Image Deblurring Using Deep Learning

Jeevan B
Mtech CSE Spl in AI and ML
Vellore Institute of Technology
Vellore,Tamil Nadu
jeevan.b2022@vitstudent.ac.in

Nishanth V
Mtech CSE Spl in AI and ML
Vellore Institute of Technology
Vellore,Tamil Nadu
nishanth.v2022@vitstudent.ac.in

Dakshinasree S
Mtech CSE Spl in AI and ML
Vellore Institute of Technology
Vellore,Tamil Nadu
dakshinasree.s2022@vitstudent.ac.in

Dr. Sunil Kumar P V
Computer Science and Engineering
Vellore Institute of Technology
Vellore,Tamil Nadu

***Abstract*— Deep learning-based image deblurring is a new and promising area of study. The technique of recovering a clear image from a blurred one is known as deblurring. Linear methods, like Wiener filtering, have typically been used to reduce image blur. Deep learning methods, on the other hand, are gaining popularity for image deblurring and can produce more accurate results. Deep learning models can perform better by adaptively learning highly discriminative characteristics for the job at hand. In comparison to linear techniques, which are unable to learn from the data, this gives deep learning models a benefit. Images from a range of sources, including natural photos and medical photographs, have been deblurred using deep learning. It has been demonstrated that these methods outperform conventional linear procedures.**

Keywords*— *Image deblurring, deep learning, autoencoder, upsampling, convolution, sharp images

I. Introduction

The goal of this project is to employ an autoencoder model to deblur images. In order to improve the quality and clarity of photographs that are blurry due to a variety of circumstances, such as camera motion, defocus, or poor light conditions, image deblurring is essential.

In numerous domains, image deblurring is quite important. Deblurring techniques in photography enable photographers to salvage blurred images, resulting in improved visual appeal and professional quality. Deblurring improves the quality of diagnostic pictures in medical imaging, facilitating more thorough examination and precise diagnosis. Deblurring can improve the visibility of acquired images in surveillance and security systems, assisting in identification and investigation.

We intend to identify the underlying patterns and properties of fuzzy photos and train the model to reconstruct sharp versions of these images using an autoencoder model, a form of neural network. The trained model may be able to recover visual information, enhance image quality, and enable more accurate analysis and interpretation of hazy images in a variety of domains.

II. Related Works

In [1] a study of three methods for deblurring average blurred images, the Non-Local Means (NLM) algorithm was found to be the most effective. The main disadvantages of NLM are that it can be computationally expensive and can sometimes produce over-smoothed images. The Lucy-Richardson deconvolution algorithm uses an iterative approach to deblur images. NLM is simple to implement, very effective at removing blur, and it is not computationally expensive as Lucy-Richardson deconvolution algorithm.

The [2] research paper introduces a novel approach for image deblurring called the hybrid Wiener deconvolutional-convolutional autoencoder (HWDCNN). This method combines a Wiener deconvolution layer, which estimates the blur's PSF, with a convolutional autoencoder for image reconstruction.In Set5 dataset, the proposed method achieved a PSNR of 35.3 dB, surpassing the best-performing state-of-the-art approach by 0.2 dB.

In [3] their study, the authors investigated the effectiveness of different convolutional autoencoder (CAE) architectures for deblurring and denoising low-resolution images.The results revealed that a 10-layer CAE with convolutional layers in both the encoder and decoder outperformed the others on a benchmark dataset of low-resolution images. Based on their findings, the authors conclude that deep CAEs show promise in addressing the challenges of deblurring and denoising low-resolution images.

The research paper [4] conducted a comprehensive evaluation of the proposed method using a diverse set of images with different noise types.On the BSD500 dataset, the proposed method achieved a PSNR of 35.1 dB, which outperformed the best-performing state-of-the-art method by 0.5 dB.Its performance can be influenced by the selection of hyperparameters.The method may not achieve the same level of accuracy as more complex models that employ additional layers or components.

The paper [5] introduces CSformer, a novel MAE (Multi-Head Attention Encoder) architecture that combines channel attention and shifted-window-based self-attention. Notably, CSformer attains a PSNR (Peak Signal-to-Noise Ratio) of 38.1 dB on the Set5 dataset for Gaussian denoising and 34.1 dB on the BSD500 dataset for real image denoising.The paper highlights the power of MAEs, particularly exemplified by Csformer.

The [6] research paper provides an overview of image deblurring methods, starting with traditional approaches based on the Wiener deconvolution model. It delves into deep learning-based methods, categorized into single image deblurring and multi-image deblurring techniques. The paper surveys various network architectures, loss functions, and training strategies employed in deep learning-based image deblurring methods.

