scientific reports

Check for updates

OPEN

A promising AI based super resolution image reconstruction technique for early diagnosis of skin cancer

Nirmala Veeramani & Premaladha Jayaraman✉

Skin cancer can be prevalent in people of any age group who are exposed to ultraviolet (UV) radiation. Among all other types, melanoma is a notable severe kind of skin cancer, which can be fatal. Melanoma is a malignant skin cancer arising from melanocytes, requiring early detection. Typically, skin lesions are classified either as benign or malignant. However, some lesions do exist that don't show clear cancer signs, making them suspicious. If unnoticed, these suspicious lesions develop into severe melanoma, requiring invasive treatments later on. These intermediate or suspicious skin lesions are completely curable if it is diagnosed at their early stages. To tackle this, few researchers intended to improve the image quality of the infected lesions obtained from the dermoscopy through image reconstruction techniques. Analyzing reconstructed super-resolution (SR) images allows early detection, fine feature extraction, and treatment plans. Despite advancements in machine learning, deep learning, and complex neural networks enhancing skin lesion image quality, a key challenge remains unresolved: how the intricate textures are obtained while performing significant up scaling in medical image reconstruction? Thus, an artificial intelligence (AI) based reconstruction algorithm is proposed to obtain the fine features from the intermediate skin lesion from dermoscopic images for early diagnosis. This serves as a non-invasive approach. In this research, a novel melanoma information improvised generative adversarial network (MELIIGAN) framework is proposed for the expedited diagnosis of intermediate skin lesions. Also, designed a stacked residual block that handles larger scaling factors and the reconstruction of fine-grained details. Finally, a hybrid loss function with a total variation (TV) regularization term switches to the Charbonnier loss function, a robust substitute for the mean square error loss function. The benchmark dataset results in a structural index similarity (SSIM) of 0.946 and a peak signal-to-noise ratio (PSNR) of 40.12 dB as the highest texture information, evidently compared to other state-of-the-art methods.

Keywords Early disease detection, Melanoma, Artificial intelligence, Neural networks, Generative adversarial network, Super resolution, Image reconstruction, Deep learning, Residual learning

Skin cancer can affect individuals of all ages, especially those exposed to ultraviolet (UV) radiation. Among the various types, melanoma stands out as a particularly aggressive form, which can be life-threatening if not detected and treated early. Fortunately, the early-stage malignant lesions are curable and can be reversed. The clinicians analyze the abnormalities on lesions such as asymmetry, border, color and diameter called ABCD features[1]. But, these features are not visible to the naked eye, there comes the image processing techniques. Medical images obtained from the dermoscopic image are processed to keenly observe for its possible indicators for confirming the malignancy. However, some suspicious lesions fail to show up the explicit features[2]. Such lesions are called intermediate skin lesions, which are the ultimate focus of this research work. It is important to be vigilant in diagnosing intermediate lesions because they may turn into melanoma over time. Through AI-based reconstruction techniques, these intermediate skin lesions are processed to obtain more intricate information. Detecting numerous dermatologically relevant diseases, including melanoma and other skin cancers, is made possible by computer-aided methods[3–8]where image reconstruction is the crucial component of medical image analysis[9]. Many researchers have shown that the acquired dermoscopic images may have lower diagnostic accuracies when carried out by inexperienced dermatologists. Moreover, it is expected to be carried

School of Computing, SASTRA University, Thirumalaisamudram, Thanjavur 613401, Tamil Nadu, India. ✉email: premaladha@ict.sastra.edu

out by an expert with a whole load of years of experience diagnosing skin lesions. However, even the experts attain the diagnostic accuracy range between 75% and 84%[10]. Also, distinguishing benign from malignant lesions is a tedious task[11]. The intermediate evolving lesions are another perplexing concept in dermoscopic image analysis. Therefore, this research aims to extract the high-frequency feature indicators of intermediate skin lesions for expedited diagnosis of melanoma. The sample images of each class of skin cancer are acquired from the benchmark datasets, as shown in Fig. 1.

Furthermore, different techniques were adopted to overcome these diagnostic challenges. Especially in terms of melanoma, it is more important to predict cancer early before it gets severe and spreads to the lymph nodes[12]. The research proposed by Celebi et al. image analysis techniques adopted for medical tumour classification[13]. A preemptive diagnosis of lesions using a non-invasive method like image reconstruction prevents metastasis. Researchers are storming using deep learning mechanisms and SR reconstruction techniques in the field of medical imaging to strongly emphasize the advancements of AI. Moreover, deep learning-based skin lesion analysis leads to accurate diagnosis[14].

Following are the few contributing works on the SR reconstructions of images. The single two-dimensional image super-resolution (SISR) reconstruction is one of the crucial areas of indagation in machine vision and image analysis[15]. The resolution of an image is increased using image processing tools and signal processing techniques, which heavily rely on SR reconstruction. Assuming that the hardware characteristics of the imaging apparatus remain unchanged. Standard SISR reconstruction has three different groups of reconstruction techniques.They are interpolation of pixels[16], different restoration[17], and deep learning[18]. Convolutional neural networks (CNNs) are proficient at identifying the potential distribution properties of data and extracting abstract features at a high level from the source image. With image inputs in CNN, the deep learning-based SR study has attracted much attention[19]. Dong et al. is known for being the very first research team to experiment with super-resolution SNN-based SISR reconstruction[20]. The network layer counts between the range of 8 to 12, which is found to have a slow convergence rate and a smaller kernel size of 2. Additionally, the extracted local features of images make it challenging to recover the helpful information on textures, resulting in degraded iterative effects. Researchers have suggested that a fast super-resolution convolutional neural network (FSRCNN) diminishes the tuned parameters and escalates the total layers[21]. It swapped with two 3×3 convolution kernels instead of a 5×5 super-resolution convolutional neural network (SRCNN) kernel.

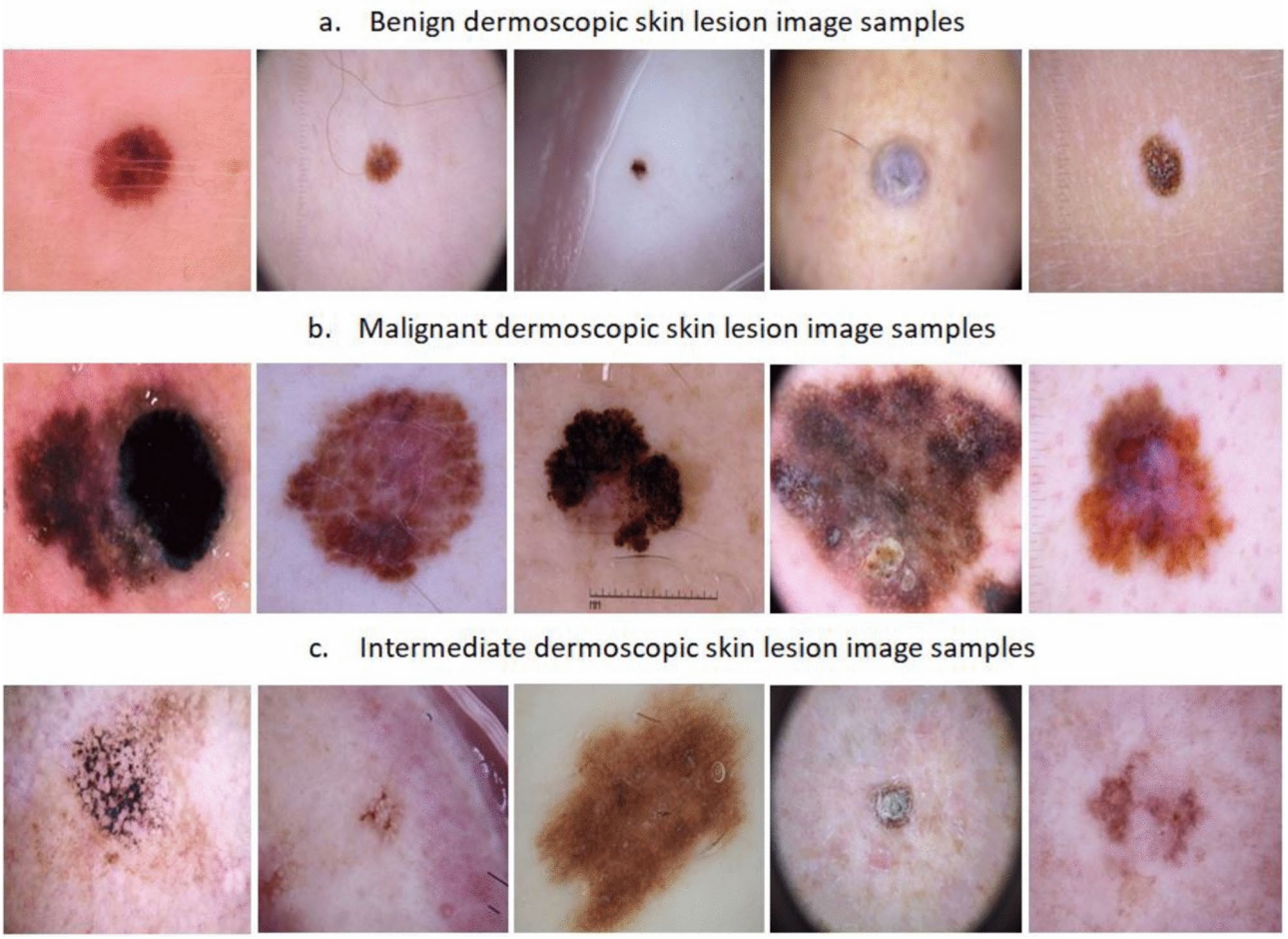

Fig. 1. Sample dermoscopic images of human skin lesion images a. Benign b. Malignant c. Intermediate.

Tolba et al. found that incorporating sub-pixel convolutional layers into the SRCNN approach would be preferable for images with poor quality[22]. Low-resolution (LR) images may be transformed efficiently and instantly into high-resolution (HR) images. Based on ResNet[23], a very deep super-resolution reconstruction (VDSR) uses 40 million parameters to increase the training and learning capacity of the network to twenty layers. It remarkably expedites the network's convergence through residual learning. It uses 40% less memory and performs better with larger network models. Researchers have also experimented with retinopathy images and named the model an enhanced deep super-resolution network (EDSR)[24]. Another study on deep laplacian pyramid networks for fast and accurate super-resolution (LapSRN) switched the L2 loss function to the Charbonnier loss function[25]. Exceptionally, the researchers trained the network under deep supervision and attained high-quality reconstruction[26].A deep residual network approach for learning the features of images is called image recognition ResNet. This residual network addressed the issue of gradient vanishing by increasing network layers[27].

In addition, experimentation with a generative adversarial network (GAN) was popularly stated as a super-resolution generative adversarial network (SRGAN). Ledig et al. is known as the first research team to employ GAN in image SR to improve image perceptions by boosting the realism of precise details[28]. In article[29], carried out the super-resolution on 2D tooth image processing with computed tomography (CT) images in the U-Net and sub-pixel networks. By adopting the SRGAN[30]enhanced super-resolution generative adversarial networks (ESRGAN), the researchers intensively used it for medical image analysis. The mean square error (MSE) loss is utilized by Wang et al. as a metric for evaluation[31]. For specialists, this technique produced positive outcomes and improved medical feature explanations, such as root canals' size, the shape of the patches, and archness.

This research ultimately focuses on diagnosing melanoma skin lesions at their early stages, including the intermediate lesions, using the morphological features obtained from the input skin lesion images. To get clarity on the abnormal lesions in the infected skin regions, the details of the lesions are carefully obtained. In general skin lesion images, the feature visibility is very minimal. Hence, the image's quality must be considered a primary concern. To focus on the perplexed lesion, the image quality aspect is dealt with prudently in this research, which has novel significant contributions as follows,

1. The quality of melanoma skin lesion images is enhanced by employing upscaling with residual attention blocks, enabling the extraction of intricate feature details from the images.
2. The generative adversarial model is utilized with loss function optimization during training, thereby reducing incorrect predictions.
3. Designed a novel super-resolution image reconstruction technique on melanoma images, leveraging stacked residual blocks for feature extraction.
4. Conducted the comparison of the proposed reconstruction techniques with other state-of-the-art methods and assessed the quality metrics.The paper is organized as section 2, which describes the existing super-resolution works and the application of GAN in medical image analysis. In section 3, our proposed architecture with its significance are given in detail. Section 4, the thorough discussion of the obtained results and the ablation, is depicted with a quantitative and qualitative analysis. Finally, section 5 tells us the limitations of the proposed work and the future road map of the work with a concluded summary.

