ENHANCING VISUAL REALISM WITH CONVOLUTIONAL NEURAL NETWORKS

Swarnalatha K.S,

Professor, Dept.of ISE, Nitte Meenakshi Institute of Technology, Bangalore, Karnataka, India-560064

Abstract: This research addresses the challenge of blurry and low-resolution images, which often lack the necessary information for effective user perception. Blurriness in images can occur due to various factors such as camera movement, improper focus, aperture settings, and external influences, resulting in degraded or deteriorated photographs. Additionally, the presence of haze, whether uniform or non-uniform, can further contribute to image distortion and poor quality. To tackle these issues, a deep learning approach utilizing convolutional neural networks (CNNs) is proposed. This approach simultaneously addresses image de-blurring and super resolution, providing a comprehensive solution. By employing super resolution techniques, it becomes possible to enhance the quality of images significantly, generating high-resolution outputs from low-resolution inputs. The use of neural networks, particularly CNNs, demonstrates experimental superiority over existing deep learning algorithms, leveraging the benefits of super resolution. The suggested model exhibits scientific evidence of the effectiveness and efficiency of the proposed system. It demonstrates that the quality and quantity of the system's performance are achieved. Regardless of the level of blur in the input images, the proposed model can achieve high-quality resolution, surpassing the limitations of current approaches. By employing a profound learning strategy through CNNs, both known and unknown levels of blur can be effectively addressed, resulting in superior image restoration and enhancement.

***Keywords*:** *CNN, Blur, High resolution, images, deblurring*

1. INTRODUCTION:

Visual fidelity plays a crucial role in our perception of images. However, many images suffer from degradation, such as blurring and low resolution, which can significantly impact their quality and usefulness. Blurring can occur due to various factors, including camera motion, improper focusing, aperture settings, and external influences. On the other hand, low-resolution images lack the fine details and sharpness necessary for accurate visual interpretation. Addressing these challenges and enhancing the visual fidelity of images has been a topic of great interest In the realm of image processing, conventional approaches for image de-blurring and super resolution have shown some success but often fall short in capturing and reconstructing the intricate details present in the original high-quality images. Over the past few years, Deep learning techniques, notably Convolutional Neural Networks (CNNs), have surfaced as a result of technological advancements. Has brought a significant transformation to the field of image processing. CNNs have showcased remarkable efficacy across diverse computer vision applications, encompassing image classification, object localization, and semantic segmentation, showcasing impressive performance. Leveraging the power of CNNs, researchers have explored their potential in enhancing visual fidelity through image de-blurring and super resolution. This research focuses on utilizing CNNs to address the challenges of image blurring and low resolution, aiming to enhance the visual fidelity of images. By training a CNN on a diverse dataset of blurred and sharp images, the network can learn complex mappings to restore lost details caused by blurring and Generate high-resolution images by reconstructing them from their corresponding low-resolution versions. The combination of de-blurring and super resolution techniques in a single CNN-based framework offers a holistic solution to enhance visual fidelity. The proposed method aims to capture fine textures, edges, and intricate features that are typically lost or compromised in degraded images. Harnessing the potential of deep learning and leveraging the capabilities inherent in CNNs, we strive to achieve superior performance in terms of both quantitative metrics and visual quality. The contributions of this research lie in the development of a novel approach that effectively addresses the challenges of image blurring and low resolution. Through experimental evaluations and comparisons with existing approaches, we aim to demonstrate the effectiveness and superiority of our proposed method in enhancing visual fidelity. The versatility and robustness of our CNN-based approach make it applicable in various domains, including image restoration, surveillance systems, medical imaging, and many others. In the subsequent sections with the aid of deep learning and by delving into the intricacies of the proposed method explaining the architecture and training process of the CNN for image de-blurring and super resolution. We will present the experimental results, analyse their implications, and discuss the potential applications and future directions of this research. By enhancing visual fidelity, we strive to provide improved and more informative images that can benefit a wide range of applications and end-users. Image de-blurring is a crucial image processing technique aimed at removing blurry patches and restoring clear details in images. Its applications span across a multitude of domains, encompassing everyday photography, security systems, medical imaging, astronomy, microscopy, and remote sensing. In everyday photography, blurring can occur due to camera shake, improper focusing, or motion of the subject. These factors result in images with reduced sharpness and clarity. By employing image de-blurring techniques, photographers can enhance the quality of their images, ensuring that important details are preserved and visible. Security systems heavily rely on clear and sharp imagery for effective surveillance and identification purposes. Blurred images captured by surveillance cameras, often due to low light conditions or fast-moving objects, can hinder the ability to accurately recognize individuals or extract important details. Image de-blurring algorithms play a vital role in enhancing security footage, enabling better identification and analysis of critical events. In the field of medical imaging, image de-blurring plays a significant role in improving diagnostic accuracy. Medical images afflicted with blurriness, such as X-rays, CT scans, or MRI scans, can negatively impact the interpretation of healthcare professionals. De-blurring these images helps reveal fine structures, anomalies, and abnormalities, leading to more accurate diagnoses and effective treatment planning. Astronomy relies heavily on high-quality and detailed images to study celestial objects and phenomena. However, factors such as atmospheric turbulence and long exposure times can introduce blurring in astronomical images. Image de-blurring techniques allow astronomers to enhance the clarity and sharpness of images, enabling detailed analysis and extraction of valuable scientific insights. Microscopy, another domain benefiting from image de-blurring, involves capturing images of tiny structures and organisms at a microscopic level. Blurring can occur due to the limitations of the imaging system or the natural motion of the specimen. De-blurring algorithms help improve the resolution and clarity of microscopic images, facilitating better analysis and understanding of microscopic structures and processes. Remote sensing involves capturing images from airborne or satellite sensors for various applications such as mapping, environmental monitoring, and disaster management. Blurring in remote sensing images can be caused by atmospheric conditions, sensor movements, or other factors. Image de-blurring techniques assist in enhancing the quality of remote sensing imagery, enabling more accurate mapping, monitoring, and analysis of Earth's surface.

