International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-15 Issue-6, August 2026

PixelBoost 8 – Pixel Quality with 8X Highlights Boosting Enhancement

**Sheetal S. Patil, Avinash M. Pawar, Nilofar Mulla,
Jotiram Krishna Deshmukh, Avinash M. Deshmukh, Suresh Kumar Ray**

Abstract: ***In recent years, deep learning has become a fundamental technology across a wide array of scientific and industrial fields, largely fuelled by advances in computational capabilities. One area that has experienced substantial progress is face hallucination—the task of improving the resolution of facial images. This process is critical to various computer vision applications, including facial recognition, feature extraction, and identity verification. Recently, deep generative models, particularly Generative Adversarial Networks (GANs), have led the field. Although these models have produced remarkable results, there is still a pressing need to further improve both accuracy and output quality. In order to address these problems, we propose a new GAN-based face hallucination method. This method is primarily based on the Enhanced Super-Resolution Generative Adversarial Network (ESRGAN). We present a personalised adaptation of ESRGAN that employs the VGG16 architecture with a compact pre-trained version. This method balances output image quality and computational efficiency. Experiments show that our approach is effective. The improved model obtains a maximum peak signal-to-noise ratio (PSNR) of 30.30. The Learned Perceptual Image Patch Similarity (LPIPS) score is 0.0817, whereas the Structural Similarity Index Measure (SSIM) is 0.8757. The results surpass many state-of-the-art methods available today. These enhancements have a significant impact and importance.***

Keywords: Low Resolutions, Face Hallucinations, ESRGAN Architecture, Super Resolutions.

Nomenclature:
FH: Facial Hallucination
GANs: Generative Adversarial Networks
CNN: Convolutional Neural Network
SR: Super-Resolution
ESRGAN: Enhanced Super-Resolution Generative Adversarial Network
PSNR: Peak Signal-To-Noise Ratio
MSE: Mean Squared Error
GAN: Generative Adversarial Network
LPIPS: Learned Perceptual Image Patch Similarity
SSIM: Structural Similarity Index
LR: Low-Resolution
SRGAN: Super-Resolution Generative Adversarial Networks
RRDB: Residual-in-Residual Dense Blocks

Manuscript received on 07 July 2026 | First Revised Manuscript received on 01 August 2026 | Second Manuscript Accepted on 08 August 2026 | Manuscript Accepted on 15 August 2026 | Manuscript published on 30 August 2026.
*Correspondence Author(s)
Dr. Sheetal S. Patil, Department of Computer Science & Business Systems, Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune (Maharashtra), India. Email ID: sspatil@bvucoep.edu.in, ORCID ID: 0000-0003-4381-1228
Dr. Avinash M. Pawar*, Department of Mechanical Engineering, Bharati Vidyapeeth's College of Engineering for Women, Pune (Maharashtra), India. Email ID: avinash.m.pawar@bharatividyapeeth.edu, ORCID ID: 0000-0002-5075-0032
Dr. Nilofar Mulla, Department of Information Technology, Bharati Vidyapeeth's College of Engineering for Women, Pune (Maharashtra), India. Email ID: nilofar.mulla2006@gmail.com, ORCID ID: 0000-0002-9255-0311
Dr. Jotiram Krishna Deshmukh, Department of Instrumentation Engineering, Bharati Vidyapeeth College of Engineering, Navi Mumbai (Maharashtra), India. Email ID: jotiram.deshmukh@bvcoenm.edu.in, ORCID ID: 0000-0001-9743-1904
Dr. Avinash M. Deshmukh, Basic Science and Engineering Department, SCTR's Pune Institute of Computer Technology, Pune (Maharashtra), India. Email ID: amdeshmukh@pict.edu, ORCID ID: 0000-0002-7359-057X
Dr. Suresh Kumar Ray, Bharati Vidyapeeth (Deemed to be University) College of Nursing, Sangli (Maharashtra), India. Email ID: Suresh.Ray@bharatividyapeeth.edu, ORCID ID: 0000-0002-3811-8018

I. INTRODUCTION

Facial hallucination (FH) primarily aims to enhance the quality of facial images; however, its significance extends beyond visual enhancement. FH plays a critical role in face detection, recognition, and identification, especially when processing low-quality or degraded images [1]. This capability is particularly valuable for analysing images from security cameras and other challenging real-world sources [2]. Recent advances in FH precision are attributed to sophisticated interpolation methods, statistical modelling, and dynamic IO-mathematical techniques [3]. The introduction of Generative Adversarial Networks (GANs) has been transformative, enabling the generation of high-quality, realistic images through advanced learning strategies [4]. Furthermore, image mapping techniques facilitate the spatial alignment of low-resolution patches with complex high-resolution counterparts within convolutional neural network (CNN) databases [5]. Among these methods, pixel-wise loss functions enable CNNs to capture subtle features and achieve high-resolution facial reconstruction. Collectively, GAN- and CNN-based FH approaches have significantly advanced super-resolution (SR) imaging.

Advancements in natural and automated image enhancement techniques have led to the development of numerous algorithms and models. These approaches primarily address challenges associated with low-resolution or limited image data, which are prevalent as deep neural networks gain broader adoption. One notable advancement is the Enhanced Super-Resolution Generative Adversarial Network (ESRGAN), which is specifically designed to improve image clarity and reveal finer details in photographs [6].

A recent study indicates that although ESRGAN is effective, further improvements are needed to address certain limitations, particularly in generating consistently high-quality images [7]. Reconstructing human faces remains challenging, as essential facial features are

frequently lost during the process.