Related works

The theoretical analysis of the earlier works is summarized in this section with the general architecture and the insights of GAN.

SRGAN

A GAN-based model comprised of two modules: A generator and a discriminator. In the SRGAN network, I_{SR} is the super-resolution image reconstructed, and I_{HR} represents high-resolution images. I_{LR} exhibits images represented in low resolution corresponding to I_{HR}, resulting in I_{HR} through the sub sampling process. Notations N_G and N_D denote generator and discriminator networks, respectively. D(x) denoted the loss communicated from the discriminator module. Fig. 2 depicts the schematic SRGAN model representation.

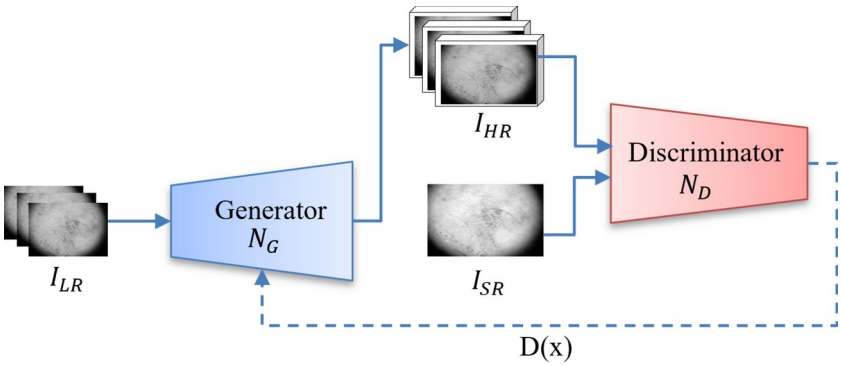

Fig. 2. Schematic block diagram of SRGAN model.

The generator produces I_{SR} lesion images from input I_{LR} images that modify the gradient following the discriminator's output and consistently improve image generation of N_G performance by reducing loss function to increase I_{SR}'s pixel per inch. The discriminator effectively functions to distinguish between two classes, and it classifies either to I_{HR} or to I_{SR}. The discriminator's loss function is employed with the cross-entropy technique, and [0, 1] is the output range. Additionally, the discriminator's optimization aim is to differentiate and identify between the resulting high and super-resolution images represented as I_{HR} and I_{SR}, respectively. In contrast, the generator's module in the network optimization target ultimately prevents the discriminator from being able to do so. Throughout the process, N_G and N_D's performance will advance through training to attain the Nash equilibrium state, which starts to smooth and lose image information. These issues are with the SRGAN despite having more significant effects on SR reconstruction.

Application of GAN in medical image analysis

GAN has a wide variety of applications. It is used in medical analyses that are categorized into a couple of significant aspects. The primary aspect of medical image generation is reconstruction. This adversarial network effectively populates many realistic melanoma images with appropriate annotations. To circumvent the redundancy and inadequacy of melanoma skin lesion datasets this kind of adversarial networks were adopted. Though many benchmark datasets are available with expert annotations in the existing datasets, there is an imbalance between the benign and the malignant illustrative classes. This affects the training of the deep learning networks for the skin lesion data. Also, it is found that the datasets are more prevalent in the classes of both malignant and benign. The problem of intermediate lesion images is left untouched with some gaps in recent research. Also, there is evidence of researchers who worked on the retinal with color fundus image generation using the GAN network. They used the binarized segmentation method to know the vessel globules and blobs[32].

A GAN network is used in the skin lesion analysis to emphasize the datasets image for training the deep neural network. This experiment employed the two class identifications for the grading results and retinal images to simulate high-resolution synthetic skin images[33]. Eventually, the retinal fundus image undergoes data augmentation techniques. Such as image flipping, flopping, scaling, and translation. Thus increased the accuracy of the existing fundus image generation with grading networks.

The secondary aspect is the discrimination direction, which can be considered a regularization strategy to improve the output of the generator network. Here, the preceding dissemination of the images from the outputs for image differentiation between several objects in image formats[34]. Notably, local details of the image achieved precise discrimination results. The researchers successfully performed the same with SR amplification of retinal images using the GAN network. It has a conjunction with internal feature loss to emphasize local structural elements[35]. GANs were chosen over other reconstruction networks because they excel at producing realistic, high-quality images by using a generator-discriminator setup. While traditional networks focus on minimizing pixel errors, GANs push the generator to create images that look natural and preserve fine details. This makes them especially effective for tasks like medical imaging, where texture and accuracy are crucial. Studies show that GANs perform well in upscaling and reducing noise, which aligns with our goal of improving image quality without losing important diagnostic features. Their ability to handle complex data distributions made them the best fit for our reconstruction needs.

Motivation & need for the study

The early diagnosis of skin cancer prevents the painful, invasive methods performed on the patients, which affect their normal lives. The keen observation of features or morphological structures hidden in suspicious skin lesion images is expected to be treated at the right time. Hence provided a technical shift in the early diagnosis of melanoma. The existing works on image quality reconstructions were carried out using the generative adversarial model to increase the number of training samples[36,37]. Also, they performed rigorous generation of synthetic images[38]. In contrast, this proposed research article presents the technical road map to improve the image's quality without losing feature information. Hence, the intricate features of the intermediate skin lesion information are preserved and used for analyzing malignancy at an early phase.

Proposed methodology

Owing to the GAN network's superior image simulation generating capabilities, it can assist in overcoming the challenges faced by existing deep learning in dermoscopy imaging. In addition address the data shortage in intermediate skin lesion analysis, and make deep learning a more generally applicable area. Additionally, the particular discriminant network of the GAN network is crucial for processing skin lesion input that can assist in producing finer local features of the network output.

The research presents a MELIIGAN reconstruction technique for skin lesion images based on SRGAN. It enhances the formation of confrontation networks by combining the attention block mechanism with the improved deep super-resolution (IDSR) network and the hybrid loss function. In fig. 3 and 4, the organized blocks of the network architecture are displayed.

1. Residual block with an attention CNN is designed for adaptive feature refinement and boost intermediate skin lesion's high-frequency edge detailed information in the image.
2. To ensure regular GAN training, designed a hybrid loss function with aggregation of the regular term as TV.
3. To achieve feature information from the intermediate skin lesions and a better fitting result, discard the Batch Normalization (BN) layer and add updated stacked residual blocks.

These contributions are majorly designed to obtain richer information from the low-resolution image. The features are then subjected to knowledge distillation that finds the features of intermediate skin lesions. It

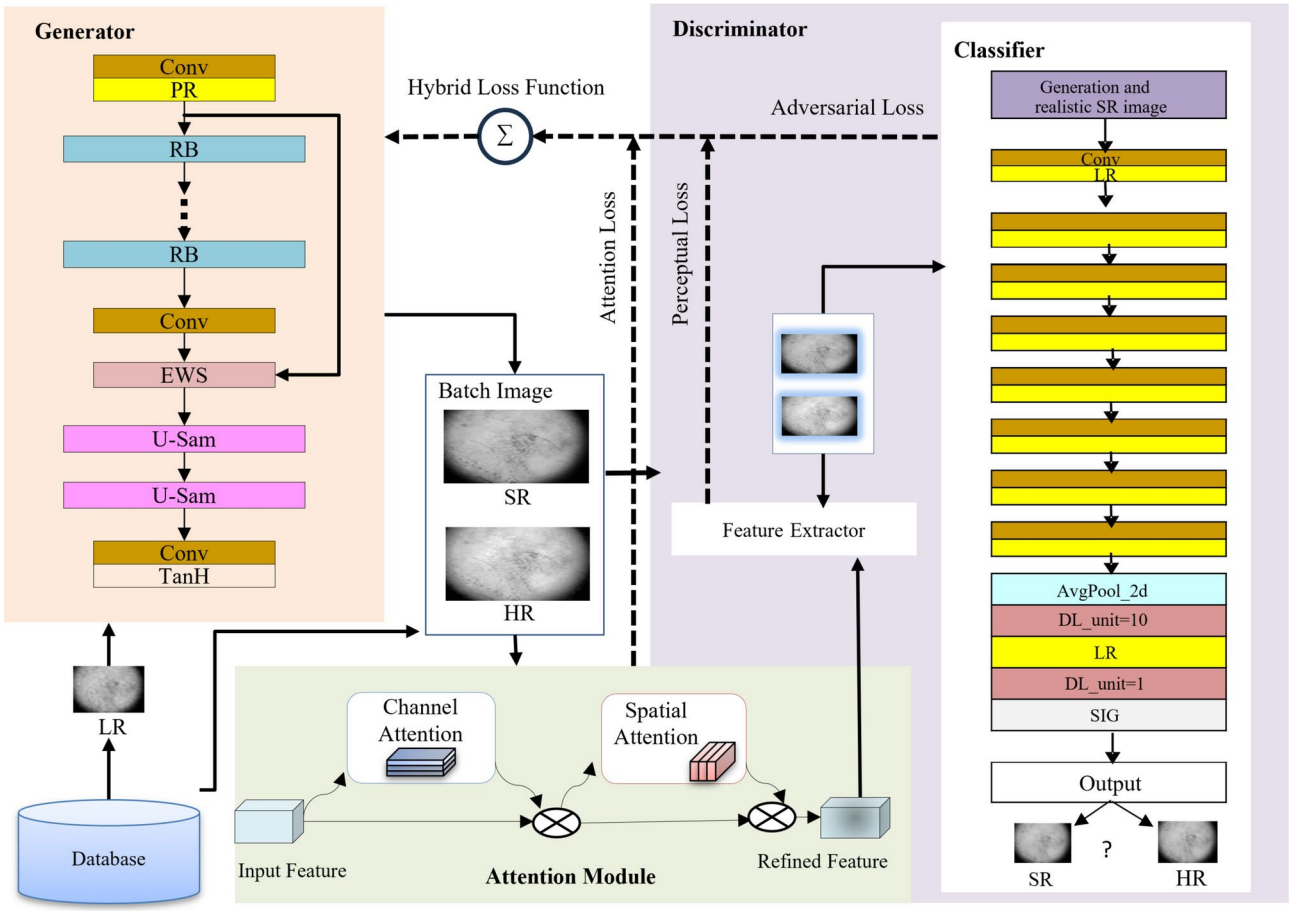

Fig. 3. Representation of the proposed MELIIGAN architectural workflow.

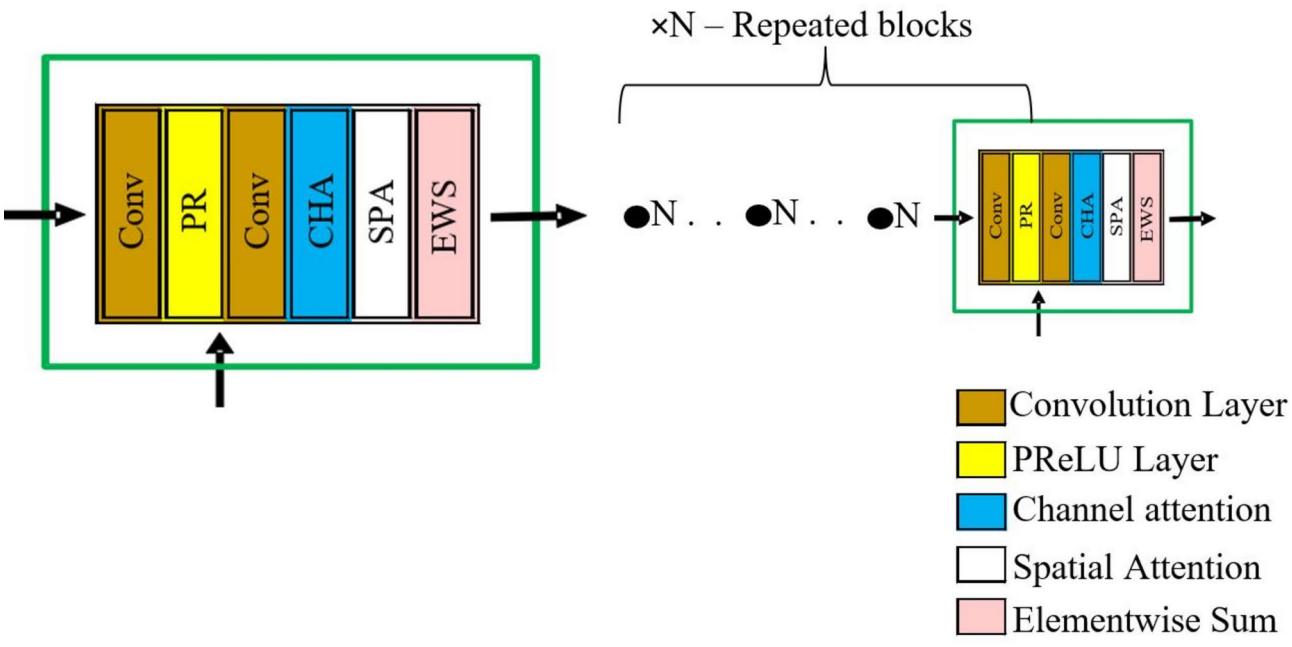

Fig. 4. Representation of residual blocks of the generator module.

removes redundant information during feature extraction that has been carried out in the generator network. Also, a non-classifier discriminator ensures training stability and enhances the generated image's quality.