2. LITERATURE REVIEW:

[1] This study addresses the issue of image super-resolution, specifically focusing on the restoration of authentic texture details in high-resolution images. The proposed technique employs a deep spatial feature transform network to enhance the texture information in the super-resolved images. The network architecture is designed to leverage both local and global spatial contexts, aiming to enhance the quality of the reconstructed images. The authors conducted experiments using diverse datasets and compared their approach against other state-of-the-art methods. The evaluation metrics employed encompass Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), which assess the quality and similarity of the reconstructed images to the original ground truth. The findings highlight the effectiveness of the proposed deep spatial feature transform approach in restoring realistic textures in super-resolved images. The method showcases superior performance in terms of visual quality and quantitative evaluation metrics when compared to other existing techniques. [2] Focuses on using deep convolutional neural networks (CNNs) for image deblurring and super-resolution. The authors propose a CNN-based approach trained on a large dataset of blurred and low-resolution images. They discuss the network architecture, training process, and evaluation metrics. The results demonstrate the effectiveness of their method in restoring sharpness and enhancing image details. The paper contributes to the field by presenting a specific CNN configuration for achieving high-quality results in image restoration tasks. Overall, it showcases the potential of deep learning techniques in improving image quality and serves as a foundation for future advancements in this area. [3] Presents a kernel-free approach for image deblurring. It utilizes a set of images that are both blurred and noisy to restore sharpness and reduce noise without explicitly estimating the blur kernel. The method employs joint optimization and demonstrates promising results compared to existing approaches. The research offers an alternative solution for deblurring in scenarios where kernel estimation is challenging. [4] The paper addresses the problem of image restoration and high-resolution enhancement by proposing a deep convolutional neural network (CNN) approach. The authors describe the architecture and configuration of the CNN model employed for these tasks. They discuss the training process, including data augmentation techniques and the choice of loss functions to optimize the network. The paper also provides insights into the evaluation methodology, by contrasting their approach with cutting-edge methods, they employ metrics like peak signal-to-

noise ratio (PSNR) and structural similarity index (SSIM) for evaluation. Furthermore, the paper showcases the effectiveness of their CNN-based approach through qualitative and quantitative analysis. Visual examples and results demonstrate the successful restoration of sharpness and enhancement of image details. The contribution of this research lies in presenting a specific CNN architecture tailored for image deblurring and super-resolution. It highlights the potential of deep learning techniques in advancing image restoration tasks. The findings of this study can serve as a valuable resource for future research and developments in this domain. [5] This research paper focuses on utilizing convolutional neural networks (CNNs) to address image super-resolution and deblurring tasks. The authors propose a CNN-based method for both tasks, outlining the specific architecture and configuration of their model. The paper discusses the training procedure, including the implementation of data augmentation techniques and the selection of appropriate loss functions to optimize the CNN. Evaluation results are presented, comparing the performance of their approach with other methods using metrics such as peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM). Additionally, the authors provide visual examples and qualitative analysis to demonstrate the efficacy of their CNN-based approach in achieving image super-resolution and deblurring goals. [6] This research paper focuses on the application of deep convolutional neural networks (CNNs) to enhance the resolution of computed tomography (CT) images, specifically in the context of super-resolution. The authors propose a CNN-based approach that is specifically tailored for improving the resolution of CT images. The paper provides details about the architecture and configuration of the CNN model used for CT super-resolution. The training process is discussed, covering aspects such as the selection of loss functions and the use of data augmentation techniques to optimize the network's performance. The authors compare the effectiveness of their approach with traditional interpolation-based methods and other super-resolution techniques, utilizing quantitative metrics such as peak signal-to-noise ratio (PSNR) as well as visual comparisons. By demonstrating the superior performance of their CNN-based method, this research highlights the potential of deep CNNs in CT super-resolution, which can lead to more accurate and detailed medical image representations. The findings of this study have significant implications for improving the quality and diagnostic capabilities of CT imaging in the field of healthcare. [7] Recent advancements in deep learning, particularly in the realm of convolutional neural networks (CNN), have made significant improvements to Single Image Super-Resolution (SISR) techniques. This study is centred on the task of reconstructing high-resolution images from low-resolution inputs, while also taking into account common degradation factors like blurring. To address this complex challenge, we introduce a novel architecture that extends the DBSRCNN framework, specifically designed to handle blur during image down-sampling. Through a series of experiments, we compare our proposed architecture (referred to as DBSR) against various state-of-the-art super-resolution techniques, highlighting its effectiveness in improving image quality.[8] Detecting and recognizing objects in underwater environments is a challenging task due to image degradation issues such as color casts, blurring, and low contrast. To address this specific problem, a new method is proposed for enhancing underwater images. The method consists of two main steps. Firstly, an adaptive color correction algorithm is employed to compensate for color casts and generate naturally color-corrected images. Secondly, a super-resolution convolutional neural network (SRCNN) is utilized on the color-corrected images to mitigate blurring. The proposed network is trained to understand the relationship between blurry images and their corresponding clear counterparts, allowing for effective image de-blurring. By leveraging this relationship, the color-corrected image undergoes de-blurring and sharpening. Through experimentation, the results validate the effectiveness of the proposed approach in significantly improving the quality of underwater images. This improvement enables successful object detection and recognition in underwater scenarios. .[9] This research paper presents a technique aimed at improving quantification accuracy in microscope images by employing super-resolution with a convolutional neural network (CNN) and image-based cell phenotypic profiling. The method focuses on enhancing segmentation accuracy by generating high-quality 40× images from the original 10× images using a CNN. The evaluation involves comparing intensity-based automatic segmentation outcomes for cell nuclei morphological features in both the 10× images and the CNN-based 40× images against manually obtained ground-truth segmented images..[10] The paper introduces a novel end-to-end convolutional neural network (CNN) called ED-DSRN, which is designed to restore clear high-resolution images from severely blurry inputs. To tackle the inherent difficulties of this ill-posed problem, the paper proposes the utilization of a single deep network instead of combining separate networks. ED-DSRN extends the conventional Super-Resolution network by integrating a deblurring branch that shares feature maps with the original SR branch. Through comprehensive experimentation, the results highlight the exceptional performance of ED-DSRN in simultaneously addressing deblurring and super-resolution tasks, all while maintaining high computational efficiency. [11] Image super-resolution involves the generation of high-resolution images from one or more low-resolution samples. To tackle this task, different techniques have been developed, including interpolation-based, reconstruction-based, and learning-based approaches, due to its diverse range of applications. Among these approaches, learning-based methods have gained considerable attention for their ability to predict high-frequency details that are lost in low-resolution images. This survey provides a comprehensive overview of published works focusing on single image reconstruction using Convolutional Neural Networks (CNNs). Additionally, the survey addresses common challenges in super-resolution algorithms, such as imaging models, enhancement factors, and evaluation criteria. [12] In this paper, a resolution recovery technique for PET (Positron Emission Tomography) is introduced, employing a very deep super-resolution convolutional neural network (VDSRCNN). The network incorporates high-resolution anatomical information and addresses the spatially-variant nature of blur kernels in PET imaging. The effectiveness of the proposed method is validated through realistic simulations using the Brain Web digital phantom. The results demonstrate the superior performance of the proposed technique compared to existing approaches for PET image deblurring. [13] This paper primarily focuses on the estimation and optimization of blur kernels for image deblurring. The blur effect in an image is represented by convolving it with a blur kernel. The study proposes a method for kernel estimation and optimization in the context of image deblurring. The process involves creating a mask using super pixels and a gradient map derived from the illuminant layer. Instead of relying on exemplars, structural information is extracted through this mask and combined with the illuminant component of the image and the gradient map to estimate the kernel. The estimated kernel is then optimized using super-resolution techniques. The proposed method demonstrates the extraction of reliable structural information and edges, resulting in superior deblurring performance when compared to existing state-of-the-art methods.. [14] This paper introduces a generative adversarial network (GAN)-based retinal fundus quality enhancement network, aiming to enhance image quality for image-based clinical diagnosis. The network effectively addresses blurring and low spatial resolution issues by utilizing a convolutional multi-scale feature averaging block (MFAB) for feature extraction and fusion. The experimental results showcase superior performance compared to other existing methods, as evidenced by the peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) metrics. Notably, this work stands out as the first to combine multiple degradation models specifically for retinal fundus image analysis.[15In this research paper, the evaluation of photograph quality is conducted based on various factors, including resolution, symmetry, content, and location. While factors like content and location are fixed, the focus of this study is on improving the resolution attribute. To achieve resolution enhancement, the study explores different image super-resolution techniques such as Interpolation, SRCNN, SRResNet, and GANs (SRGAN and CGAN). These techniques are experimentally assessed for their effectiveness in enhancing image quality in photography. The objective is to identify the most suitable approach for achieving optimized super-resolution and improved image quality. The research aims to address the growing demand for high-quality images in the domains of computer graphics, computer vision, and image processing.