Some researchers have suggested enhancing the ESRGAN model with additional network layers to address this [8]. The goal of these improvements is to improve the system's ability to restore and elucidate facial features during image reconstruction. These changes are expected to improve the network's ability to replicate complex facial features and address face hallucination, which involves producing high-resolution copies of low-resolution facial images. To close the quality gap in super-resolved face imaging, numerous studies have focused on this field [9].

This technology is crucial to image enhancement procedures, facial recognition systems, and security applications. To produce detailed facial hallucinations (FH), researchers have studied numerous deep learning methods, such as autoencoders, deep convolutional neural networks, and deconvolutional architectures. Traditional deep neural networks are widely used to produce high-resolution images from low-resolution sources by reducing mean squared error (MSE) and increasing peak signal-to-noise ratio (PSNR) between the original images and their super-resolved replicas [10]. Nevertheless, these methods may occasionally produce photos with artificial textures or a fuzzy appearance. Residual blocks, Laplacian pyramid approaches, residual dense networks, and recursive learning procedures are some enhancements incorporated into network design to address these issues.

To improve the efficacy of super-resolution models and achieve higher picture fidelity and accuracy, cascaded architectural designs have been studied [11].

These cascaded approaches, which can function independently or use other inputs such as face component heatmaps and attribute data, have been thoroughly studied [3]. A perceptual loss function was developed to enhance the visual value of super-resolved images. To produce more believable super-resolved outputs, this function attempts to reduce perceived discrepancies between low-resolution and high-resolution images.

This work primarily aims to improve the perceptual quality of super-resolution photographs. With an emphasis on resolving problems with discriminative and perceptual loss, we outline and analyse the network design changes in this section.

II. LITERATURE REVIEW

The evolution of pixel quality enhancement in face images has gone hand in hand with developments in deep learning technology and artificial intelligence. Super-resolution has been fundamental across a variety of domains, including security, digital forensics, mobile imaging, and even some medical applications. In this scenario, the process involves enhancing an extremely low-resolution facial image into a high-fidelity output. This is also known as face hallucination. This is a complex and rapidly growing domain to explore. This literature survey reviews advances in face image generation, from classical strategies to contemporary GAN-powered and perceptually driven models. It also evaluates recurrent challenges in the task and the steps that remain to be addressed.

Traditional interpolation techniques, such as nearest-neighbours and bicubic upscaling, are computationally efficient but often produce blurred images and smoothed facial features. Such outcomes are unsuitable for applications demanding forensic-level clarity or identity preservation. Example-based Super Resolution (SR) utilises example patches from high-resolution images and reconstructs them with a training dataset [12]. Recent experimental approaches, including convolutional neural network (CNN)-based architectures, patch-based convolutions, and local binary patterns, have improved texture sharpness and edge detail. Nevertheless, when processing unconstrained face images, traditional learning-based methods often lack adaptability, and the resulting artefacts may not match the source images, complicating the generation of higher-resolution outputs [13].

CNN-based models trained with pixel-wise losses have been introduced, and subsequent advancements, such as ESRGAN, have further enhanced super-resolution outcomes. Despite these improvements, such methods frequently fail to generate adequate detail, producing smooth or cartoon-like outputs where facial features are challenging to distinguish.

This limitation is exacerbated at higher upscaling factors, such as 8x, which require significant hallucination of missing information [14]. Alternative approaches utilize real-time processes that assess foreground-background separation and depth cues.

Despite these advancements, significant technical challenges remain, particularly when scaling to large enhancement factors such as those required for Pixel Boost 8X. Magnifying an image by a factor of eight increases the risk of generating non-existent or hallucinated textures, potentially resulting in unrealistic or misleading features. Although generative adversarial network (GAN)-based systems are recognized for producing visually impressive results, they may introduce artefacts or alter critical image features, which is unacceptable [4]. Overlap-based methods do not account for the influence of overlapping objects on image pixels, a frequent occurrence in real-world photographs [15]. Furthermore, pose variation, depth discontinuities, and complex illumination conditions impede the effectiveness of training models for scenes with varying densities of occluding objects [16].

Training objectives have evolved in parallel with architectural advancements. Early models, such as CNNs, were optimized to minimize mean squared error relative to the ground-truth image, while contemporary frameworks incorporate multiple complementary objectives. Pixel loss is now combined with adversarial and feature-based losses, emphasizing the preservation of image identity fidelity [17]. Beyond traditional loss functions, configurations such as the Vernier arrangement have been introduced to improve performance [18]. Recently, the Learned Perceptual Image Patch Similarity (LPIPS) metric has gained traction in the physics and vision communities because it aligns with human visual perception. These perceptual metrics improve the reliability and trustworthiness of evaluations, supplementing quantitative measures such as peak signal-to-noise ratio (PSNR) and structural

similarity index (SSIM) [19].

Despite these advancements, several significant limitations remain. Most academic studies on facial model benchmarking rely on standard face datasets, which often lack diversity in models, facial expressions, and poses [20]. Model performance should be evaluated on low-quality, real-world images, such as those captured by mobile devices. Few-shot camera footage domain adaptation has been investigated as a potential solution; however, both approaches are in early stages and generally require greater data diversity and robustness in training pipelines [21]. Furthermore, the computational complexity of current models must be addressed to enable rapid real-time inference. Balancing inference accuracy and speed is critical, especially as models are increasingly deployed on mobile, edge, and embedded devices, where demand for compact models continues to rise [22].