Design of residual attention block

The residual model is widely practiced in deep CNN, which resolves the vanishing gradient problem due to the more profound escalation of the network by availing of simple short connections. To increase the network's depth and improve the extraction of deeper layers, including the channel and spatial attention blocks added to the residual module[39]. Fig. 5 depicts the structure of the residual attention network structure of the proposed work. The majority of the information in LR skin lesion images is low-frequency information, and there is typically significantly less high-frequency information. The new proposed approach eliminates the BN layers in the deep stacked residual attention block to circumvent the blurring of the intricate details to be fetched from the image.

The central part of the skin lesion image datasets consists of the smooth areas in the skin lesion images, including the body of the image information, which is simpler to recreate. However, the significance of edge key details is diminished during the image reconstruction course. It is well known that human visual inspection has a poor sensory effect on loss in high-frequency information and a deficient degree of edge-sharpness. So, here, the geometrical dimension of each input feature map is compressed to calculate the channel attention value in the proposed model. Moreover, the spatial data is kept aggregated using the average pooling method.

The concurrent usage of max pooling and average pooling is practiced to boost the high-frequency input's fitting effect and strengthen the network's representation capability. It provided us with realistic, contingent reconstructed images. To create T_{avg}^{C} and T_{max}^{C}, the average set feature and maximum set feature altered the spatial information, i.e. the geometrical information in which the several scalar values are utilized as aggregated features map using the average and the maximum pooling layer. Then, this average image feature set can extricate the particular channel's comprehensive skin lesion details. The top-notched feature can exercise with the highly detailed frequency information, scrutinized towards the channel that holds the high-frequency information. Eventually, these couple of data are inputted into a shared network. The shared network generates

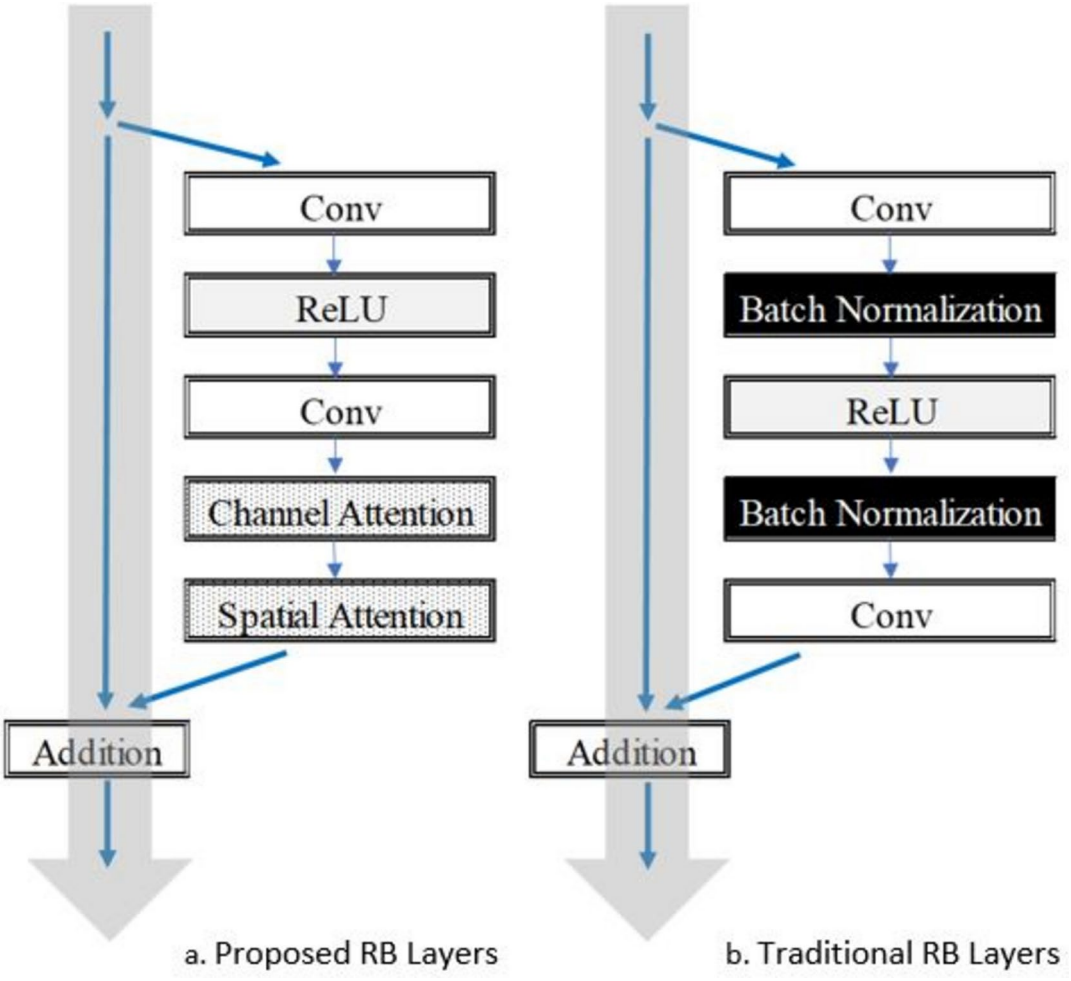

Fig. 5. Comparison of proposed architecture over traditional residual blocks.

a channel attention map $W_c \epsilon R^{c\times 1\times 1}$, which is a composition of some hidden squashing functions layer and the composition of the multi-layer perceptron[40].

According to the above-given description, calculated channel attention is given in equations (1) and (2).

$$W_c\,(T) = \sigma\,(MLP\,(P_{avg}\,(T))) + MLP\,(P_{max}\,(T)) \tag{1}$$

$$W_c\,(T) = \sigma\left(N_1\left(N_0\left(T^C_{avg}\right)\right)\right) + \left(N_1\left(N_0\left(T^C_{max}\right)\right)\right) \tag{2}$$

The notations used in the expression are as follows: $'\sigma'$ denoted the sigmoid activation function. $C \times H \times W$ is the measure to T size.$'N_0'$ Is the matrix w.r.t $C \times \frac{C}{d_r}$. N_1 Is the matrix w.r.t $\frac{C}{d_r} \times C$. Here, d_r Exhibits the decline rate. $W_c\,(T)$ represents the number of channels with C as the size factor.

Furthermore, specifying the class between benign and malignancy in the skin lesion images in SR lesion reconstruction is hard. Obtaining the intra-class differences between the benign and intermediate lesions is also challenging. Because it has a shallow high-frequency feature to assess the markers in the images, extracting the salient features from the targeted image is highly impossible in the LR images.

Therefore, a spatial attention module is added to generate an attention map stipulating the spatial association among the features[41]. The residual value is nearly close to zero in the high-frequency regions. The residual value is consistently more remarkable in the higher-frequency regions than low-frequency regions. It constructed a pooling layer and a channel axis to highlight futuristic detailed information areas of the skin image as shown in Fig. 6. Depicts the image flow under the residual network at the generator and discriminator module of the proposed architecture. Fusion occurs with the maximum pooling and the average pooling.

The information features are mapped in the channel axis. Here, T^C_{avg} and T^C_{max} represent the two-dimensional average and other max pooling functions, as given in equations (5) and (6). Finally, the spatial attention map is created by fusing the feature information between and surrounding them using a typical 3x3 convolutional layer.

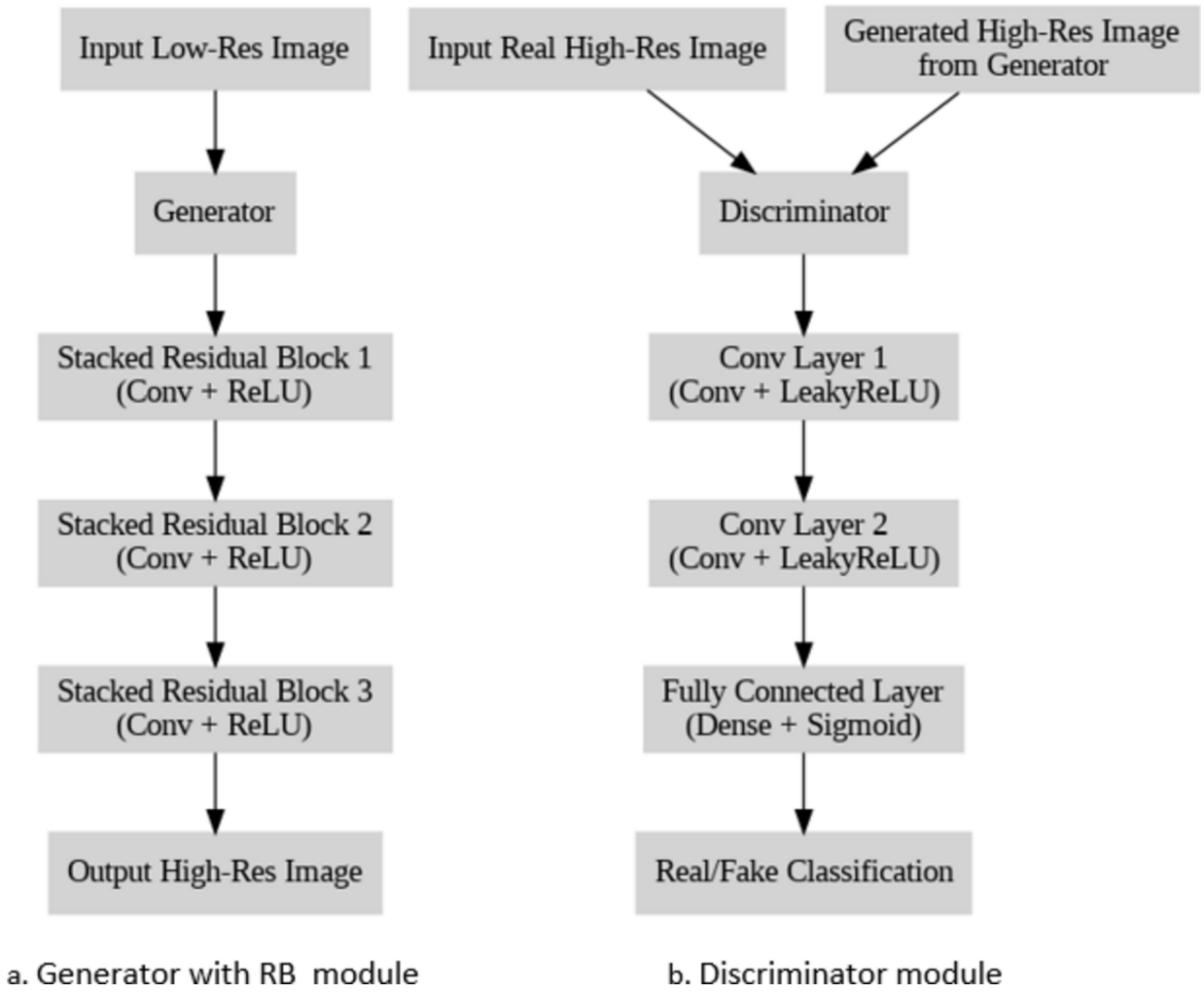

Fig. 6. Representation of proposed residual attention module layer architecture.