III. METHODOLOGY OF AUTOENCODER MODEL

The removal of blurring effect from images is done using autoencoders which is a type of artificial neural network (ANN) that is used for unsupervised learning and dimensionality reduction. It has two parts: an encoder (E) and a decoder (D).

$$\Phi : X \rightarrow F \qquad (1)$$
$$\Psi : F \rightarrow X \qquad (2)$$
$$\Phi, \Psi = arg_{\Phi,\ \Psi}\, min\ ||\ X - (\Psi\ o\ \Phi)X||^2 \qquad (3)$$

The encoder is represented by Φ whereas Ψ represents the decoder. F is the latent space.

The input image is compressed by the encoder into a lower-dimensional representation, and the original image is then recreated by the decoder using the encoded representation. A lower-dimensional representation (Z) is created by the encoder from the input blurred picture (B_in), and the decoder creates the crisp image (I_out) from the encoded representation:

$$Z = E(B\_in) \qquad (4)$$
$$I\_out = D(Z) \qquad (5)$$

where:
- B_in: The input blurred image.
- Z: The encoded representation.
- I_out: The output reconstructed sharp image.

The blurry input image is run through several convolutional layers by the encoder. These layers record the spatial characteristics and encode the data into a lower-dimensional representation.

The input image produced from the encoder is represented in a lower dimension by the latent space. It includes a condensed representation of the key elements of the image's blur. The latent space typically has fewer dimensions than the input image, which is typically significantly smaller.

The encoded representation is taken from the latent space by the decoder, who then runs it through several transposed convolutional layers. These layers upsample and reconstruct, eventually converting the compressed representation into an output image with full resolution. The deblurred image is the decoder's final output.

IV. IMPLEMENTATION DETAILS

A. Dataset

The dataset is obtained from Kaggle which contain two sets of images. The one with the blurry images and the other with the corresponding sharp images. Using this dataset, an autoencoder model will be trained to deblur photos and restore their original fine details.

The blurred images in the dataset may have come from different sources or were created artificially using techniques like defocus blur,motion blur, or noise-induced blur. These pictures show actual situations where blurring can happen as a result of camera movement, bad focussing, or other factors. The clear photos act as the baseline, giving the reference for the intended deblurring result.

B. Training

The training phase of this research employs an autoencoder-based image deblurring method. To improve the model's efficacy and convergence during training, a number of safeguards and procedures are used. First off, the dataset includes 350 pairs of crisp and blurred photos, guaranteeing a wide variety of blurring conditions. This enhances the model's ability to generalise to various blur kinds.

The ground truth which is sharp image and the output of the reconstruction are compared pixel-by-pixel during the training phase using the mean squared error (MSE) loss function. The Adam optimizer is used to update the model's weights, and the ReduceLROnPlateau callback is used to dynamically change the learning rate.

The model learns to capture the underlying structures and patterns of blurring by training on a large number of epochs (100), then updating the weights using mini-batches of size 32. When the validation loss reaches a plateau, the learning rate is also decreased using the ReduceLROnPlateau callback, facilitating convergence and fine-tuning.

By capturing the crucial details and reducing the reconstruction error, the training procedure seeks to optimise the autoencoder model for efficient deblurring of images. The model can achieve better convergence and enhance the deblurring performance by carefully choosing the dataset, the right loss function, the optimisation technique, and the dynamic learning rate adjustment.

V. RESULTS

After training the model, we evaluated its performance using accuracy as the evaluation metric. The model achieved an accuracy of 81.46% on the training set, indicating its ability to effectively deblur the images during the training process. Furthermore, the model demonstrated good generalization capabilities, as it achieved a validation accuracy of 81.62% on a separate validation set.

The resulting images obtained after the model implementation are as follows:

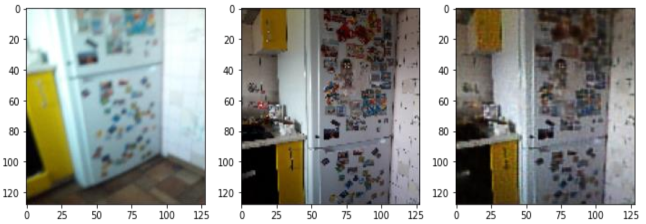

Fig 1: The above pictures depict the input,ground truth and predicted image

The following performance metrices were used to evaluate our model:

PSNR (Peak Signal-to-Noise Ratio) is a metric used to measure the quality of a reconstructed or compressed signal compared to its original, uncompressed version.