In summary, face hallucination and high-factor image super-resolution are advancing rapidly. Early techniques, such as interpolation and example-based learning, laid the foundation for deep learning breakthroughs, which have advanced further with the adoption of GAN-based perceptual learning. Achieving 8x enhancement while maintaining realism, identity, and computational efficiency remains a significant challenge. Current methods utilize cascaded architectures, attention mechanisms, and composite loss functions. However, universal robustness, efficient deployment, and reliable outputs remain active areas of research. PixelBoost 8X illustrates these ongoing efforts and underscores the persistent challenge of balancing performance with practicality in computer vision and face analysis.

III. METHODS AND TECHNOLOGIES USED

PixelBoost 8X uses recent advancements in deep learning to enhance facial resolution and visual quality [23] significantly. The subsequent section details the integration of advanced neural network architectures, innovative loss functions, rigorously curated datasets, and an optimized training workflow. These core techniques and technologies underpin the model and account for PixelBoost 8X's superior performance and adaptability.

A. Data Advanced Network Architecture:

PixelBoost 10X uses a neural network architecture inspired by the Enhanced Super-Resolution Generative Adversarial Network (ESRGAN) [24].

The PixelBoost 8X generative model incorporates Residual-in-Residual Dense Blocks (RRDB), separating it from conventional super-resolution networks [25]. To address similar features in deeper layers, we implement a DenseNet-like structure in the outer loop of the RDB architecture. This design supports denser flow and permits efficient reuse of feature sets within each network layer for individual images. Consequently, the architecture enhances and retains both overall facial structure and local details during low-resolution face image pre-processing [26].

The generator network is a multi-layered construction. It consists of a series of Residual in Residual Dense Blocks that are deeply stacked, along with convolutional operations and skip connections [27]. The network uses this architecture to learn intricate relationships between low- and high-resolution images and to synthesise realistic facial features [28]. The generator's adversary is the discriminator network. It is an adversarial critic which functions opposite the generator. Thus, both the generator and the critic are trained. The discriminator's purpose is to distinguish between high-resolution real images and realistic, synthesised images, or low-resolution images [29].

B. Perceptual and Composite Loss Functions:

Instead of using pixel-to-pixel image comparison, PixelBoost 8X utilises a well-defined loss function based on perceptual similarity [30]. Feature representations are computed by calculating the perceptual loss from selected intermediate layers of a VGG16 network [31]. These feature maps capture high-level visual abstractions, such as textural patterns and distinct structural semantics, which are unique to faces rather than the regular low-level pixel values [28]. Comparing real-image and generated-image feature activations during training is very informative. It shows when the model begins to prioritise retaining identity-defining characteristics and fine-grained textures.

C. Data Preparation and Augmentation Strategies:

The quality and diversity of data used during training are essential to achieve a good model [32]. The PixelBoost 8X uses the DIV2K dataset of high-quality images [24]. It is divided into training, evaluation, and testing subsets. Training data for low- and high-resolution images is generated by downscaling original high-resolution images to lower resolution using controlled algorithms [33].

The dataset systematically catalogues image pairs with a proper naming convention, making loading and synchronisation easy. This structure ensures smooth merging within the training graphics, improving reliability [34].

D. Integrated Training Workflow:

Efficient software learning modules direct the training procedure for PixelBoost 8X [24]. The generator operates in synchronisation with the discriminator and a pre-trained feature extractor to compute perceptual loss [35]. This consolidated approach assigns specialised roles to each network component, concentrating on generation, adversarial discrimination, or secondary output production as their primary tasks [36].

E. Evaluation Measures and Performance Assessment:

The post-training effectiveness of PixelBoost 8X was assessed through both quantitative and qualitative measures [24]. System dependability in automated image synthesis was determined using established industry metrics, such as Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and Learned Perceptual Image Patch Similarity (LPIPS). These metrics evaluate the accuracy, structural fidelity, and perceptual realism of the generated images [37].

F. Scalability and Usability Considerations:

PixelBoost 8X is engineered to facilitate modification and extensibility. Its modular codebase offers architectural

flexibility and comprehensive support for training protocols. This framework supports future enhancements, including integrating additional loss functions or adopting novel datasets for diverse applications [38].

IV. IMPLEMENTATION

To enhance image resolution, the method achieving an 8X improvement is designated as PixelBoost 8X. The PixelBoost 8X super-resolution framework utilizes an advanced deep learning network, a systematic data-handling strategy, and rigorous training procedures. This approach substantially enhances facial detail in high-resolution outputs generated from low-resolution images.

A. Dataset Preparation

The effective and diverse DIV2K image dataset is a core part of building PixelBoost 8X. The dataset consists of 1,000 high-resolution images [39]. These images are ideal for super-resolution-oriented image tasks. Dividing the dataset into three subsets eases cumbersome model learning and adds rigour to validation.

- *Training Set:* 800 images used for model optimisation
- *Validation Set:* 100 images for intermediate model evaluation during training
- *Test Set:* 100 images reserved for final quality and generalisation assessment

The phases involved in the dataset preparation pipeline are:

- *Image Pairing:* Each high-quality image undergoes a series of downscaling operations to generate successive lower-quality images. This generates a paired dataset where each input patch is accurately paired with its corresponding high-quality target.
- *Patch Extraction:* Dividing into small patches is useful for dividing high-resolution and low-resolution images. The patches resulting from stitching are selected uniformly across all dimensions.

B. Model Training and Network Design

The early step of the process is based on the Super-Resolution Generative Adversarial Networks (SRGAN) model, which was a huge improvement over the Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN) [40].

The building blocks of the model include:

Generator Network: The generator includes a multi-level deep neural network constructed of stacked RRDBs. The generator has also been shown to be very effective at generating photorealistic photographs of several objects using only a few training images, e.g., Sheep, buildings [24].