It is activated using the sigmoid function to obtain the value of spatial attention. Moreover, spatial attention is calculated in equations (3) and (4). .

$$W_s(T) = \sigma\left(f^{3\times3}\left(\left[P_{avg}(T); P_{max}(T)\right]\right)\right) \tag{3}$$

$$W_s(T) = \sigma\left(f^{3\times3}\left(T^s_{avg}; T^s_{max}\right)\right) \tag{4}$$

where the notations used in the expression are as follows, $f^{3\times3}$ indicates the 3 × 3 convolution, 'σ' indicates the sigmoid activation function. T^C_{avg} and T^C_{max} are calculated as follows,

$$T^s_{avg}(h,w) = \frac{\sum_{c=1}^{c}\left(T^c_2(h,w)\right)}{c} \tag{5}$$

$$T^s_{max}(h,w) = \max_{\{c=1..C\}}\left(T^c_2(h,w)\right) \tag{6}$$

where the notations used in the expression are as follows, (h,w) depicts the spatial location,$h\epsilon\{h=1..H\}$, $w\epsilon\{w=1..W\}$. T^C_2 representation of the c^{th} channel, $c\epsilon\{c=1..C\}$. C denotes the characteristic graph w.r.t. number of channels of a loss function.

Significance of proposed residual attention block

In addition to breadth and depth, attention is a crucial component of architectural design. Embedded residual attention design has been tried in multiple trials to enhance the suggested framework's performance on various skin lesion images. In addition to indicating where attention should be directed, attention enhances the depiction of the region of interest by emphasizing more informative aspects and hiding less useful ones. Also, the major value is that it is implemented as a stack of residual blocks in the context of residual learning to utilize better the non-redundant features from the L_R inputs and intermediate features. Additionally, the skip connections give the model an advantage by allowing it to ignore the large amount of low-frequency information in the image.

Engineering of the hybrid loss function

Loss functions have an important role in the performance of deep networks. Recently, various hybrid loss functions have been used in different models[42]. In the proposed work, the design of the loss function is the combination of perceptual and content loss.The ultimate aim to construct this hybrid loss function is to cut down the noise during image generation. It also adopted the Charbonnier loss function to produce the pixel loss function. Then, it incorporates the TV loss for the true purpose of minimizing the noise in the images[43]. This enhances the model's ability to accurately classify melanoma images by combining the strengths of different loss functions[44]. This approach aims to improve both the accuracy and true positives, leading to more reliable detection results on suspicious or intermediate skin lesions.

High-frequency information significantly impacts the MSE loss function during the SISR implementation. Additionally, it can increase the resulting image's PSNR value. However, it also produces many phantoms and erroneous pixels that alter the view and texture in some image regions. On the other hand, the MSE loss limitation is ineffective for a slight loss, and the penalties are imposed for tremendous losses that are higher. The output image is too smooth and has a shallow loss value, making it unsuitable for the terminal layer of the CNN with the activation function called sigmoid. The L1 loss function strictly constrains the high-frequency information from the low-resolution image samples to be recovered. It also takes a stretch of duration to rebuild. The reconstruction of skin lesion images eliminates the foggy effect and uneven pixels. The resultant loss function is compared with the Charbonnier loss function, which requires marginal resources for training the input images. It is considerably less than the state-of-the-art methods and also resulted in better PSNR value than the L1 and MSE loss functions as given in equation (7).

$$L_{c\left(y,\widehat{y}\right)} = \varepsilon E_{x,y} \sim P_{data(x,y)}\left(\sigma\left(y - G(x)\right)\right) \tag{7}$$

where the original image is represented as y, G(x), is the true output, and $\widehat{y}$ is the predicted output image resulting from the generator. Its mathematical expression is given as $\sigma(x) = \sqrt{x^2 + \varepsilon^2}$, and the value of ϵ is 10^{-7}. The noise present in the image significantly affects the results during the SISR reconstruction procedure. By including common regular expressions, the images can be kept smooth and eliminate compression artifacts. And so that the restoration of the image process could have left behind. And the TV loss in term of regularization is represented as δ. As given in equation (8), this research article adds the regular term TV as a hyper parameter.

$$L_{TV} = \delta\left\|\widehat{y}\right\|_{TV} \tag{8}$$

Equation (9) has all four components of the loss function that are employed in the discriminator network: the Charbonnier loss function L_c, the perceptual loss function termed as L_p, the adversarial loss function as L_a, and the ruling term as L_{TV}. The MELIIGAN's loss function parameters are $\alpha = 10^{-4}$,and $\beta = 10^{-3}$.

$$L = L_c + \alpha L_p + \beta L_a + \varepsilon L_{TV} \tag{9}$$

Charbonnier's loss function substitutes the true MSE, which competes with the perceptual loss function used in the overall calculation function. When the coefficient value is too high, the color difference is more pronounced than in the original image. The most negligible color and contrast differentiation between the original image occurs at the value of $\varepsilon = 2 \times e^{-8}$. The ultimate choice is $\varepsilon = 2 \times e^{-8}$ as the image has some spots. where $\alpha = 6 \times 10^{-3}$, $\beta = 10^{-3}$ and $\varepsilon = 2 \times e^{-8}$.

Process in normalization module

The CNN employed with BN layers specifically influence regularization up to certain limits, accelerating convergence and staving off model fitting. Hence, the BN is frequently employed in classification problems. However, the BN layer will prompt instability in training and significantly sway the network's output quality. Because SR reconstruction aims to interpolate pixels to make small images larger[45]. The color distribution of images is normalized after ensuing the BN layer[46]. However, medical image analysis must preserve the contrast information from the clinical images. Alternatively, skipping the normalization step in optimizing an algorithm makes features effective and cuts the training time. It allows casting up convolutional layers to the available resources to create an incentivized with the effect of the valuation. Moreover, it relinquished the BN layer and substituted the PReLU[47]activation function for the ReLU[48], as given in the network architecture in Fig. 6.

Experimental setup and results

Dataset and evaluation metrics

The high-quality, resolution images of the international skin imaging collaboration (ISIC) 2020 challenge dataset[49]that were utilized have different skin lesion images, including the intermediate skin lesions. This experiment uses a handcrafted set of 1350 training and 100 test samples. To be precise 40% of melanoma samples were included in the training and balanced between the other classes. The labels for these samples were obtained from the benchmark dataset available on Kaggle. Here, the validation set utilized 100 images from the same dataset. it is indeed considered during training. However, masks were not included, as our focus is on feature extraction and classification, rather than segmentation. This is why segmentation masks were not incorporated at this stage of the study. Also, the sample training images are tested using the PH2[50]and med node[51] datasets. All the experimental trials are performed using the LR and HR images at a four-scale. The input skin lesion images are cropped with the dimensions of 88 $\times$ 88, which refers to the height and width, respectively. Here, the size of the batches was set to 18, and the training epochs were trailed up to 85 due to the restrictions of the experimental setup. The hyper-parameters used in this exploratory study are $\alpha = 10^{-4}$, $\beta = 10^{-3}$, $\gamma = 10^{-3}$ and $\delta = 2 \times 10^{-6}$. The generator and discriminator module both employ the adam optimizer.

In addition, the system configuration that supports this research experiment is NVIDIA GeForce RTX 3080 GPU, CUDA 11.1, windows 10 operating system, and PyTorch 1.8.1 were used intensively to train and trail assessments. To confirm the effectiveness, the skin lesion images were rebuilt using our technique with the open-source benchmark ISIC 2019 dataset available at kaggle source[52]. The ISIC skin cancer dataset contains 3540 images of human skin lesions captured using dermoscopy. These were merged with the other 2260 skin lesion images for the inclusiveness of more skin lesion conditions.

The most critical indicators for evaluating our proposed work are PSNR and SSIM. While SSIM and PSNR are widely used metrics for evaluating SR performance, they primarily assess similarity rather than perceptual quality improvement. Our study used these metrics to quantify structural fidelity and pixel-wise similarity, which are crucial in medical imaging to ensure diagnostic accuracy. Here in the reconstruction, the pixels of the two images are said to have differences with the positions in the arrangement of pixels. In that case, the objective index of the PSNR facilitates ascertaining the likeliness of the two images, as given in equation (10). This method is widely adopted to assess the grade of the reconstruction of images. On the other hand, this SSIM deals with image contrast, luminousness and structure. These evaluation indicators are more intact with the level of the human visualization system. So, as the value gets higher, the images look alike, and the possible peak value of SSIM is set to the default of the value one.

$$PSNR = 10log \frac{255^2 \times W \times H \times C}{\sum_{p=1}^{W} \sum_{q=1}^{H} \sum_{r=1}^{c} \left[\overline{x}\,(p,q,r) - x\,(p,q,r)\right]^2} \tag{10}$$

where W, H, and C exhibit the width of an image, height dimension of an image, and several channels in image inputs, respectively. $\overline{x}$ indicates the generated super-resolution image, x denotes the input image. The p, q, and r are the indices representing the position of a pixel in the image along the width, height, and color channels, respectively.

$$SSIM(X,Y) = \frac{(2_{\mu X \mu Y} + C_1)\,(2_{\sigma XY} + C_2)}{(\mu_X^2 + \mu_Y^2 + C_1)\,(\sigma_X^2 + \sigma_Y^2 + C_2)} \tag{11}$$

where the mean of X and Y is exhibited as a notation μ_X and μ_Y respectively. Also, σ_X and σ_Y are the mean intensity values of images X and Y. Here, σ_{XY} represents the covariance value of image X and image Y. The constants are given as C_1 and C_2 in the given equation (11).

Results and discussions

The effectiveness of our research methodology is confirmed with the ablation study on three different skin lesion image datasets. These are compared with the other widely used classic SR reconstruction approaches considered. It includes the SRGAN, ESRGAN, ESPCN and Bicubic to assess and analyze the performance of the method provided in this proposed work. Also, it has got the piece of influence from IGAN[53] and compared it with the existing SRGAN and other reconstruction models to ascertain the efficacy of our model. Qualitative and quantitative evaluations are also conducted using handcrafted skin lesion image datasets.

Quantitative assessment

PSNR and SSIM are employed as indicators to measure the SR reconstruction methods. To quantitatively assess the improvements in the proposed approach, a comparative analysis is done with existing sophisticated SR reconstruction methods on reconstruction jobs[54]. Right now, the SRGAN network is among the top super-scoring networks. Because it is concerned about subjective visual effects and objective indications for the first time[55]. According to Table 1, the proposed network's PSNR index beats the previous value of 11.58 dB. The rise of the SSIM index to 0.132 shows that our proposed MELIIGAN effectively identifies and learns to distinguish the mapping association between the images of LR and HR images. This result significantly improves the features to distinguish the pre-malignant lesions from the benign lesions.

The SSIM index, which evaluates structural similarity between two images, reflects the model's ability to preserve and enhance perceptual features such as textures, edges, and fine details during image reconstruction. The observed improvement in SSIM indicates that the model has effectively captured these key features, thereby improving the model's ability to differentiate between pre-malignant and benign lesions. This enhancement allows for more accurate and robust lesion classification based on image quality and structural consistency, which is crucial for medical diagnosis.

Qualitative assessment

The grey level transformation during image pre-processing was used to simplify the images and highlight important features. By converting the images to grey scale, we removed unnecessary color information, allowing the model to focus on the intensity of the pixels, which is more relevant for detecting details like lesion boundaries. This not only reduces the complexity of the images but also helps the model focus on the key features that matter most in medical analysis, making the process more efficient and effective. The resolution of the skin lesion image reconstructed, as shown in Fig. 7. Using the ESPCN approach, is much better than that of the image reconstructed using interpolation of the bicubic method. However, the high-frequency details are inadequate. The dense texture is extremely smoothed, which does not define well the feature's identification of the intermediate skin lesions and the image that has flaws in the originality[56,57].

Overall, the image quality recovered using the ESRGAN approach is better. The thick texture of the image generates more explicit details on the image edges and texture information. The texture information of the referential image has significant differences. In the SRGAN method's reconstruction of the lesion image, it has artifacts, is somewhat blurry, and has jagged edges[58,59]. However, skin lesion images produced by the proposed technique are more similar to the reference image in terms of the detailed grains of the image. Moreover, the MELIIGAN outperformed in image brightness and contrast quality.

Several trials ascertain the effectiveness of the ablation study on the skin lesion images from the ISIC 2020 skin lesion image datasets. The study for ablation is performed in three trials in the experiment steps. The changes to the feature extraction module have been made to understand the pixel distraction in the generated image. The features extracted are based on a single scale rather than the 4x scale. i.e., the kernel size of 3 is used. The recorded PSNR and the SSIM scores drastically fell by 3 dB and 3.2 percentage, respectively. In the second trial, the changes to the residual blocks and the BN are now included between the two sub-pixel layers of the convolution network. By experimenting, observed that 4.62 dB decreased the value of PSNR and a 5 per cent decrease in the scores of SSIM. Finally, while performing the third trial of the ablation experiment, the aggregated hybrid loss function is removed. By removing the hybrid loss function, which has the functionalities of the aggregated regular term TV and Charbonnier loss function, the scores are random after several iterations. The value stabilized and lies in the range of fall by 6.58 dB and the 8 percentage for PSNR & SSIM respectively.