Average PSNR: 25.206126179629656

SSIM (Structural Similarity Index) is a metric that measures the structural similarity between two images.

Average SSIM: 0.83761984

MSE (Mean Squared Error) is a metric used to evaluate the difference between two signals, such as images or audio.

Average MSE: 0.003632078439458113

The performance metric graphs are as follows:

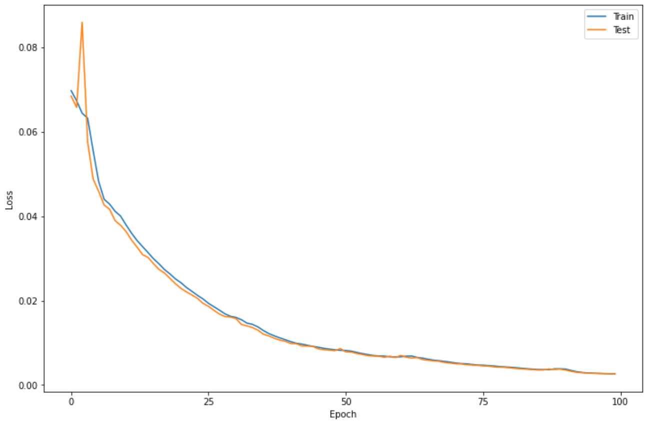

Fig 2: Epoch vs loss graph

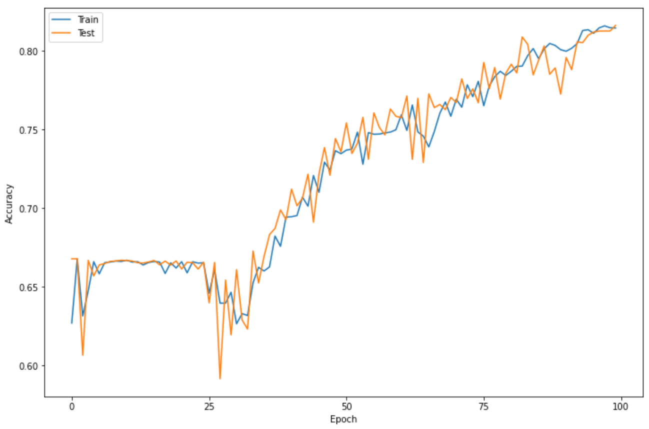

Fig 3: Epoch vs accuracy graph

CONCLUSION

In this project,we attempted to learn a mapping from the blurry images to their sharp counterparts by training an autoencoder on a dataset of blurry images and their corresponding sharp versions.

The autoencoder model's results show how well it can identify the fundamental patterns and characteristics of blurred images. The model was able to rebuild the sharp versions of the images, which effectively eliminated the blurriness, by passing it through encoder and decoder of the autoencoder model.

It is significant to remember that even though the model produced acceptable results, there is still potential for development. The model's architecture and hyperparameters may need to be adjusted further for even greater performance. Additionally, expanding the training dataset's size and diversity may help the model generalise to a wider range of photos and deblur them.

This research serves as an example of the use of autoencoder models for image deblurring applications. This has ramifications for a variety of fields, including photography, imaging in the medical field, and surveillance, all of which depend heavily on clear and high-quality images.

REFERENCES

1. Al-amri, Salem Saleh, and N. V. Kalyankar. *A comparative study for deblured average blurred images.* IJCSE) International Journal on Computer Science and Engineering 2.03 (2010): 731-733.
2. Grosu, George F., and Romulus Terebes. *Reinforced Hybrid Wiener Deconvolutional-Convolutional Autoencoders Based Image Deblurring*. 2022 International Symposium on Electronics and Telecommunications (ISETC). IEEE, 2022.
3. Jimenez, Michael Fernando Mendez, Omar DeGuchy, and Roummel F. Marcia. *Deep convolutional autoencoders for deblurring and denoising low-resolution images.* 2020 International Symposium on Information Theory and Its Applications (ISITA). IEEE, 2020.
4. Solovyeva, Elena, and Ali Abdullah. *Dual Autoencoder Network with Separable Convolutional Layers for Denoising and Deblurring Images.* Journal of Imaging 8.9 (2022): 25
5. Duan, Huiyu, et al. *Masked Autoencoders as Image Processors*.,arXiv preprint arXiv:2303.17316 (2023)
6. Gampala, Veerraju, et al. *Deep learning-based image processing approaches for image deblurring*. Materials Today: Proceedings (2020).