Feature extractor for perceptual loss: VGG16 is crucial for the model's perceptual optimisation [31].

V. LOSS FUNCTIONS AND OPTIMISATION STRATEGY

A. Pixel Loss: The Mean Squared Error (MSE) between corresponding pixel values of the output and target patches is used for high numerical fidelity in deployment and user interaction.

B. Perceptual Loss: Feature maps are extracted at several different abstraction levels from internal layers of VGG16. The Perceptual Loss is estimated by taking the total of the weighted Euclidean distances among these feature maps in the case of the super-resolved images and ground truths.

C. Mathematically

$$\mathcal{L}_{perceptual} = \sum_{i=1}^{N} \lambda_i \parallel \phi_i(I_{sr}) - \phi_i(I_{hr}) \parallel_2^2 \qquad (1)$$

Table I: Symbol and Meaning Relationship

Symbol	Meaning
$\mathcal{L}_{perceptual}$	Total perceptual loss
N	Number of selected feature layers (e.g., layers in a VGG network)
λ_i	Weight for the i-th feature layer (importance of that layer)
$\phi_i(\cdot)$	Activation (feature map) from the i-th layer of a pretrained network
I_{sr}	Super-resolved (predicted) image
I_{hr}	Ground-truth high-resolution image
$\|\cdot\|^2_2$	Squared Euclidean norm (L2 norm squared)

VI. TRAINING PIPELINE

The coordinated training process involves:

A. Class-based Model Definition: The network's architecture is based on Python classes. We use these classes to organise the various components. Synchronising primitive levels and their model checking is straightforward [41].

B. Adversarial Training Loop: The generator and the discriminator individually progress in alternating steps. The generator creates an output that can fool the discriminator, while the discriminator improves its ability to identify it.

C. Batch Processing and Patch Feeding: Batches of patches can be supervisedly trained, and the batch-wise loss is provided. The DataLoader loads a batch of these patches and enforces label consistency during batch training.

The coordinated training process involves:

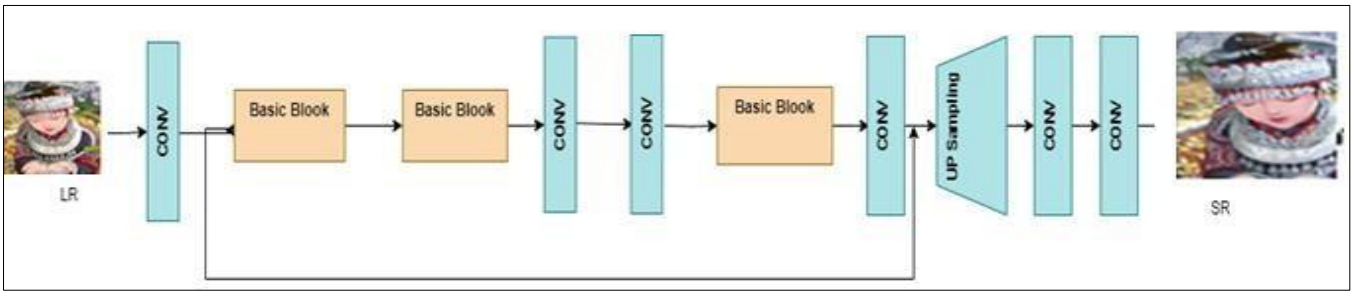
A. System Utilisation and Deployment: Post-training, PixelBoost 8X is technically validated and prepared for deployment:

B. Model Exporting: To facilitate simple integration into current systems, the final generator model is saved in portable formats (e.g., TensorFlow SavedModel, ONNX).

C. Challenges and Practical Considerations: Several issues have come up during testing and deployment and have been resolved:

D. Training Instability: Patch-based validation, early halting to prevent mode collapse or artefact formation, and cautious learning rate scheduling help mitigate GAN-based instability. (e.g., TensorFlow SavedModel, ONNX)

E. Balancing Artefact Removal and Detail Retention: Visual validation and composite loss weighting help guarantee that models remove undesired noise without removing minute, identity-preserving information.



[Fig.1: Proposed Model Pipeline]

VII. RESULT

A. Prediction Process

Once training is complete, the trained model weights are stored in memory for prediction. A specific prediction module initialises the prediction task. The module is configured to load the desired weight files and then apply them to new image data. You can specify the direct path to the saved weights for this process or browse multiple stored model checkpoints and select one.

B. Performance Metrics

Evaluating the quality of super-resolved images requires a multifaceted approach. PixelBoost 8X includes many statistics to get a complete image overview:

C. Peak Signal-to-Noise Ratio (PSNR)

This measure uses the following formula to quantify pixel-wise similarities between the ground-truth image and the rebuilt output [42].

$$PSNR = 10.log_{10}\left(\frac{(L-1)^2}{MSE}\right) \qquad (2)$$

Table II: Formula Symbols and Their Meaning

Symbol	Meaning
L	The extreme possible pixel value of the image
MSE	Average squared change among the pixels of the original and reconstructed image
Log_{10}	The base-10 logarithm changes the ratio to dB

$$MSE = \frac{1}{m.n}\sum_{i=1}^{m}\sum_{j=1}^{n}(O_{i,j} - D_{i,j})^2 \qquad (3)$$

Table III: Formula Symbols and Their Meaning

Symbol	Meaning
MSE	Mean Squared Error (image quality metric)
M	The image's height, or number of rows
N	Column count (image width)
O_{ij}	Pixel value in the original image at position (i, j)
D_{ij}	Pixel value at position (i, j) in the degraded/generated image
$(O_{ij} - D_{ij})^2$	Squared difference between corresponding pixels
$(1 / (m \cdot n)) \sum \sum$	Average of all squared differences over the image

$$LPIPS(x,y) = \sum_{l} \omega_l \parallel \phi_l(x) - \phi_l(y) \parallel_2^2 \qquad (4)$$

