labled Gradient Boosting System for Student Activity Analysis, International Journal of All Research Education and Scientific Methods (IJARESM), ISSN: 2455-6211, Volume 3, Issue 7, pp:68-73.

[13]. L. Kong, A. Ikusan, R. Dai, and J. Zhu. "Blind Deep fake Prediction for Object Detection," in Proceedings of IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), San Jose, California, pp. 216-221, November 2019.

[14]. N. Jain and P. Peddi, Gender Classification Model based on the Resnet 152 Architecture, 2023 IEEE International Carnahan Conference on Security Technology (ICCST), Pune, India, 2023, pp. 1-7, doi: 10.1109/ICCST59048.2023.10474266.

[15]. W. Szegedy, Y. Liu. "Going deeper with convolutions. In: Computer Vision and Pattern Recognition," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, Massachusetts, pp. 1-9, June 2015.

[16]. Sun, Y., Wang, X., & Tang, X. (2013). Deep convolutional network cascade for facial point detection. In Proceedings IEEE conference on computer vision and pattern recognition (pp. 3476-3483).