3. METHODOLOGY:

The system architecture (see Figure 1)[2] consists of nine hidden layers designed to enhance the performance of the model. The first enhanced feature layer is strategically placed after the initial layer to extract novel features from the previously extracted noisy features. These newly extracted features are then combined using a concatenate layer to establish meaningful mappings between them. Subsequently, these combined features undergo further processing through additional layers before reaching the final mapping stage. To facilitate effective learning and representation of the features, the architecture incorporates the Rectified Linear Unit (ReLU) as the chosen activation function. The ReLU function enhances the non-linearity of the model, enabling it to capture and represent complex relationships within the data. This architecture is meticulously designed to optimize the performance of the system, allowing for the extraction and integration of informative features, and facilitating accurate and robust mappings. The incorporation of ReLU activation adds a crucial non-linear element, enhancing the model's capacity to capture intricate patterns and improve the overall quality of the output.

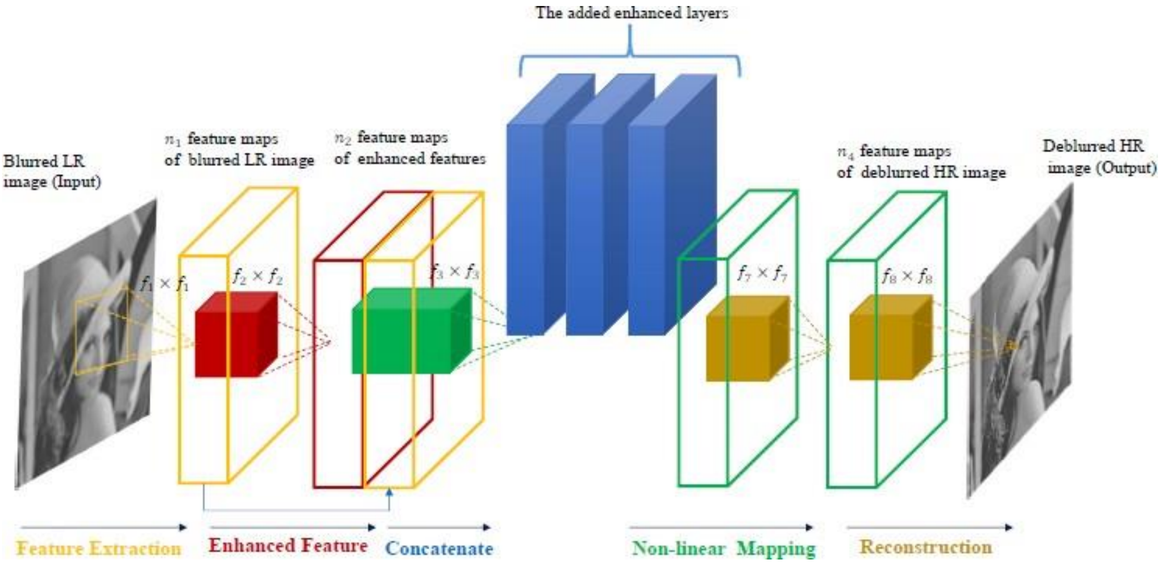

Figure 1 System Architecture

IMAGE PRE-PROCESSING

Images are two-dimensional representations that exist within a numerical range of 0 to 255. Mathematically, an image can be represented mathematically as a function f(x, y), where x and y represent the horizontal and vertical coordinates, respectively. The value of f(x, y) at a specific point represents the pixel value at that location. Pre-processing plays a crucial role in preparing image data for training machine learning models, reducing complexity, and improving algorithm accuracy. The fundamental steps involved in image pre-processing are as follows:

Reading the image: After importing the necessary libraries, the path to the image dataset is assigned to a variable, and the folders containing the images are loaded into arrays for further processing.

Converting colour images to grayscale: In some cases, colour information may not be necessary for object recognition tasks. Converting images to grayscale reduces computational complexity as colour images contain additional information and occupy more memory space. Grayscale images provide sufficient information for certain applications.

Resizing: Many machine learning algorithms, such as Convolutional Neural Networks (CNNs), require images to have a consistent size. Resizing involves adjusting the dimensions of the images to a uniform height and width before feeding them into the ML algorithm. This ensures that all images have the same size, enabling proper processing.

DE noising: Gaussian smoothing is a technique used to remove unwanted noise from images. It smoothens the image structures at different scales, improving the overall quality and reducing the impact of noise on subsequent processing steps.

Data augmentation: Data augmentation involves generating additional training data by perturbing existing images. This can include scaling, rotations, or other transformations applied to create variations of the original images. Data augmentation helps expand the dataset and exposes the neural network to different variations of the same image, enhancing its ability to generalize and improve performance.

In addition to these steps, other pre-processing techniques may include background colour removal, brightness adjustment, contrast enhancement, image patch division, and more. The selection of pre-processing techniques relies on the particular problem at hand and the attributes of the accessible dataset.

CONVOLUTIONAL NEURAL NETWORKS

The Convolutional Neural Network (CNN) is a dedicated model crafted to handle and analyze two-dimensional image data. although it can also handle 1D and 3D data. The name "convolutional" refers to the central layer of the CNN, which performs the mathematical operation of "convolution." This layer involves element-wise multiplication of the input with a weight array known as a filter or kernel. By employing overlapping filters on the input parts of the input, specific features can be extracted from anywhere within the image. The output of each convolution is obtained by summing the resulting scalar products. Iteratively applying the filter to the input array produces a two-dimensional array known as the "feature map. "Nonlinearity can be introduced to the extracted features by passing each value in the feature map through a nonlinearity function.