The scores are tabulated and shown in Table 2. It is now inferred that the single-scale feature extraction experiment on the skin lesion images reconstructed the poor visual quality images relatively shallow with our proposed work. Incorporating all the modules in the channel and spatial attention blocks of the enhanced deep super-resolution reconstruction produces better results. Moreover, it elevates the high-frequency information in the image to differentiate the intermediate lesions in the images.

Dataset	Metric	Bicubic	ESPCN	SRGAN	ESRGAN	IGAN	Proposed
ISIC 2020	PSNR	22.26	23.22	25.95	27.49	32.64	40.12
	SSIM	0.7352	0.783	0.7497	0.8234	0.864	0.9465
PH2	PSNR	20.43	25.8	28.62	29.12	31.78	38.84
	SSIM	0.618	0.7124	0.7599	0.7801	0.8689	0.9314
Mednode	PSNR	20.85	22.87	23.88	27.89	31.34	37.51
	SSIM	0.6587	0.6890	0.6975	0.803	0.8445	0.9215

Table 1. Quantitative evaluation of image super-resolution reconstruction effect.

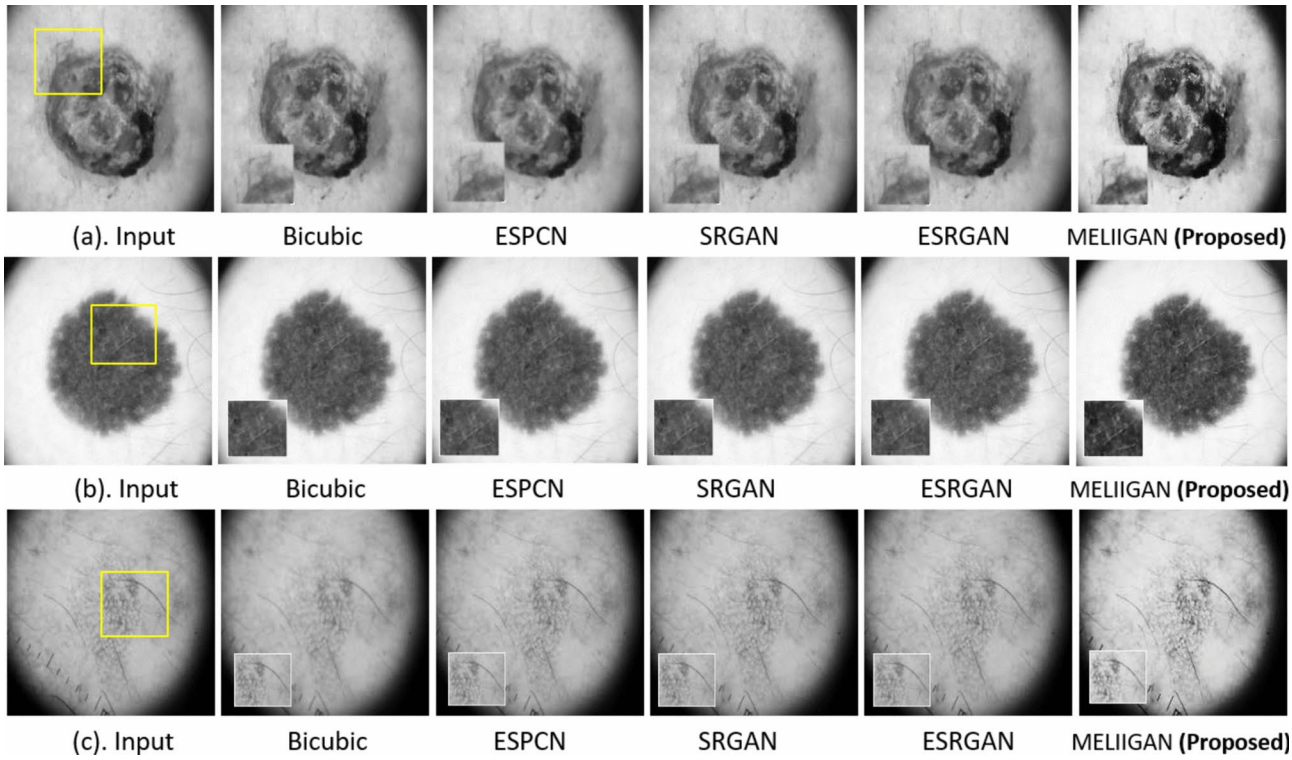

Fig. 7. Experimental results of our proposed work with comparison (**a**). Malignant lesions (**b**). Benign lesions, and (**c**). Intermediate skin lesions.

Ablation experiments	PSNR (dB)	SSIM
1X - Single scale	33.50	0.7813
With Batch Normalization	34.11	0.7964
Without Hybrid loss function	31.04	0.8317
2X - Double scale	37.72	0.8958
Proposed work	**42.12**	**0.9475**

Table 2. Ablation trail values with different scaling factors and loss functions.

Significance of the proposed work

The proposed MELIIGAN technique is designed for dermoscopic images that improve an image's resolution and details by creating a higher-resolution rendition from a lower-quality input. This technique can be implemented in several micro-level images to receive a focus of attention in various applications by fine-tuning. Because of its capacity to increase image quality and preserve the finer morphological details that were not visible in the original image.

One of the most notable advantages of the proposed work is the improvement in image categorization and fine-level feature detection accuracy. The proposed work has the following significance on medical images as shown in Fig. 8 and 9.

The inference is evident from the obtained features extracted from the MELIIGAN framework super-resolution skin lesion images. Here, the texture and fine details of the lesions are clear, facilitating the finding of new bio indicators. This high-frequency information from the images is ruled out to know the specific occurrences of shapes and pigments of the intermediate skin lesion for early diagnosis.

Improved feature extraction

In Fig. 10 histogram, the intentional variations of the resolution changes are impacted. The difference between input normal skin lesion images and resolution improved images deliberately. It shows the hidden pointed features in the resulting images. This facilitated the classifier model to recognize the intermediate class of the skin lesion better when compared to the traditional practice.

The pointed arrows and numbers represent the high-frequency image information indicating the pigments, streaks and finer patterns. These granular details are visible in the raw dermoscopic images. The obtained fine features are used to calculate the disease severity and possibility of the lesion metastasizing to melanoma, and several other findings are possible.

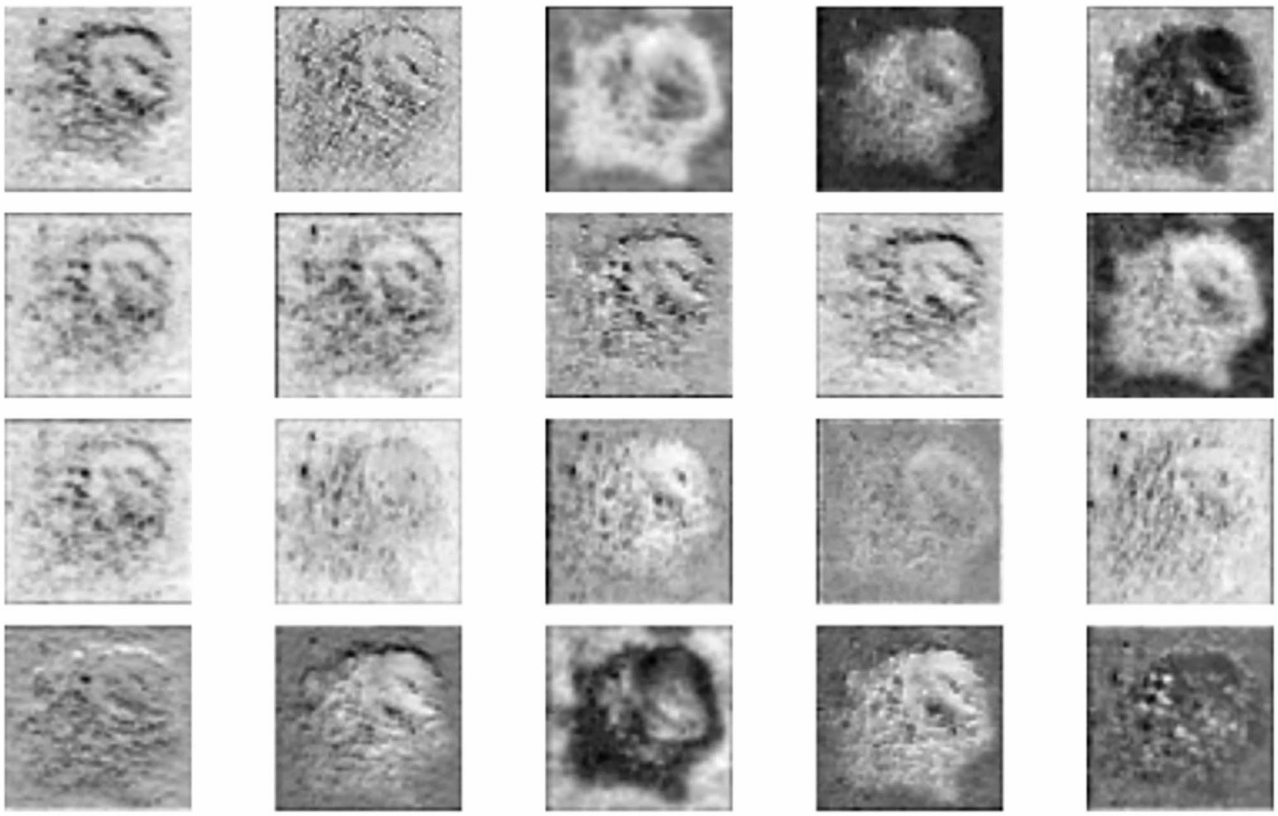

Fig. 8. The feature maps obtained from the low-resolution dermoscopic image input.

The scale factors are highly considered in extracting fine features. The scale factor is chosen upon the different trials on all three datasets. As a result, the scale factor $\times 4$ on the dermoscopic image performs well on the intricate fine lines. These details in the super-resolution images are then post-processed to assess the morphological structures of new imprinting. The bold text values in Table 3. represent the best values of the experiment. Also, this comparison of scale evaluation shows that the proposed MELIIGAN outperforms the existing methods concerning dermoscopic skin lesion images.

Enhanced context understanding

In medical images, the context of a feature within an image is critical for classification. However, the proposed work depicts the surroundings and context more clearly, assisting the classifier in grasping the distances between distinct parts. This additional information can contribute to more efficient classifications, particularly when diagnostics feature arrangement is important. Mitigating degradation factors: lower-resolution images are frequently marred by compression artifacts, noise, and blurriness. These factors can hurt classification algorithm performance by introducing false or useless information. It helped mitigate these degradation effects, as shown in Table 4 & 5, lowering the impact of artifacts and noise on the classification technique. The number of misclassified intermediate skin lesion samples is reduced drastically. Also, in the MELIIGAN-generated high-resolution images, the classification utilizes the high-frequency information exposed in the proposed work.

The increased discriminative power of the proposed work improves image perception quality by including visually consistent features with the scene. These extra characteristics can be very useful in distinguishing across visually identical classes, as shown in the confusion matrix Fig. 11 and 12. For example, medical imaging can make discriminating between malignant and benign tumours difficult. The proposed image reconstruction technique works precisely in the classification. It helps the classifier make more accurate predictions by highlighting tiny anomalies in the tumour shape and hidden image features. As the generates super resolution images preserves the highly intricate features of the image. This sophisticated the improved context of the image for diagnosing the intermediate skin lesions.Moreover, it is evident from the confusion matrix, that infers the models' ability to classify the intermediate lesions well after the reconstruction method.