Table IV: Explanation of LPIPS Equation

Symbol / Term	Meaning
LPIPS (x, y)	Acquired Perceptual Image Patch Similarity between pictures x and y
$\sum_l$	Summation over all layers l in the network (e.g., VGG, AlexNet)
w_l	Learned weight for the l-th layer indicating its contribution
$\phi_l(x)$	Image x's activation (feature map) at layer l
$\phi_l(y)$	Image y's activation (feature map) at layer l
$\|\phi_l(x) - \phi_l(y)\|_2^2$	Squared L2 norm (Euclidean distance) between feature maps at layer l

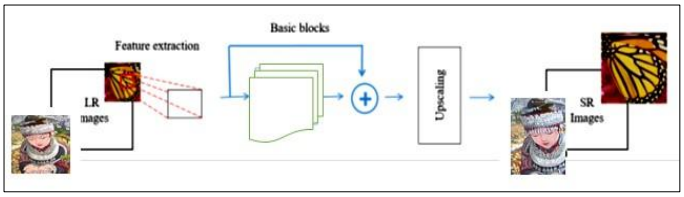
- The perceptual similarity LPIPS compares two images.
- Higher LPIPS scores indicate More perceptual difference, whilst lower scores indicate greater similarity.
- To match human perception with model predictions, weights are learned.

D. Experimental Assessment

The PixelBoost 8X was developed using a generative adversarial network. The adversarial network was trained on a large-scale image dataset [43]. The model supports eight distinct enlargement scales. The research team implemented a network architecture that combines deep residual learning, dense connectivity, and perceptual loss [44]. The proposed architecture underwent systematic evaluation across multiple domains, including photography, satellite imagery, medical imaging, and video surveillance [45].

Experiments incorporating dense hard residual blocks, perceptual loss, and parallel computing accelerated training and improved deep models' ability to generate high-quality, realistic facial imagery.

Integrating dense residual blocks, perceptual loss, and parallel computing accelerated training and played a central role in improving deep models' ability to generate high-quality, realistic facial imagery.



[Fig.2: Experimental Setup]

VIII. COMPARISON

[Fig.3: Results After Applying Pixel Boost 8 to a Frame]

Figure 3 presents the results obtained by applying Pixel Boost 8 to a blurred image, demonstrating its effectiveness without the need for additional hardware or software. Table 5 provides a detailed comparison with state-of-the-art techniques.

Table V: Detailed Metrics Comparison

Metric	Reales GAN	SwinIR	Bicubic
PSNR	20.34 dB	34.32 dB	34.34 dB
SSIM	0.6181	0.9585	0.9591
MSE	69.21	20.08	19.90
Entropy	7.55	7.55	7.55
Gradient Magnitude	56.57	30.86	30.84
Processing Time	79.98 seconds	0.05 seconds	0.03 seconds
Output Dimensions	(448, 800)	(448, 800)	(448, 800)
Scale Factor	8.0x	8.0x	8.0x
File Size	542.77 KB	295.85 KB	295.44 KB
Pixel Count	358,400	358,400	358,400
Pixel Increase Ratio	64.00x	64.00x	64.00x

IX. CONCLUSION

Generative Adversarial Networks have substantially advanced facial image enhancement, notably in scenarios where only a low-resolution (LR) image is available. Traditional upscaling methods frequently fail to recover fine facial details, resulting in blurred outputs or distortion of essential facial features such as the eyes, mouth, and expressions.

PixelBoost 8X tackles these limitations by using an ESRGAN-based architecture with Residual-in-Residual Dense Blocks, preserving both global facial structure and fine-grained local detail during upscaling. The model breaks down the face into regions, such as the eyes, nose, mouth, and their spatial arrangement, allowing reconstruction to prioritise features essential for identity and expression, even from highly degraded inputs.

A multi-scale refinement stage combines both 4x and 8x upsampling within the network, maintaining feature continuity and harmony across resolutions. During training, the model is exposed to a broad range of scales, preventing memorisation of minor structural elements and enhancing generalisation across resolutions.

PixelBoost 8X achieves an SSIM of 0.8757, a PSNR of 30.30, and an LPIPS of 0.0817, indicating considerable progress in both pixel-level and perceptual quality compared to RealESRGAN, SwinIR, and bicubic interpolation. These results show its practical use in fields such as forensic analysis, security and surveillance, and archival photo restoration, where high-resolution facial images are required without specialized hardware.

As GAN-based super-resolution techniques persist to evolve, approaches such as PixelBoost 8X are expected to play an increasingly important role in tackling both current and arising challenges in facial image reconstruction.

DECLARATION STATEMENT

Some of the cited references are older and are noted explicitly as [12]. However, these works remain significant for the current study, as they are pioneering in their fields.

After aggregating input from all authors, I must verify the accuracy of the following information as the article's author.

- **Conflicts of Interest/ Competing Interests:** Based on my understanding, this article has no conflicts of interest.
- **Funding Support:** This article has not been funded by any organizations or agencies. This independence ensures the research is conducted objectively, without external influence.
- **Ethical Approval and Consent to Participate:** The content of this article does not necessitate ethical approval or consent to participate, with supporting documentation.
- **Data Access Statement and Material Availability:** The adequate resources of this article are publicly accessible.

  Dataset 1- The supporting data for this article are available at: https://www.kaggle.com/datasets/ronaldross/video-steganography.