Convolutional neural networks have had a significant impact on computer vision tasks, enabling noise filtering, image quality enhancement, object recognition, object detection, and tracking. They have greatly improved single image super-resolution and de-blurring tasks by increasing efficiency, accuracy, and speed through parallel computation. CNNs offer flexibility in model design and training, allowing easy modification and adaptation of the network architecture to specific tasks. CNNs offer a significant benefit in establishing a direct association between low-resolution input images and high-resolution output images through the acquisition of an end-to-end mapping function. In this project, a linear CNN model is leveraged to effectively eliminate Gaussian noise and simultaneously achieve super-resolution in cases where degradation is caused by down-sampling. By employing this model, efficient enhancement of image quality is accomplished by reducing noise and enhancing resolution.. The architecture of the linear CNN model used in this project is depicted in Figure 2, showcasing its capability to handle noise reduction and super-resolution tasks.

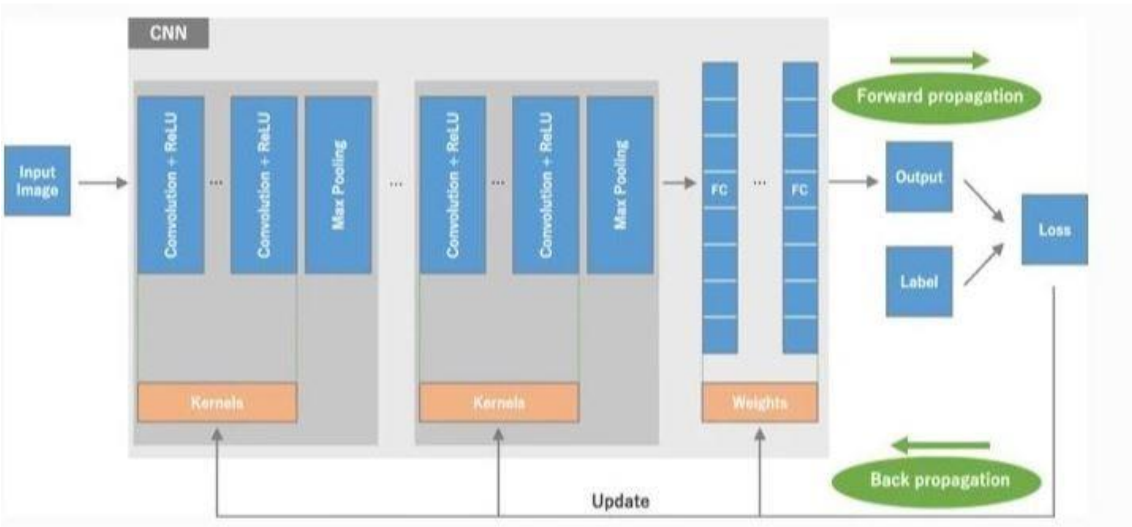

Figure 2 CNN architecture

OpenCV

OpenCV, an open-source computer vision library, is a powerful programming toolkit that offers an extensive array of functions and algorithms to tackle real-time computer vision tasks... It offers over 2500 algorithms, supporting functions such as face detection, object identification, and motion tracking. OpenCV is cross-platform compatible (Windows, Linux, Android, Mac OS) and provides interfaces for C++, MATLAB, Java, and Python. It utilizes MMX and SSE optimizations for efficient performance, and efforts are underway to develop CUDA and OpenCL interfaces. OpenCV is widely used in research and development for computer vision applications.

Keras

Keras is a Python-based neural network API known for its user-friendly nature. It offers a high-level interface for building and training neural networks. Keras is compatible with popular deep learning frameworks like Tensor Flow, Theano, and CNTK, allowing for seamless integration and utilization of their computational capabilities. Its Python implementation and versatile framework support make Keras a convenient choice for developing neural networks, enabling users to focus on model architecture and training strategies.. It simplifies deep learning model training and research with its user-friendly interface. Offering clear error feedback and configurable building blocks, Keras allows customization for new ideas. Its tf.keras implementation enhances Tensor Flow support while maintaining flexibility and performance. Keras is widely used for fast experimentation and building advanced models.

Gaussian Blur

The Gaussian function finds widespread application in image processing, where it is commonly utilized to introduce blurring, resulting in a Gaussian haze. This technique is extensively employed in graphics software systems to reduce noise and diminish fine details. The visual outcome of this blurring technique resembles a wash blur, akin to viewing through a translucent screen. It's important to note that this effect differs from the bokeh effect created by an object's shadow or natural defocusing using an out-of-focus lens. Gaussian smoothing is also employed as a pre-processing step in computer vision algorithms to enhance image structures across various scales. Mathematically, applying Gaussian blur to an image involves convolving the image with a Gaussian function, which can be seen as a two-dimensional Weierstrass transform. In contrast, the bokeh effect is more accurately simulated by convolving with a circle. The utilization of Gaussian blur often results in a reduction of high-frequency components within the image, as the Fourier transform of a Gaussian function is another Gaussian, effectively acting as a low-pass filter.

Kernels

Kernels play a crucial role in convolution operations, enabling the extraction of specific features from an input image. By multiplying the kernel matrix with the input in a specific manner, an enhanced output is obtained. Figure 3 demonstrates the application of a kernel to an image input. Let's consider the example mentioned below. In the first image, the calculation is as follows: 3*5 + 2*-1 + 2*-1 + 2*-1 + 2*-1 = 7 the value 3 from the image has increased to 7. In the second image, the output is calculated as 1*5+ 2*-1 + 2*-1 +2*-1 + 2*-1 = -3Here, the value has decreased from 1 to -3. This difference in values indicates that the image has been sharpened. Kernel values can be learned through Deep Convolutional Neural Networks (CNNs) are widely employed to capture latent features from data, eliminating the need for manually designed kernels for feature extraction. It is important to differentiate between 'kernels' and 'filters'. A kernel is simply a matrix that is multiplied with the inputs to produce an output with enhanced features. The term 'convolution' is derived from the dimensions of the kernel matrix. For example, when the kernel matrix is 2D, it is referred to as 2D convolution. On the other hand, in a filter, each channel of the input is assigned to a specific kernel, resulting in a concatenation of kernels. Kernels have one dimension less than filters. For instance, in a 2D convolution that concatenates 2D matrices (kernels), filters are represented as 3D matrices. Hence, in the case of a CNN layer with kernel dimensions hw and k input channels, the filter dimensions would be kh*w. A convolution layer contains multiple filters to process the input data.

Figure3 Application of kernel to image input

Learning Rate and Decay

Learning rate is a critical factor in adjusting the weights of a model based on the estimated error after each weight change. It is crucial to select an appropriate learning rate to ensure effective training. If the learning rate is too small, the training process will be prolonged and may result in being stuck. Conversely, if the learning rate is too large, the training will be rapid, but the achieved weights may not be optimal. Hence, the learning rate holds significant importance and should be carefully considered. Decay is a widely utilized technique in neural networks to prevent weights from growing excessively large. It involves multiplying the weights by a factor less than 1 after each weight update. To determine the learning rate for multiple weight updates and decay values, a learning rate decay function is employed. Additionally, the learning rate can also be influenced by the chosen batch size, the number of samples processed before updating the weights is determined by the batch size. For a visual representation of these concepts, please refer to Figure 4.