Capturing fine textures

Textures and tiny patterns are frequently used to distinguish across classes. Super-resolution improves the visibility of these textures. Particularly allowing the classifier to distinguish between classes defined by detailed texture changes. Improved texture representation, for example, can greatly aid the classification of various material types or geological formations. Finally, the super-resolution improves image classification accuracy by delivering higher-quality images with increased details, lower degradation factors, improved context comprehension and

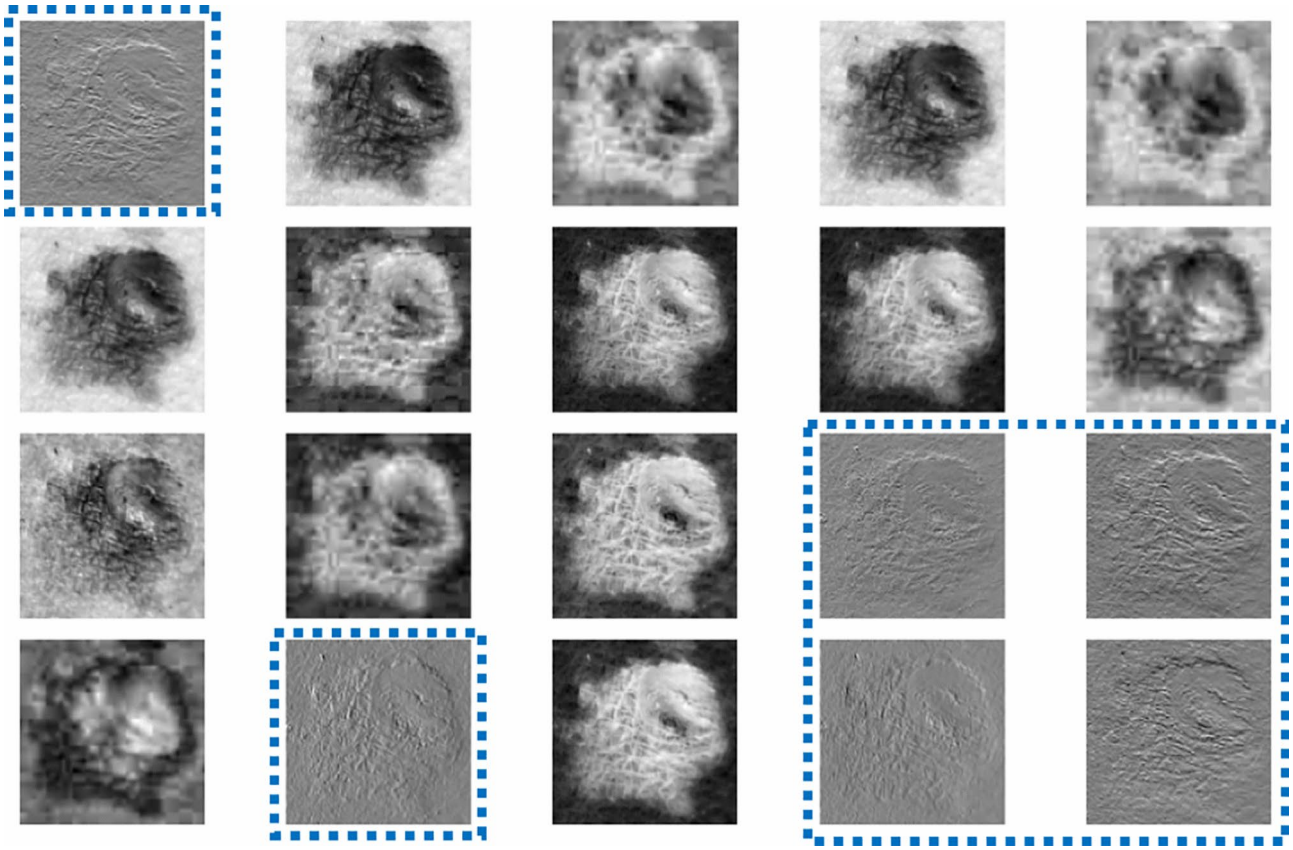

Fig. 9. The feature maps obtained from the proposed MELIIGAN super-resolution framework with significant high-frequency features - Dotted boxes represent the intricate high-frequency information on the image.

a more accurate texture representation. These enhancements directly contribute to classification models' greater discriminative ability, making them better. Now, it is able to accurately categorize images, especially when dealing with superficially similar or complicated classes. Incorporating super-resolution into the preprocessing phase can enhance image classification system performance, as shown in the t-distributed stochastic neighbour embedding (t-SNE) visual representation Fig. 13. The class wise results concerning the conventional features show that the clusters are closely packed for malignant and intermediate classes. The features using the intricate edge components facilitate us to identify the intermediate skin lesions in nearly stages. This also helps us prevent unattended lesions that get transformed into the malignancy stage. This t-SNE confirms that the proposed method allows us to separate the conjoined clusters using the intricate feature embedding with the original data.

Limitation of the proposed work

Our proposed MELIIGAN generates high-resolution images with hair but has significant issues identifying the quality of high-frequency features. The pre-processing with practical steps is needed to feed as an input in this GAN model. The greater occlusion of hair and ink patches influences the extraction of features. Thus, sample images from the dataset are given below in Fig. 14. In the extended work, GAN will develop an effective pre-processing technique for this hair removal[60]and utilize it in the MELIIGAN for high-resolution image generation. The generated high-resolution images also contain the details of the hair strands present in the skin that ultimately affect the performance of high-resolution generation[61]. Since the ultimate aim is to find the high-frequency features from the images[62].

The proposed work is implemented using the hand-picked images from the benchmark ISIC skin lesion dataset that excluded the over-hairy images. These over-occluded hairy images are more noisy in the generated super-resolution images. It has no point in identifying the intricate features in hairy images before applying any hair removal technique.

Conclusion and future directions

This research presents a MELIIGAN, an improved melanoma reconstruction technique in medical image analysis for skin lesion images. It is ultimately used to increase high-frequency (i.e., edge details of the input skin lesion image) by building residual blocks in the CNN model by designing an attention mapping function. Training the network by discarding the BN layer from the residual block also influences the image reconstruction quality. Finally, the TV regular term was employed with the hybrid loss function to smooth the training results before replacing the MSE loss function with the more durable Charbonnier. Hence, the proposed MELIIGAN

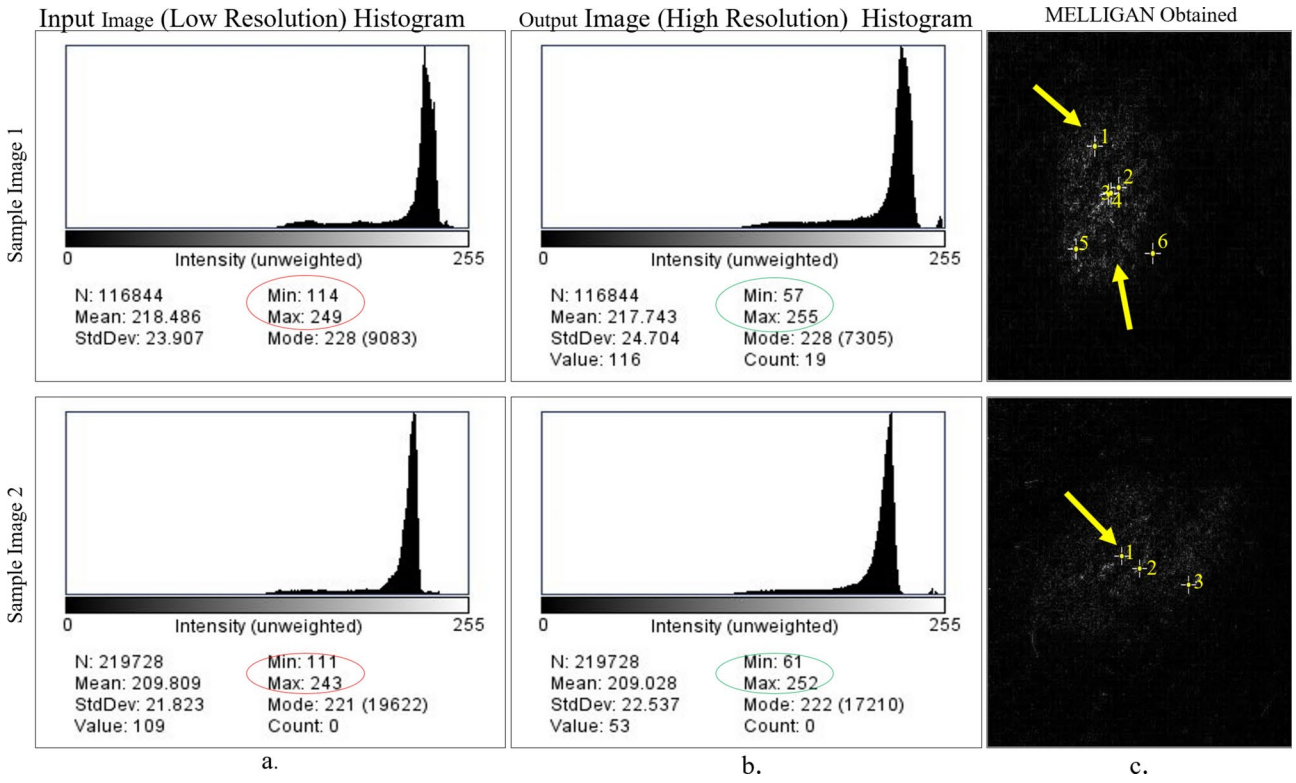

Fig. 10. Row 1 & 2: a. Input dermoscopic image b. Resulted high resolution image from the proposed work c. The difference features between a & b (Note: For better image visibility, refer to the web version of this work).

Dataset	Metrics	Scale	Bicubic	ESPCN	SRGAN	ESRGAN	IGAN	Proposed
ISIC 2020	PSNR	2	25.66	27.77	28.2	28.54	28.66	**28.86**
		3	22.39	23.84	24.28	24.59	**25.75**	25.64
		4	22.42	23.61	25.03	27.28	32.49	**32.87**
	SSIM	2	0.8499	0.869	0.8711	0.8744	0.9042	**0.9159**
		3	0.7882	0.8156	0.8233	0.8288	0.869	**0.8856**
		4	0.7304	0.7602	0.7941	0.8203	0.7828	**0.8249**
PH2	PSNR	2	27.26	29.37	29.9	30.14	30.26	**31.46**
		3	23.99	25.44	25.88	26.29	**26.45**	25.6
		4	22.02	23.21	25.73	28.88	31.49	**32.89**
	SSIM	2	0.8459	0.865	0.8671	0.8604	0.8402	**0.891**
		3	0.7942	0.8216	0.8293	0.8348	0.835	**0.8716**
		4	0.7164	0.7462	0.7601	0.7863	0.8628	**0.91**
Mednode	PSNR	2	26.32	24.68	26.12	28.45	29.56	**31.42**
		3	22.75	22.87	25.87	26.78	**26.89**	26.75
		4	21.23	23.5	24.2	28.63	32.55	**33.49**
	SSIM	2	0.7799	0.859	**0.8711**	0.8086	0.8342	0.8167
		3	0.7882	0.8156	0.8233	0.8288	0.8227	**0.8596**
		4	0.6504	0.7002	0.7941	0.8003	0.8489	**0.9082**

Table 3. Scale Factor evaluation with corresponding PSNR and SSIM metrics on skin lesion datasets.

technique produces rich and more transparent textures in the high-frequency features of the malignant lesion images and improves the SR reconstruction results. The extended work will primarily concentrate on finding new biological indicators for melanoma early detection[63]in terms of its structure and achieving higher-level artifact occluded image reconstruction with multi-modal image fusion[64] for future implementation.

Classification without Proposed MELIIGAN			
No.of Misclassified samples	Benign	Malignant	Intermediate
ISIC 2020 dataset	156	676	1168
Mednode dataset	218	731	1043
PH2 dataset	156	606	1238

Table 4. Number of Misclassified images in the class: Intermediate.

Classification with Proposed MELIIGAN			
No.of Misclassified samples/	Benign	Malignant	Intermediate
ISIC 2020 dataset	3	46	43
Mednode dataset	9	4	8
PH2 dataset	4	3	6

Table 5. Proposed work results on a number of misclassified images in the class: intermediate on different datasets.