  Dataset 2- The supporting data for this article are available at: https://www.kaggle.com/datasets/matthewjansen/ucf101-action-recognition.
- **Author's Contributions:** The authorship of this article is contributed equally to all participating individuals.

REFERENCES

1. (2022). Low-resolution face recognition based on feature-mapping face hallucination. Computers and Electrical Engineering 101. DOI: 10.1016/j.compeleceng.2022.108136
2. Aakerberg, A., Nasrollahi, K. & Moeslund, T. B. (2021). Real-world super-resolution of face images from surveillance cameras. IET Image Processing 16(8), pp. 1234-1245. DOI: 10.1049/ipr2.12359
3. Grm, K., Scheirer, W. J. & Struc, V. (2020). Face Hallucination Using Cascaded Super-Resolution and Identity Priors. IEEE Transactions on Image Processing 29(1), pp. 2150-2165. DOI: 10.1109/TIP.2019.2945835
4. Wang, Z., She, Q. & Ward, T. E. (2021). Generative Adversarial Networks in Computer Vision: A Survey and Taxonomy. ACM Computing Surveys 54(2). DOI: 10.1145/3439723
5. Tustison, N. J., Avants, B. B. & Gee, J. C. (2019). Learning image-based spatial transformations via convolutional neural networks: A review. Magnetic Resonance Imaging 64, pp. 142-153. DOI: 10.1016 /j.mri.2019.05.037
6. (2023). ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks - Technical Documentation and Model Overview. ESRGAN ReadTheDocs. https://esrgan.readthedocs.io/en/latest/index.html
7. Zhu, Z., Lei, Y., Qin, Y. & Zhu, C. (2023). IRE: Improved Image Super-Resolution Based on Real-ESRGAN. IEEE Access 99, p. 1. DOI: 10.1109/ACCESS.2023.3256086
8. Choi, Y. & Park, H. (2023). Improving ESRGAN with an additional image quality loss. Multimedia Tools & Applications 82(2), p. 3123. DOI: 10.1007/s11042-022-13452-4
9. Tomar, A. S., Arya, K. & Rajput, S. S. (2024). Learning face super-resolution through identity features and distilling facial prior knowledge—Expert Systems with Applications 202. DOI: 10.1016/j.eswa.2024.125625
10. Bashir, S. M., Wang, Y., Khan, M. & Niu, Y. (2021). A comprehensive review of deep learning-based single image super-resolution. PeerJ Computer Science 7. DOI: 10.7717/peerj-cs.621
11. Zhang, W., Zhao, W., Li, J., Zhuang, P., Sun, H., Xu, Y. &