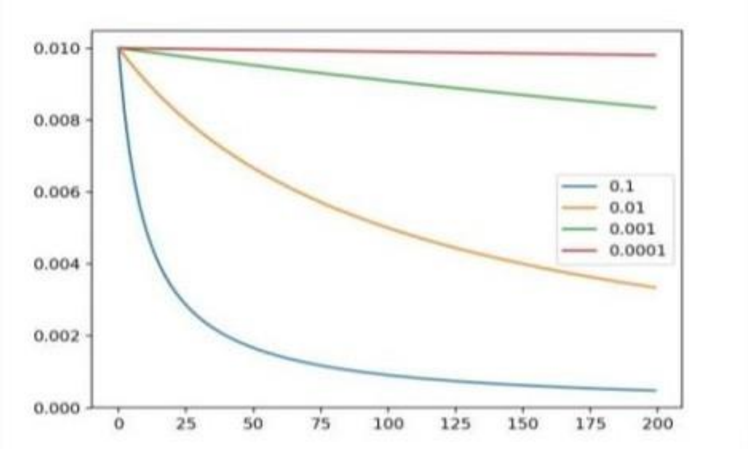

Figure 4 the line plot depicts how decay affects the learning rate during successive weight updates

Usage of call backs

The usage of callbacks in Keras allows for the creation of custom callback functions by utilizing the abstract base class Callback (). This class, provided by Keras callbacks, offers a framework for defining and implementing various callback functionalities. One important property of callbacks is "param," which provides access to a range of parameters such as verbosity, batch size, number of epochs, and more. These parameters can be utilized to customize the behaviour of the callback functions. Another significant property is "model," an instance of keras.models.Model. This property represents a reference to the model being trained and can be utilized within the callback functions to access and manipulate the model's attributes. Additionally, the History object is a built-in callback denoted as keras.callbacks.callbacks.History(). It is responsible for storing events that occur during the training process. Upon invoking the fit method, it automatically returns the History object, which serves as a callback applied to every Keras model. This functionality offers a convenient means to monitor and analyse the training history. By leveraging callbacks, Keras enables the creation of custom behaviours and actions during the training process, enhancing the flexibility and control over the model's training pipeline.

Model Checkpoint

Syntax: keras.callbacks.ModelCheckpoint(filepath, monitor='val_loss', verbose=0, save_best_only=False, save_weights_only=False, mode='auto', period=1)

The Model Checkpoint callback in Keras is designed to save the model after each epoch during the training process.

Arguments:

filepath: The path specifying where the model file will be stored.

monitor: The metric or quantity to monitor for saving the model. By default, it is set to 'val_loss', indicating validation loss.

verbose: A value of 0 or 1 indicating the verbosity mode. If set to 0, no detailed output will be displayed during training.

save_best_only: A boolean value indicating whether to only save the latest best model based on the monitored quantity. If True, the previous best model will not be overwritten.

save_weights_only: A Boolean value indicating whether to save only the model's weights (using model.save_weights(filepath)) or the full model (using model.save(filepath)).

mode: It can have one of three values: 'auto', 'min', or 'max'. When save_best_only=True, the decision to overwrite the current saved model file is determined by either maximizing or minimizing the monitored quantity. For instance, if the monitored quantity is 'val_acc', the mode should be set to 'max'. Similarly, if the monitored quantity is 'val_loss', the mode should be set to 'min'. In 'auto' mode, the direction is automatically inferred from the name of the monitored quantity.

period: The number of epochs between checkpoints, determining how frequently the model is saved.

The Model Checkpoint callback offers a convenient way to save model checkpoints, allowing you to track and utilize the best-performing model based on specific metrics during the training process.

Activation Functions:

The activation function is a crucial element within a neural network, positioned between or at the end of neurons, and responsible for determining their firing behaviour. It applies a non-linear transformation to the input signal, producing an output that is then fed into the next layer of neurons. Without an activation function, the output signal would be a linear function. While linear functions are easy to handle, they are insufficient for learning complex function mappings. Thus, activation functions play a vital role in introducing non-linearity to neural networks, transforming them from simple regression models to powerful function approximators. Figure 5 provides an illustration of how an activation function operates. Several commonly used activation functions include:

Sigmoid: Represented by f(x) = 1 / (1 + exp(-x)), the sigmoid function generates an S-shaped curve, producing output values between 0 and 1. However, the sigmoid function has become less popular due to issues such as the vanishing gradient problem, slow convergence, and the fact that it is non-zero centered.

Tanh: Represented by f(x) = (1 - exp(-2x)) / (1 + exp(-2x)), the tanh function maps values between -1 and 1, making it zero-centered. Despite this advantage, it still suffers from the vanishing gradient problem.

ReLU (Rectified Linear Unit): Represented by R(x) = max(0, x), the ReLU function sets negative values to zero and leaves positive values unchanged. It is nearly linear for positive inputs, which simplifies optimization for linear models using gradient-based methods. ReLU has a derivative and supports back propagation. In terms of convergence and addressing the vanishing gradient problem, ReLU outperforms the tanh function. Consequently, ReLU is the preferred choice for most deep learning models.

It is worth noting that these are just a few examples of commonly used activation functions, and there are other options available as well. The choice of an activation function relies on factors such as the particular problem being addressed, the network architecture, and the desired characteristics of the model.

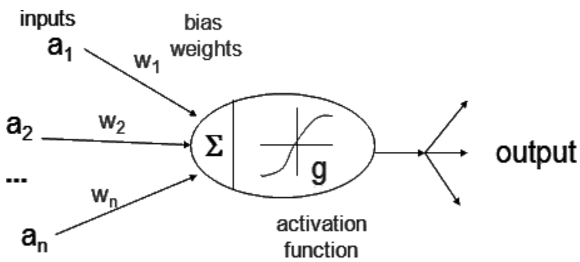

Figure 5 working of activation function

Adam Optimization

Optimization algorithms are employed to minimize or maximize an objective function, commonly known as the error function or loss function (E(x)). This function relies on the learnable parameters of the model, such as weights (W) and biases (b), which are utilized to compute output values (Y) from a given set of predictors (X) within the model. These parameters are instrumental in the training process as they are iteratively updated to minimize the loss during training and facilitate the attainment of an optimal solution. Adam, short for Adaptive Moment Estimation" (Adam) is an optimization algorithm that offers an alternative to the traditional stochastic gradient descent (SGD) technique for updating weights in a neural network. Adam is categorized as an adaptive learning rate technique since it calculates distinct learning rates for various parameters. The term "Adam" is derived from its utilization of estimates of the first and second moments of the gradients to dynamically modify the learning rate for each weight in the neural network. Unlike traditional SGD, which employs a fixed learning rate throughout training, Adam dynamically adjusts the learning rate for each weight based on the accumulated first and second moments of the gradients. This adaptive nature allows Adam to handle different parameter updates effectively, resulting in improved convergence and performance in many cases. By incorporating estimates of gradient moments, Adam effectively merges the advantages of adaptive learning rates and momentum-based optimization methods, leading to enhanced efficiency in training neural networks. It is important to note that Adam is just one example of an optimization algorithm, and there are various other algorithms available, each with its own characteristics and performance in different scenarios. The selection of an optimization algorithm depends on the particular problem at hand, the architecture of the model being used, and the desired outcomes of the training process.