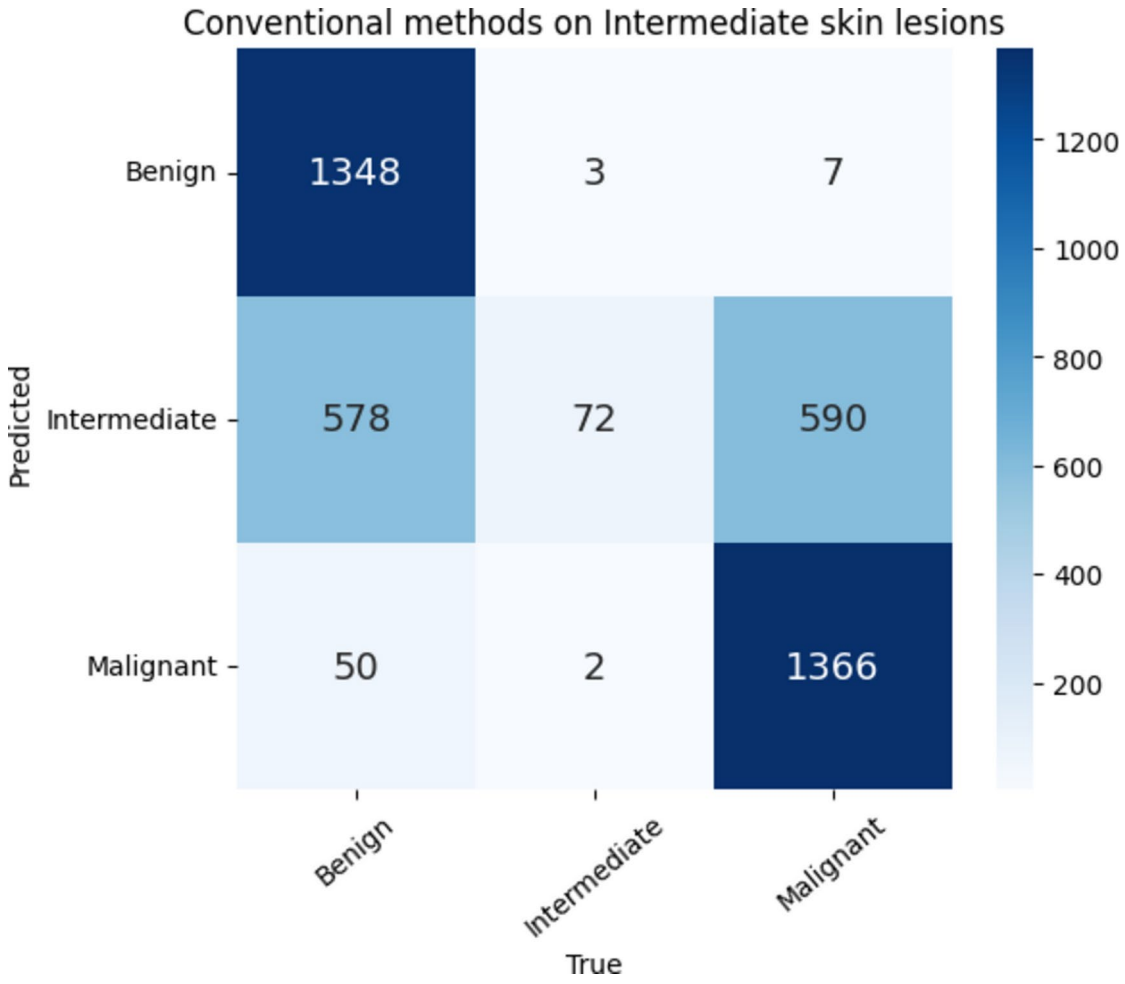

Fig. 11. Classification results on handcrafted samples of dermoscopic input skin lesion images.

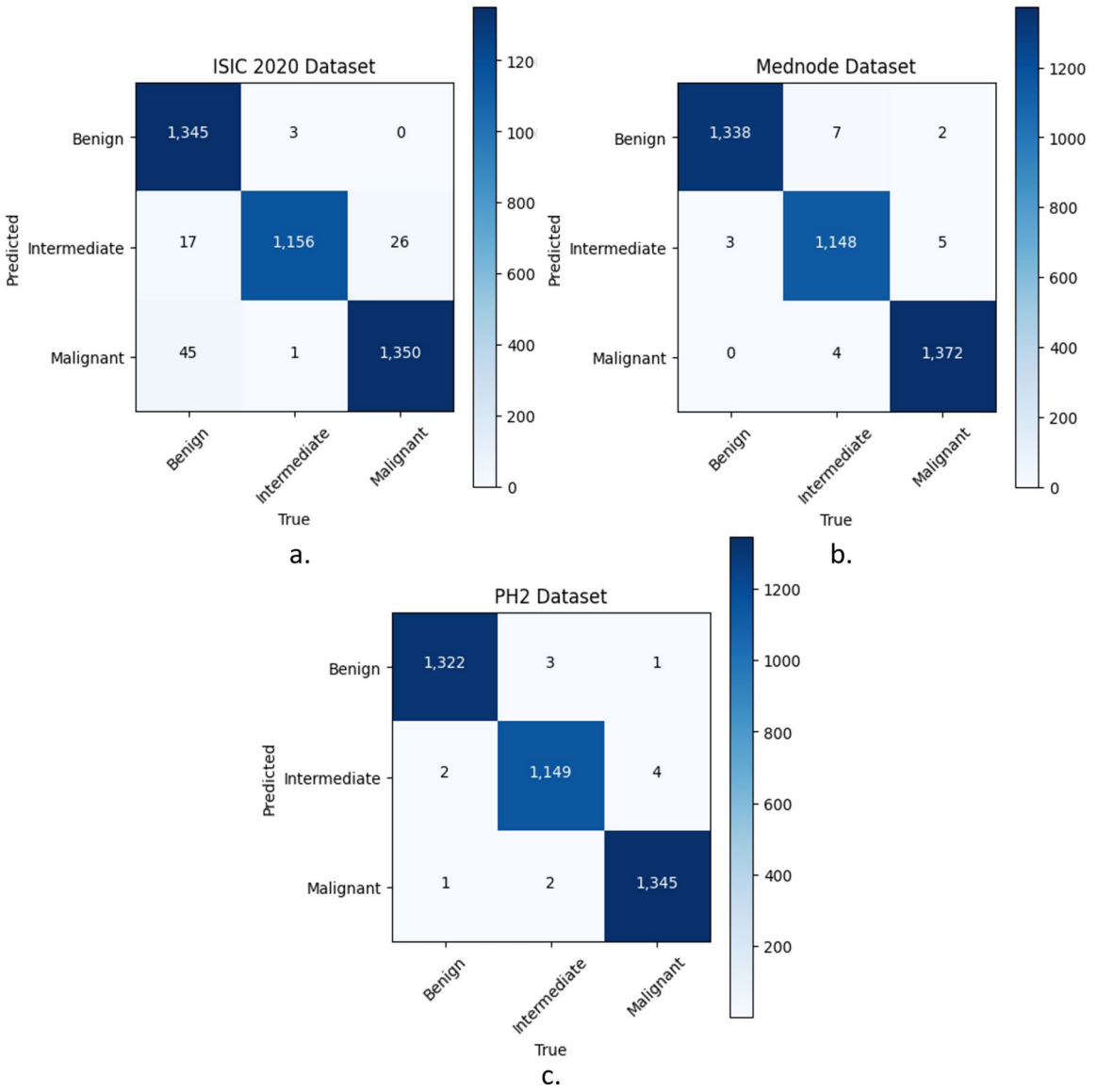

Fig. 12. Classification results on resolution increased output skin lesion images with proposed MELIIGAN.

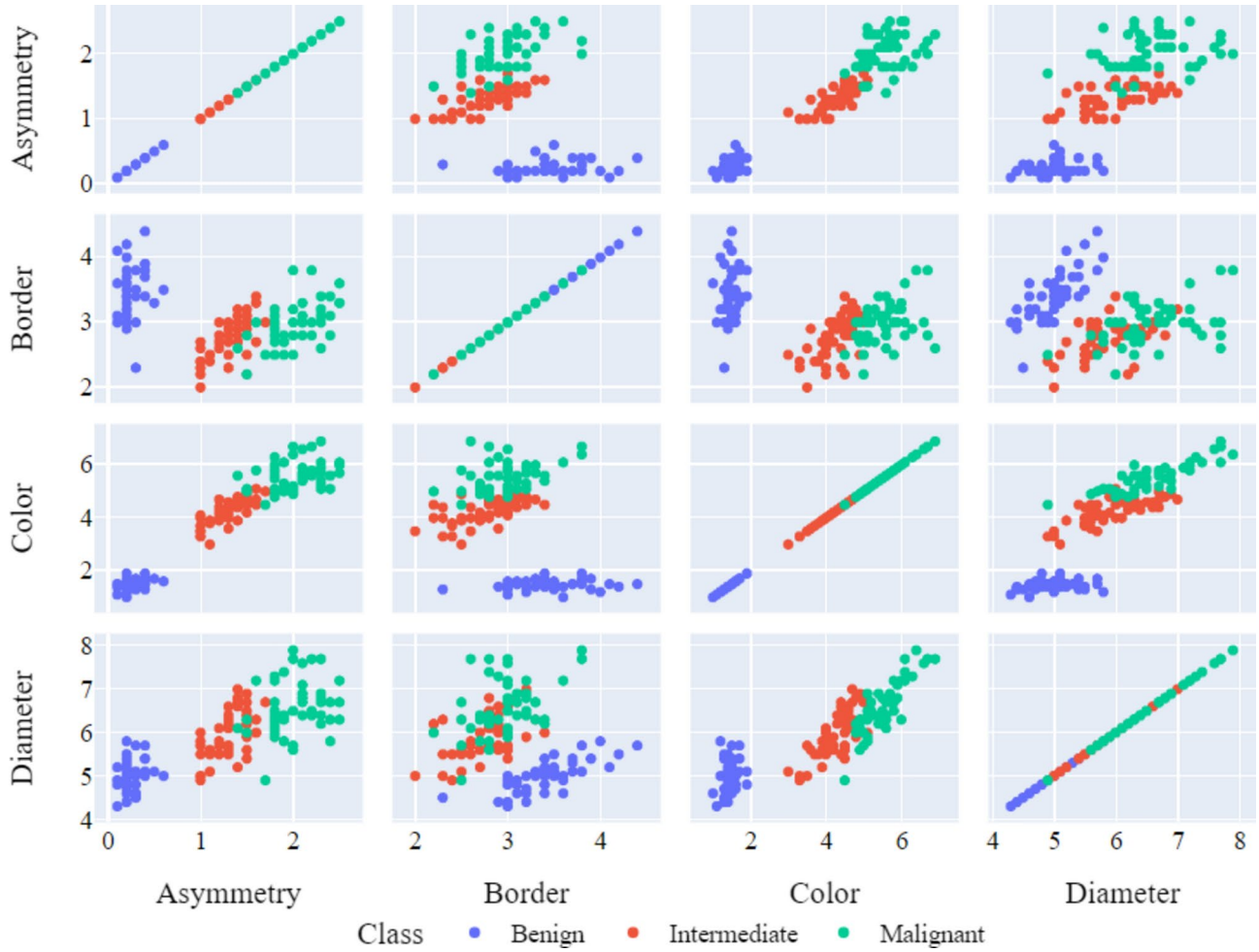

Fig. 13. t-SNE Visualization of ISIC sample data using the proposed super-resolution constructions.

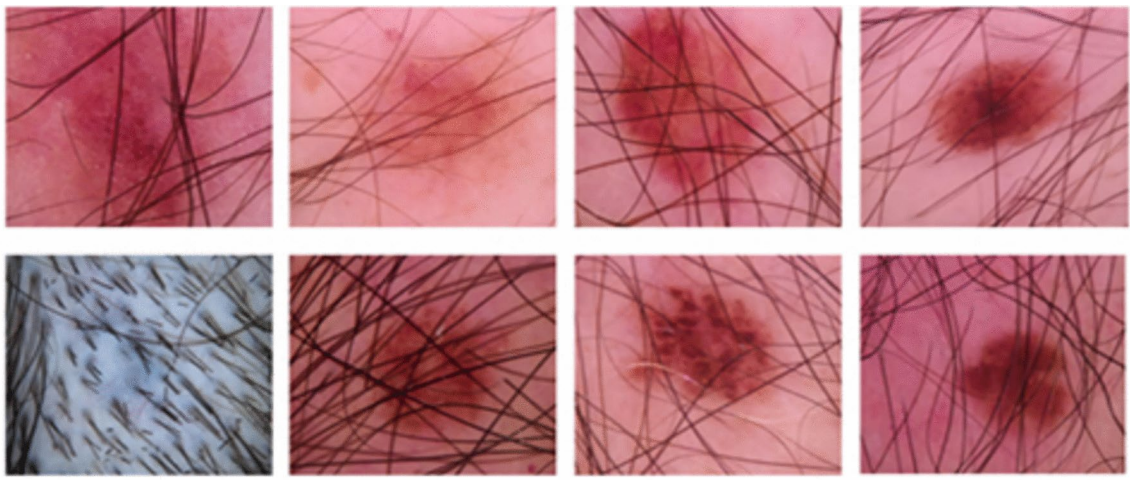

Fig. 14. Sample Skin lesion images from the benchmark datasets with occlusion of hair artifacts.