Li, C. (2024). CVANet: Cascaded Visual Attention Network for Single Image Super-Resolution. Neural Networks 170, pp. 622-634. DOI: 10.1016/j.neunet.2023.11.049
12. Zhu, Y., Zhang, Y. & Yuille, A. L. (2014). Single Image Super-Resolution Using Deformable Patches. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 2917-2924. DOI: 10.1109/CVPR.2014.373, works remain significant, see the declaration
13. Anwar, S., Khan, S. & Barnes, N. (2020). A Deep Journey into Super-resolution: A Survey. ACM Computing Surveys 53(3). DOI: 10.1145/3390462
14. Yu, X., Fernando, B., Hartley, R. & Porikli, F. (2020). Semantic Face Hallucination: Super-Resolving Very Low-Resolution Face Images with Supplementary Attributes. IEEE Transactions on Pattern Analysis and Machine Intelligence 42(11), pp. 2926-2943. DOI: 10.1109/TPAMI.2019.2916881
15. Rana, M. S., Nibali, A. & He, Z. (2022). Selection of object detections using overlap map predictions. Neural Computing and Applications 34. DOI: 10.1007/s00521-022-07469-x
16. Hu, X., Fan, Z., Jia, X., Li, Z., Zhang, X., Qi, L. & Xuan, Z. (2021). Towards effective learning for face super-resolution with shape and pose perturbations—Knowledge-Based Systems 220. DOI: 10.1016/j.knosys.2021.106938
17. Dastmalchi, H. & Aghaeinia, H. (2022). Super-resolution of very low-resolution face images with a wavelet-integrated, identity-preserving adversarial network. Image and Vision Computing 107. DOI: 10.1016/j.image.2022.116755
18. (2020). Facial Image Synthesis and Super-Resolution with Stacked Generative Adversarial Network. Neurocomputing 402, pp. 359-365. DOI: 10.1016/j.neucom.2020.03.107
19. Arabboev, M., Begmatov, S., Rikhsivoev, M., Nosirov, K. & Saydiakbarov, S. (2021). A comprehensive review of image super-resolution metrics: classical and AI-based approaches. Acta IMEKO 13(1). DOI: 10.21014/actaimeko.v13i1.1679
20. (2021). Py-Feat: Python Facial Expression Analysis Toolbox. PMC10751270. DOI: 10.3389/fnins.2021.635019
21. He, K., Pu, N., Lao, M. & Lew, M. S. (2023). Few-shot and meta-learning methods for image understanding: a survey. International Journal of Multimedia Information Retrieval 12. DOI: 10.1007/s13735-023-00279-4
22. Lopes, A., Santos, F. P., Oliveira, D. d., Schiezaro, M. & Pedrini, H. (2024). Computer Vision Model Compression Techniques for Embedded Systems: A Survey. Computers & Graphics 123. DOI: 10.1016/j.cag.2024.104015
23. Zhang, Z., Shi, Y., Zhou, X., Kan, H. & Wen, J. (2020). Shuffle block SRGAN for face image super-resolution reconstruction. Measurement and Control 53(78), pp. 1429-1439. DOI: 10.1177/0020294020944969
24. Mishra, A., Lee, B. & Lee, B. (2025). PixelBoost: Leveraging Brownian Motion for Realistic-Image Super-Resolution. IEEE Transactions on Multimedia 99, pp. 1-13. DOI: 10.1109/TMM.2025.3623517
25. Zhang, Y., Tian, Y., Kong, Y., Zhong, B. & Fu, Y. (2018). Residual Dense Network for Image Super-Resolution. arXiv preprint arXiv:1802.08797. DOI: 10.48550/arXiv.1802.08797
26. Li, M., Zhang, Z., Yu, J. & Chen, C. (2020). Learning Face Image Super-Resolution Through Facial Semantic Attribute Transformation and Self-Attentive Structure Enhancement. IEEE Transactions on Multimedia 99, p. 1. DOI: 10.1109/TMM.2020.2984092
27. Musunuri, Y. R. & Kwon, O. (2021). Deep residual dense network for Single Image Super-Resolution. Electronics 10(5). DOI: 10.3390/electronics10050555
28. Fan, Z., Hu, X., Chen, C., Wang, X. & Peng, S. (2020). Facial image super-resolution guided by adaptive geometric features. EURASIP Journal on Wireless Communications and Networking 2020. DOI: 10.1186/s13638-020-01760-y
29. (2021). Improved Face Image Super-Resolution Model Based on Generative Adversarial Network. MDPI 11(5). DOI: 10.3390/23134331
30. (2022). Deep Learning for Face Super-Resolution: A Techniques Review. APSIPA Transactions on Signal and Information Processing 11(1), pp. 1-15. DOI: 10.1561/116.20240045
31. Johnson, J., Alahi, A. & Fei-Fei, L. (2016). Perceptual Losses for Real-Time Style Transfer and Super-Resolution. arXiv preprint arXiv:1603.08155. DOI: 10.1007/978-3-319-46475-6_43
32. Georgopoulos, M., Oldfield, J., Nicolaou, M. A., Panagakis, Y. & Pantic, M. (2021). Mitigating Demographic Bias in Facial Datasets with Style-Based Multi-attribute Transfer. International Journal of Computer Vision 129. DOI: 10.1007/s11263-021-01448-w
33. (2020). Facial image super-resolution guided by adaptive geometric features. Journal on Wireless Communications and Networking. DOI: 10.1186/s13638-020-01760-y
34. Joze, H. R., Zharkov, I., Powell, K., Ringler, C., Liang, L., Roulston, A., Lutz, M. & Pradeep, V. (2020). ImagePairs: Realistic Super Resolution Dataset via Beam Splitter Camera Rig. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. DOI: 10.1109/CVPRW50498.2020.00267
35. Tejy, A. R., Halder, S. S., Shandeelya, A. P. & Pankajakshany, V. (2020). Enhancing Perceptual Loss with Adversarial Feature Matching for Super-Resolution. arXiv preprint. https://arxiv.org/pdf/2005.07502
36. Liu, Y., Liu, Y. & Liu, Y. (2017). Tensor Super-Resolution with Generative Adversarial Nets. IEEE Transactions on Image Processing 26(1), pp. 1-12. DOI: 10.1109/TIP.2016.2610190
37. Liu, Y., Sun, D., Wang, F., Lim, K. P., Chiew, T. K. & Lai, Y. (2022). Improved Face Image Super-Resolution Model Based on Generative Adversarial Network. IEEE Transactions on Circuits and Systems for Video Technology 32(1), pp. 1-14. DOI: 10.1109/TCSVT.2021.3071234
38. Kim, D., Kim, M., Kwon, G. & Kim, D. (2019). Progressive Face Super-Resolution via Attention to Facial Landmark. arXiv preprint arXiv:1908.08239. DOI: 10.48550/arXiv.1908.08239
39. Eugenesiow. (2026). Dataset Card for Div2k. Hugging Face. https://huggingface.co/datasets/eugenesiow/Div2k
40. Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Loy, C. C., Qiao, Y. & Tang, X. (2018). ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks. ECCVW 2018. DOI: 10.1109/ICCVW.2018.00135
41. (2021). Synchronizing billion-scale automata. Information Sciences 574, pp. 162-175. DOI: 10.1016/j.ins.2021.05.072
42. (2024). Peak signal-to-noise ratio. Wikipedia. https://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio
43. Karras, T., Aittala, M., Hellsten, J., Laine, S., Lehtinen, J. & Aila, T. (2020). Training Generative Adversarial Networks with Limited Data. arXiv preprint arXiv:2006.06676. DOI: 10.48550/arXiv.2006.06676
44. He, K., Zhang, X., Ren, S. & Sun, J. (2016). Deep Residual Learning for Image Recognition. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778. DOI: 10.1109/CVPR.2016.90
45. (2024). Learned Perceptual Image Patch Similarity (LPIPS). PyTorch-Metrics 0.8.0 documentation. https://torchmetrics.readthedocs.io/en/v0.8.0/image/learned_perceptual_image_patch_similarity.html

AUTHOR'S PROFILE

Dr. Sheetal S. Patil is an accomplished academician, researcher, and educator in Computer Engineering, with expertise in Artificial Intelligence, Machine Learning, Deep Learning, Cybersecurity, Internet of Things (IoT), Blockchain Technology, and Healthcare Informatics. She earned her PhD in Computer Engineering from Bharati Vidyapeeth Deemed University, Pune, and has actively contributed to research, teaching, and innovation in emerging computing technologies. Her scholarly contributions include publications in reputed journals, conference proceedings, and international research platforms, covering topics such as medical image synthesis using deep learning, AI-enabled assistive technologies, IoT security, blockchain-based data management, and emotion recognition from EEG signals. Dr Patil is dedicated to fostering research-driven education and interdisciplinary collaboration. Through her work, she strives to develop intelligent, secure, and socially impactful technological solutions that address real-world challenges in healthcare, communication, and digital ecosystems.