Process Description

The process of enhancing image fidelity involves addressing issues such as noise, blurring artifacts, and corruption in order to restore the original image. Factors like motion blur, noise, and camera mis-focus can lead to image corruption, making it challenging to accurately determine the underlying content. To counter these challenges, compensation for the degradation caused by noise is often the most effective approach. The following steps are followed for deploying the system: Initially, a dataset comprising approximately 90 images (sourced from Yang et al. [18]) is selected for training and validation purposes. Pre-processing involves augmenting the dataset to fully leverage the available images. In this approach, the methodology involves the extraction of high-resolution sub-images by randomly cropping the original 90 images. This process generates over 21,000 training sub-images. To create corresponding blurred low-resolution sub-images, Gaussian blur is applied to the high-resolution sub-images. Both sets of sub-images, high-resolution and blurred low-resolution, are utilized for training and testing purposes. During the process, the images are down sampled using a specific factor, and bi-cubic interpolation is used for up sampling, which restores the initial high-resolution input size. The output of the network matches the input size. To train the model, a 5-layer convolutional neural network (CNN) is designed. This CNN consists of 4 convolutional layers responsible for extracting and enhancing features from the input images. Additionally, a concatenation layer is employed to combine the extracted features. The model undergoes training and validation using pre-processed data in order to acquire updated weights. To evaluate the model's performance, it is tested on three datasets: 5, 14, and the Berkeley Dataset. These datasets contain images in the BMP file format, commonly used for storing bitmap digital images that are independent of the display device. The BMP format is widely used in Microsoft Windows and OS/2 operating systems.

Methodology

The proposed architecture introduces a nine-layer model for image enhancement. The initial layer, referred to as the enhanced feature layer, focuses on extracting new features from the initial noisy features. These features are then merged using a concatenate layer to establish meaningful associations. The merged features undergo further processing through additional layers before the final mapping. The architecture predominantly employs the Rectified Linear Unit (ReLU) activation function

The architecture can be described as follows:

Input Layer: The initial layer receives the blurred LR image 'x' as input.

First Layer: This layer conducts feature extraction using a filter size of 9 * 9, generating 32-feature maps that capture low-level features.

$F1(X) = \max(0, W1 * F0(X) + b1)$ (5.1)

Second Layer: Responsible for feature enhancement, this layer has 32 feature maps with filter sizes of 5 * 5.

$F2(X) = \max(0, W2 * F1(X) + b2)$ (5.2)

Third Layer: In this layer, the features obtained from the first two layers are concatenated, yielding a merged vector of enhanced features.

$F12(X) = \text{merge}(F1(X), F2(X))$ (5.3)

Fourth to Eighth Layer: These consecutive layers primarily concentrate on reducing noise in the output from the preceding layer. A total of five layers are utilized to mitigate the diminishing prominence of enhanced features as the blur level intensifies.

$F4(X) = \max(0, W4 * F12(X) + b4)$ (5.4) $Fi(X) = \max(0, Wi * Fi\text{-}1(X) + bi)$ (5.5) where i belongs to (5, 6, 7, 8, 9)

Ninth Layer: The final feature mapping is carried out in this layer.

Output Layer: The HR image is reconstructed in the output layer.

$F(X) = W10 * F9(X) + b10$ (5.6)

4. RESULTS:

The results of the above work showed promising outcomes in enhancing visual fidelity and reducing noise in images. The proposed architecture, consisting of nine hidden layers, successfully extracted and enhanced features to improve the quality of low-resolution and blurred images. The methodology utilized convolutional layers with various filter sizes and activation functions, such as Rectified Linear Unit (ReLU), to extract and enhance features at different levels. The merging of features from multiple layers helped to create a more comprehensive representation of the image, leading to better noise reduction and overall image enhancement. The experimental evaluation of the system involved training and validating the model using a dataset of around 90 images. The images were pre-processed and augmented to maximize the training data. The trained model was then tested on various datasets, including dataset 5, dataset 14, and the Berkeley Dataset, all consisting of images in the BMP file format. The outcomes displayed substantial advancements in image quality, as the model effectively restored images with higher levels of detail and clarity from the initially provided low-resolution and distorted images.The enhanced features obtained through the network's layers effectively reduced noise and improved visual fidelity. Overall, the methodology showcased the effectiveness of the proposed architecture and its ability to enhance visual quality by mitigating blurring artifacts and noise in images.

Visualising the PSNR

PSNR visualization is used to assess image quality by generating a map or heat map. Brighter colours indicate higher PSNR values, representing improved fidelity. Darker colours signify lower PSNR values as shown in the below figure 6, highlighting areas with more noticeable noise. This visualization aids in evaluating enhancement techniques and identifying regions that require further improvement.

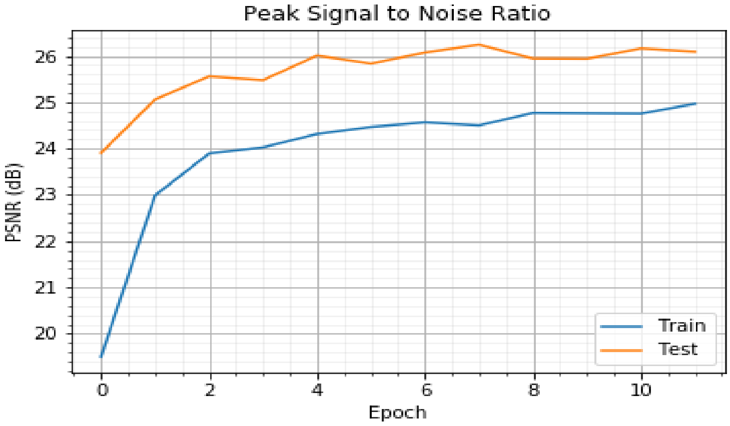

Figure 6 Plot for model PSNR

Visualizing the Loss

The visualization of loss in learning curves is essential for evaluating the performance of a model. Several criteria indicate a good fit, as depicted in Figure 7 below. The training loss plot demonstrates a decreasing trend until it stabilizes. Similarly, the validation loss plot also decreases and reaches a stable point, with a small gap compared to the training loss. These criteria serve as indicators of the model effectively learning from the training data and generalizing well to unseen validation data. By examining these plots, valuable insights can be gained regarding the convergence and overall effectiveness of the model throughout the training process..