Data availability

The benchmark research materials used in this research can be downloaded from https://www.fc.up.pt/addi/ph2%20database.html, https://www.cs.rug.nl/~imaging/databases/melanoma_naevi/ and Kaggle ISIC 2020 benchmark dataset. Code availability for data preparation and the model of MELIIGAN is provided as Python note-

books at https://github.com/Nirmala-research/Image-Reconstruction. It will be made publicly available soon after the publishing of the work.

Received: 12 September 2024; Accepted: 6 February 2025
Published online: 11 February 2025

References

1. Jacinth Poornima, J., Anitha, J., Priya Henry, A. & Hemanth, D. J. in *Melanoma classification using machine learning techniques* 178–185 (IOS Press, 2023).
2. Nirmala, V., Shashank, H., Manoj, M., Satish, R. G. & Premaladha, J. in *Skin cancer classification using image processing with machine learning techniques* 1–15 (Auerbach Publications).
3. Balaha, H. M. & Hassan, A.E.-S. Skin cancer diagnosis based on deep transfer learning and sparrow search algorithm. *Neural Computing and Applications* **35**, 815–853 (2023).
4. Goceri, E. *Automated skin cancer detection: where we are and the way to the future*, 48–51 (IEEE, 2021).
5. Göçeri, E. *Convolutional neural network based desktop applications to classify dermatological diseases*, 138–143 (IEEE, 2020).
6. Göçeri, E. *Impact of deep learning and smartphone technologies in dermatology: Automated diagnosis*, 1–6 (IEEE, 2020).
7. Goceri, E. & Karakas, A. A. *Comparative evaluations of cnn based networks for skin lesion classification*, 1–6 (2020).
8. Goceri, E. Classification of skin cancer using adjustable and fully convolutional capsule layers. *Biomedical Signal Processing and Control* **85**, 104949 (2023).
9. Sukanya, S. & Jerine, S. Skin lesion analysis towards melanoma detection using optimized deep learning network. *Multimedia Tools and Applications* **82**, 1–23 (2023).
10. Jones, O. et al. Artificial intelligence and machine learning algorithms for early detection of skin cancer in community and primary care settings: a systematic review. *The Lancet Digital Health* **4**, e466–e476 (2022).
11. Veeramani, N., Premaladha, J., Krishankumar, R. & Ravichandran, K. S. in *Hybrid and automated segmentation algorithm for malignant melanoma using chain codes and active contours* 119–129 (Elsevier, 2023).
12. Stătescu, L. *et al.* Catching cancer early: The importance of dermato-oncology screening. *Cancers* **15** (2023). https://www.mdpi.com/2072-6694/15/12/3066.
13. Celebi, M. E. et al. Guest editorial skin image analysis in the age of deep learning. *IEEE Journal of Biomedical and Health Informatics* **27**, 143–144 (2023).
14. Celebi, M. E. et al. Guest editorial: Image analysis in dermatology. *Medical Image Analysis* **79**, 102468–102468 (2022).
15. Li, H. et al. Srdiff: Single image super-resolution with diffusion probabilistic models. *Neurocomputing* **479**, 47–59 (2022).
16. Ren, F., Liu, Y., Zhang, X. & Li, Q. Reversible information hiding scheme based on interpolation and histogram shift for medical images. *Multimedia Tools and Applications* **82**, 1–27 (2023).
17. Evangelista, D., Morotti, E. & Piccolomini, E. L. Rising: A new framework for model-based few-view ct image reconstruction with deep learning. *Computerized Medical Imaging and Graphics* **103**, 102156 (2023).
18. Johnson, P. M. et al. Deep learning reconstruction enables prospectively accelerated clinical knee mri. *Radiology* **307**, e220425 (2023).
19. Le, P.-H., Le, Q., Nguyen, R. & Hua, B.-S. *Single-image hdr reconstruction by multi-exposure generation*, 4063–4072 (2023).
20. Dong, C., Loy, C. C., He, K. & Tang, X. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence* **38**, 295–307 (2015).
21. Zhou, C. & Xiong, A. Fast image super-resolution using particle swarm optimization-based convolutional neural networks. *Sensors* **23**, 1923 (2023).
22. Tolba, B., El-Malek, A. H. A., Abo-Zahhad, M. & Elsabrouty, M. Meta-transfer learning for super-resolution channel estimation. *Journal of Ambient Intelligence and Humanized Computing* **14**, 2993–3001 (2023).
23. Zhang, Y., Huang, Y., Wang, K., Qi, G. & Zhu, J. Single image super-resolution reconstruction with preservation of structure and texture details. *Mathematics* **11**, 216 (2023).
24. Lee, S., Lieu, Q. X., Vo, T. P., Kang, J. & Lee, J. Topology optimization using super-resolution image reconstruction methods. *Advances in Engineering Software* **177**, 103413 (2023).
25. Gajera, B. et al. Ct-scan denoising using a charbonnier loss generative adversarial network. *IEEE Access* **9**, 84093–84109 (2021).
26. Gendy, G., Sabor, N., Hou, J. & He, G. *Real-time channel mixing net for mobile image super-resolution*, 573–590 (Springer, 2022).
27. Watson, J. *et al. Heightfields for efficient scene reconstruction for ar*, 5850–5860 (2023).
28. Ledig, C. *et al. Photo-realistic single image super-resolution using a generative adversarial network*, 4681–4690 (2017).
29. You, A., Kim, J. K., Ryu, I. H. & Yoo, T. K. Application of generative adversarial networks (gan) for ophthalmology image domains: a survey. *Eye and Vision* **9**, 1–19 (2022).
30. VARGAS, K. I. R. *UR-SRGAN: a generative adversarial network for real-world super-resolution with a U-Net-based discriminator*. Master's thesis, Universidade Federal de Pernambuco (2022).
31. Wang, X. *et al. Esrgan: Enhanced super-resolution generative adversarial networks*, 0–0 (2018).
32. Yu, Z. et al. Retinal image synthesis from multiple-landmarks input with generative adversarial networks. *Biomedical engineering online* **18**, 1–15 (2019).
33. Zhou, Y., Wang, B., He, X., Cui, S. & Shao, L. Dr-gan: conditional generative adversarial network for fine-grained lesion synthesis on diabetic retinopathy images. *IEEE Journal of Biomedical and Health Informatics* **26**, 56–66 (2020).
34. Kong, G., Fan, H. & Lobaccaro, G. Automatic building outline extraction from als point cloud data using generative adversarial network. *Geocarto International* **37**, 15964–15981 (2022).
35. Mahapatra, D., Bozorgtabar, B. & Shao, L. *Pathological retinal region segmentation from oct images using geometric relation based augmentation*, 9611–9620 (2020).
36. Goceri, E. *Image augmentation for deep learning based lesion classification from skin images*, 144–148 (IEEE, 2020).
37. Goceri, E. Comparison of the impacts of dermoscopy image augmentation methods on skin cancer classification and a new augmentation method with wavelet packets. *International Journal of Imaging Systems and Technology* **33**, 1727–1744 (2023).
38. Goceri, E. Gan based augmentation using a hybrid loss function for dermoscopy images. *Artificial Intelligence Review* **57**, 234 (2024).
39. Huang, S., Wang, J., Yang, Y., Wan, W. & Li, G. Lbcrn: lightweight bidirectional correction residual network for image super-resolution. *Multidimensional Systems and Signal Processing* **34**, 341–364 (2023).
40. Taud, H. & Mas, J.-F. Multilayer perceptron (mlp). *Geomatic approaches for modeling land change scenarios* 451–455 (2018).
41. Gong, C., Chen, X., Mughal, B. & Wang, S. Addictive brain-network identification by spatial attention recurrent network with feature selection. *Brain Informatics* **10**, 1–11 (2023).
42. Goceri, E. Polyp segmentation using a hybrid vision transformer and a hybrid loss function. *Journal of Imaging Informatics in Medicine* **37**, 851–863 (2024).
43. Deng, K., Wen, Y., Li, K. & Zhang, J. Hybrid model of tensor sparse representation and total variation regularization for image denoising. *Signal Processing* **217**, 109352 (2024).

44. Göçeri, E. An application for automated diagnosis of facial dermatological diseases. *İzmir Katip Çelebi Üniversitesi Sağlık Bilimleri Fakültesi Dergisi* **6**, 91–99 (2021).
45. Yin, J., Xu, S.-H., Du, Y.-B. & Jia, R.-S. Super resolution reconstruction of ct images based on multi-scale attention mechanism. *Multimedia Tools and Applications* **82**, 1–17 (2023).
46. Gaur, D., Folz, J. & Dengel, A. Training deep neural networks without batch normalization. arXiv preprint arXiv:2008.07970 (2020).
47. Dureja, A. & Pahwa, P. Analysis of non-linear activation functions for classification tasks using convolutional neural networks. *Recent Patents on Computer Science* **12**, 156–161 (2019).
48. Ramachandran, P., Zoph, B. & Le, Q. V. Searching for activation functions. arXiv preprint arXiv:1710.05941 (2017).
49. Codella, N. *et al.* Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). arXiv preprint arXiv:1902.03368**1** (2019).
50. Mendonça, T., Ferreira, P. M., Marques, J. S., Marcal, A. R. & Rozeira, J. *Ph 2-a dermoscopic image database for research and benchmarking*, 5437–5440 (IEEE, 2013).
51. Giotis, I. et al. Med-node: A computer-assisted melanoma diagnosis system using non-dermoscopic images. *Expert systems with applications* **42**, 6578–6585 (2015).
52. Rotemberg, V. et al. A patient-centric dataset of images and metadata for identifying melanomas using clinical context. *Scientific data* **8**, 34 (2021).
53. Qiu, D., Cheng, Y. & Wang, X. Improved generative adversarial network for retinal image super-resolution. *Computer Methods and Programs in Biomedicine* **225**, 106995 (2022).
54. Reznik, Y. Another look at ssim image quality metric. *Electronic Imaging* **35**, 305–1 (2023).
55. Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing* **13**, 600–612 (2004).
56. Shao, D., Qin, L., Xiang, Y., Ma, L. & Xu, H. Medical image blind super-resolution based on improved degradation process. *IET Image Processing* **17**, 1615–1625 (2023).
57. Creswell, A. et al. Generative adversarial networks: An overview. *IEEE signal processing magazine* **35**, 53–65 (2018).
58. Zhao, X., Zhang, Y., Zhang, T. & Zou, X. Channel splitting network for single mr image super-resolution. *IEEE transactions on image processing* **28**, 5649–5662 (2019).
59. Qiu, D., Cheng, Y. & Wang, X. Dual u-net residual networks for cardiac magnetic resonance images super-resolution. *Computer Methods and Programs in Biomedicine* **218**, 106707 (2022).
60. Jaworek-Korjakowska, J. *et al. Skin_hair dataset: Setting the benchmark for effective hair inpainting methods for improving the image quality of dermoscopic images*, 167–184 (Springer, 2022).
61. Sen, A. P., Pradhan, T., Rout, N. K. & Kumar, A. Comparison of algorithms for the removal of impulsive noise from an image. *e-Prime-Advances in Electrical Engineering, Electronics and Energy* **3**, 100110 (2023).
62. Jayatilake, S. & Ganegoda, G. *Melanoma skin cancer detection from dermoscopic images using computer vision*, 1–6 (IEEE, 2022).
63. Pahwa, E., Luthra, A. & Narang, P. Lvrnet: Lightweight image restoration for aerial images under low visibility. arXiv preprint arXiv:2301.05434**1**, 167–184 (2023).
64. Jayaraman, P. et al. Wavelet-based classification of enhanced melanoma skin lesions through deep neural architectures. *Information* **13**, 583 (2022).

Acknowledgements
Our earnest thanks to SASTRA Deemed University for providing all the facilities at the Computer Vision and Soft Computing Laboratory to proceed with the research.

Author contributions
The contributions for this research article are as follows: formal analysis, Conceptualization, methodology, algorithm implementation, draft preparation, done by N.V. Validation, formal analysis, visualization, supervision, and project administration by P.J. All authors have read and agreed to the published version of the manuscript.

Funding
The author(s) received no financial support for this article's research, authorship, and publication.

Declarations

Ethical approval
No ethical approval is required for this research endeavour.

Competing interests
The authors declare no conflicts of interest.

Additional information
Correspondence and requests for materials should be addressed to P.J.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.