Dr. Avinash M. Pawar is an accomplished academician, researcher, and academic administrator in Mechanical Engineering, with over 25 years of teaching and professional experience. He currently serves as Vice Principal and Head of the Department at Bharati Vidyapeeth's College of Engineering for Women, Pune, Maharashtra, India. He has contributed extensively to engineering education, academic administration, curriculum development, institutional development, and student mentoring. Dr Pawar holds a PhD in Mechanical Engineering from Bharati Vidyapeeth University and is a recognised PhD. Research Guide at Savitribai Phule Pune University (SPPU). He completed his M.E. and B.E. in Mechanical Engineering, both with First Class distinction. His teaching and research areas include Artificial Intelligence, Machine Learning, Deep Learning, Advanced Manufacturing Processes, Electrochemical Machining (ECM), Engineering Graphics, Computer-

Aided Manufacturing, Heat Transfer, and Computational Fluid Dynamics. His research contributions include 40+ peer-reviewed publications, with several papers indexed in Scopus and Web of Science. His Scopus Author Profile (Author ID: 59231132400) records 24 Scopus-indexed documents, 91 citations, 88 citing documents, and an h-index of 5, with a Field-Weighted Citation Impact (FWCI) of 1.60. His research also encompasses sustainable engineering and advanced manufacturing, with publications contributing to areas associated with several UN Sustainable Development Goals. He holds three patents related to Electrochemical Machining tool design. Dr Pawar has authored three textbooks adopted by SPPU and has actively contributed to workshops, Faculty Development Programmes, syllabus revision, paper setting, examination coordination, and academic quality initiatives. He has served as a reviewer and editorial board member for reputed international journals and conferences, including IEEE-sponsored events. He is actively associated with professional, academic, research, and institutional activities and has contributed to NSS, environmental initiatives, youth development, and social outreach. His achievements reflect a sustained commitment to research, innovation, academic excellence, leadership, and holistic engineering education.

Dr. Nilofar Mulla is currently serving as an Associate Professor at Bharati Vidyapeeth's College of Engineering for Women, located in Pune, Maharashtra, India. She brings an impressive 17 years of experience in Computer Engineering to the institute. She earned her PhD in Computer Engineering, with a specialization in Software Engineering, which underscores her deep expertise and commitment to the advancement of the discipline. In her academic and professional journey, she has significant research interests, particularly in the domains of Software Engineering, Image Processing, and Machine Learning. Her work in Software Engineering includes developing and improving methodologies, tools, and techniques that enhance software development processes and ensure high-quality software projects. Moreover, her focus on Machine Learning reflects her engagement with cutting-edge technologies that enable systems to learn from data, improve their performance over time, and make intelligent decisions.

Dr. Jotiram K. Deshmukh is an accomplished academician, researcher, and academic administrator in Instrumentation Engineering, with over 26 years of teaching and professional experience. He currently serves as Head of the Department of Instrumentation Engineering at Bharati Vidyapeeth College of Engineering, Navi Mumbai. He has contributed extensively to engineering education, academic administration, curriculum development, institutional development, and student mentoring. Dr Jotiram K Deshmukh holds a PhD in E & TC Engineering from SPSU. He completed his M.E. and B.E. in Instrumentation Engineering, both with First Class. His teaching and research areas include Process Automation, Control Engineering, IoT, and Sensors**.** His research contributions include 21+ peer-reviewed publications, with several papers indexed in Scopus and Web of Science. He has 4 Indian patents.

Dr. Avinash M. Deshmukh is an accomplished academician, researcher, and academic administrator in Mechanical Engineering, with over 26 years of teaching and professional experience. He currently serves as Controller of Examination at SCTR's Pune Institute of Computer Technology, Pune, Maharashtra, India. He has contributed extensively to engineering education, academic administration, curriculum development, institutional development, and student mentoring. Dr. Avinash M Deshmukh holds a Ph.D. in Mechanical Engineering from College of Engineering (COEP) Pune. Affiliated to Savitribai Phule Pune University. He completed his M.E. and B.E. in Mechanical Engineering, both with First Class. His areas of teaching and research include the domain of Mechanical Engineering, HVAC&R, Heat Transfer, Additive Manufacturing, Engineering Drawing. Computer-Aided Manufacturing. His research contributions include 7+ peer-reviewed publications, with several papers indexed in Scopus and Web of Science. Having 4 grant Indian patents in credit.

Dr. Suresh Kumar Ray is Principal and Head of the Department of Community Health Nursing at Bharati Vidyapeeth (Deemed to be University) College of Nursing, Sangli. He holds an M.Sc. in Nursing and a Ph.D. in Nursing, and has over 21 years of teaching, research, and academic leadership experience. A committed educator and researcher, Dr. Ray has published 72 national and international research papers in global databases with 250+ citations, authored 3 books and 2 book chapters, and guided undergraduate, postgraduate, and Ph.D. scholars. He has received several academic honours, including the Global Academic Research Excellence Award and Best Researcher Award from the Rotary Club, Pune; the Leadership Award; and the Educationist Award. He has also participated in an international faculty exchange programme in Sweden.