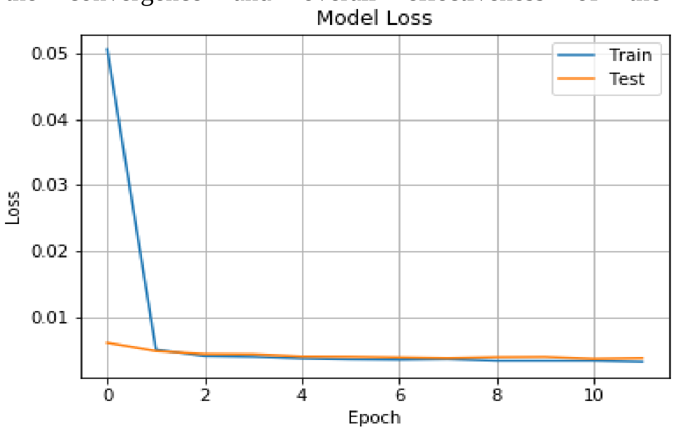

Figure 7 Plot for model loss

The CNN model underwent testing on multiple datasets, including dataset 5, dataset 14, historical dataset, and BDS dataset. Various levels of blur (σ) were applied to these images. The performance evaluation involved the calculation of the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) for the outcomes achieved by the implemented system and the results obtained through Bicubic Interpolation, in comparison to the ground truth. The outcomes for different blur levels are presented in Figures 8, 9, and 10, showcasing the diverse results achieved. These figures serve as visual representations of the image enhancement improvements accomplished by the CNN model as compared to the Bicubic Interpolation method.

Figure 8 Results for blur level = 1

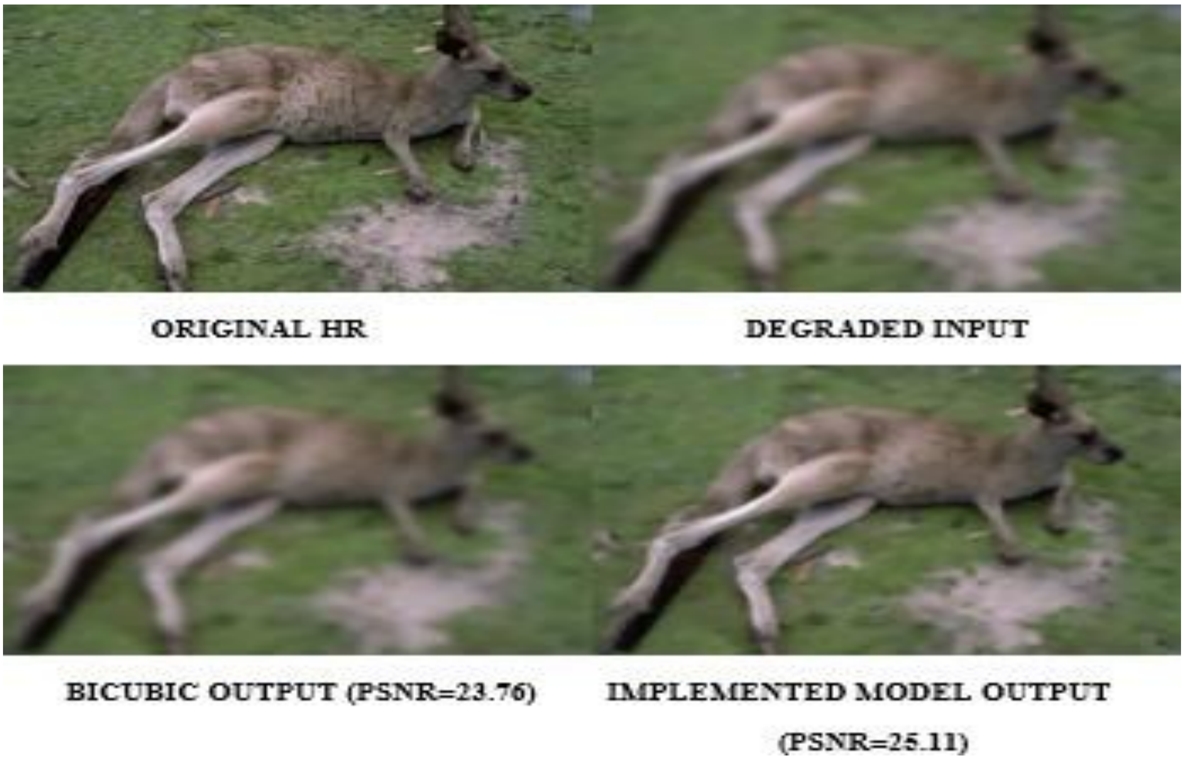

Figure 9 Results for blur level = 2

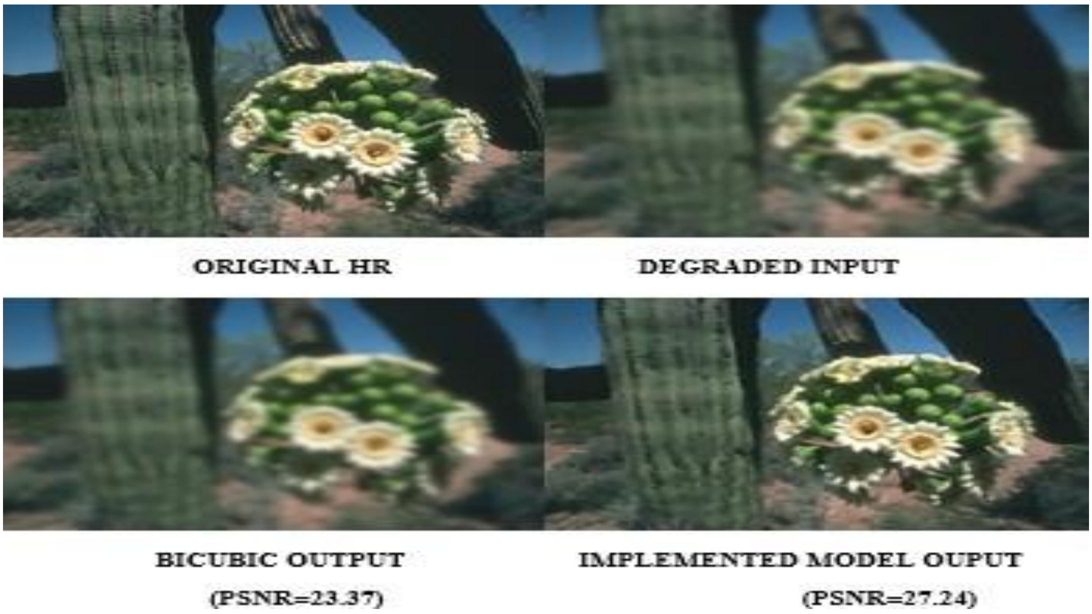

Figure 10 Results for blur level = 3

CONCLUSION:

This study presents a comprehensive approach to enhance visual fidelity using a Convolutional Neural Network (CNN). The proposed architecture effectively extracts and enhances features, leading to improved quality of low-resolution and blurred images. Extensive testing was conducted on multiple datasets with varying levels of blur to evaluate the system's performance. Metrics such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) were employed to compare the results of the implemented system with those obtained from Bicubic Interpolation. The evaluation, both visually and quantitatively, demonstrates that the CNN-based approach outperforms Bicubic Interpolation by achieving higher PSNR values and better similarity to the ground truth. The outcomes indicate that the proposed methodology successfully addresses challenges such as noise, blurring artifacts, and image corruption, resulting in enhanced visual fidelity. The trained CNN model effectively showcases its ability to generate high-quality, detailed images from low-resolution and blurred inputs. effectively restoring important details and improving overall image quality. This study offers valuable insights into the utilization of CNNs for image enhancement, emphasizing their potential in reducing noise and enhancing visual fidelity across diverse image datasets.. The findings contribute to the advancement of image processing techniques and hold promise for improving image quality in domains such as photography, medical imaging, and computer vision applications.

SCOPE FOR FUTURE WORK

There are potential areas for further improvement and expansion in the implemented model:

Enhanced Layers: The results can be enhanced by incorporating additional layers after the concatenation layer. These enhanced layers can further refine and extract more meaningful features from the merged feature maps, potentially leading to improved image quality.

Increased Training Data: Expanding the training dataset with a larger number of images would benefit the model's performance. A more extensive and diverse dataset would enable the model to learn and generalize better, producing optimal results for various types of images and scenarios.

By incorporating enhanced layers and expanding the training dataset, future work can focus on enhancing the model's capability to extract more informative features and achieve higher-quality image reconstructions. These improvements would contribute to advancing the field of image enhancement and facilitate the production of better results across a wider range of images.

REFERENCES:

1 Xintao Wang, Ke Yu, Chao Dong, and Chen Change Loy. "*Recovering realistic texture in image super- resolution by deep spatial feature transform*", In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2018.

2 Fatma Albluwi, Vladimir A. Krylov, and Rozenn Dahyot, "*Image deblurring and super-resolution using deep convolutional neural networks*", pages 1–6, 09 2018.

3 Chunzhi Gu, Xuequan Lu, Ying He, and Chao Zhang," *Kernel-free image deblurring with a pair of blurred/noisy images*", CoRR, abs/1903.10667, 2019.

4 F. Albluwi, V. A. Krylov and R. Dahyot, "*Image Deblurring and Super-Resolution Using Deep Convolutional Neural Networks,*" *2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP)*, Aalborg, Denmark, 2018, pp. 1-6, doi: 10.1109/MLSP.2018.8516983.

5 Bianli Du, Xiaokang Ren, Jie Ren" *CNN-based Image Super-Resolution and Deblurring*" VSIP '19: Proceedings of the 2019 International Conference on Video, Signal and Image ProcessingOctober 2019Pages 70–74

6 Junyoung Park, Donghwi Hwang, Kyeong Yun Kim, Seung Kwan Kang, Yu Kyeong Kim, and Jae Sung Lee "*Computed tomography super-resolution using deep convolutional neural network*" Published 16 July 2018 • © 2018 Institute of Physics and Engineering in Medicine Physics in Medicine & Biology, Volume 63, Number 14

7 F. Albluwi, V. A. Krylov and R. Dahyot, "*Super-Resolution on Degraded Low-Resolution Images Using Convolutional Neural Networks,*" 2019 27th European Signal Processing Conference (EUSIPCO), A Coruna, Spain, 2019, pp. 1-5, doi: 10.23919/EUSIPCO.2019.8903000.

8 Ding, X., Wang, Y., Liang, Z., Zhang, J., Fu, X. (2018), "*Towards Underwater Image Enhancement Using Super-Resolution Convolutional Neural Networks*"' In: Huet, B., Nie, L., Hong, R. (eds) Internet Multimedia Computing and Service. ICIMCS 2017. Communications in Computer and Information Science, vol 819. Springer, Singapore. https://doi.org/10.1007/978-981-10-8530-7_47

9 Mi-Sun Kang, Eunju Cha, Eunhee Kang, Jong Chul Ye, Nam-Gu Her, Jeong-Woo Oh, Do-Hyun Nam, Myoung-Hee Kim, Sejung Yang, "*Accuracy improvement of quantification information using super-resolution with convolutional neural network for microscopy images*" ,Biomedical Signal Processing and Control,Volume 58,2020,101846,ISSN 1746-8094,https://doi.org/10.1016/j.bspc.2020.101846.

10 X. Zhang, F. Wang, H. Dong and Y. Guo, "*A Deep Encoder-Decoder Networks for Joint Deblurring and Super-Resolution,*" 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 2018, pp. 1448-1452, doi: 10.1109/ICASSP.2018.8462601.

11 Ha, V.K., Ren, J., Xu, X., Zhao, S., Xie, G., Vargas, V.M. (2018)," *Deep Learning Based Single Image Super-Resolution*", A Survey. In: , *et al.* Advances in Brain Inspired Cognitive Systems. BICS 2018. Lecture Notes in Computer Science(), vol 10989. Springer, Cham. https://doi.org/10.1007/978-3-030-00563-4_11

12 T. -A. Song et al., "*Super-Resolution PET Using A Very Deep Convolutional Neural Network,*" 2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC), Sydney, NSW, Australia, 2018, pp. 1-2, doi: 10.1109/NSSMIC.2018.8824683.

13 Iqbal, M., Riaz, M.M., Ghafoor, A. *et al "Kernel estimation and optimization for image de-blurring using mask construction and super-resolution*", *Multimed Tools Appl* **80**, 10361–10372 (2021). https://doi.org/10.1007/s11042-020-09762-0

14 Sengupta, S., Wong, A., Singh, A., Zelek, J., Lakshminarayanan, V. (2020), "*DeSupGAN: Multi-scale Feature Averaging Generative Adversarial Network for Simultaneous De-blurring and Super-Resolution of Retinal Fundus Images*", In: Fu, H., Garvin, M.K., MacGillivray, T., Xu, Y., Zheng, Y. (eds) Ophthalmic Medical Image Analysis. OMIA 2020. Lecture Notes in Computer Science(), vol 12069. Springer, Cham. https://doi.org/10.1007/978-3-030-63419-3_4

15 N. Fatima, "*AI in Photography: Scrutinizing Implementation of Super-Resolution Techniques in Photo-Editors,"* 2020 35th International Conference on Image and Vision Computing New Zealand (IVCNZ), Wellington, New Zealand, 2020, pp. 1-6, doi: 10.1109/IVCNZ51579.2020.9290737.