

Proceedings

Image Resolution Enhancement Using Convolutional Autoencoders [†]

Pawan Kumar * and Nikita Goel

Virtuous Transactional Analytics Pvt. Ltd., Noida 201309, India

* Correspondence: pawan.gupta@vitrana.com

† Presented at the 7th Electronic Conference on Sensors and Applications, 15–30 November 2020; Available online: <https://ecsa-7.sciforum.net/>.

Published: 15 November 2020

Abstract: Resolution is an important characteristic to determine the nature and features of the image. Enhancing the resolution strengthens the features hidden within the image, and make the image sharper and more informative. The image quality is improved when noise is removed/suppressed from it. The proposed model provides a technique to enhance the resolution of different types of images, obtained from imaging devices, using a convolutional autoencoder. A convolutional neural network (CNN) architecture is developed by adding different layers to the neural network. An autoencoder capable of encoding and decoding the structure of the images is proposed to enhance their resolution. The model tends to learn the lower-dimensional features of unclear images and provide a high resolution to them by predicting and enhancing their dimensions. The model is trained on low-resolution images and the corresponding high-resolution images, and a convolutional auto-encoder is implemented to denoise the image to introduce high-resolution in the blurred or corrupted images. The model overcomes the limitations of the existing denoising filter techniques and provides a higher level of image quality enhancement.

Keywords: autoencoder; convolutional neural network; image resolution enhancement

1. Introduction

Resolution enhancement is an important pre-processing step in image processing as it helps in refining the characteristics of the image. Medical image classification and diagnosis algorithms use image enhancement as the very first step. The medical images can be enhanced using contrast enhancement techniques, like an average filter, neighborhood operation, imadjust, bilateral ratinex, and sigmoid function [1]. These techniques commonly use parameters like mean square error (MSE), peak signal to noise ratio (PSNR), root mean squared error (RMSE), and normalization coefficient (NC). In addition to medical applications, image enhancement is used for underwater images, infrared images, and haze visibility enhancement [2]. The image resolution enhancement can also be applied to the cropped section of the image to extract features within the cropped area [3]. A comparative approach can be implemented to measure the performance of the different types of filters on image enhancement, where PSNR can be the most appropriate index for comparing the filtered and non-filtered images [4]. One of the traditional approaches used superimposing of the multiple low-resolution frames to obtain a high-resolution image by using projections onto convex sets (POCS) [5]. Dual-tree complex wavelet transformation (DT-CWT) technique was also popularly used for enhancing the satellite image resolution [6]. To intensify the denoising process of the images, a variation of autoencoder can be used to decrease the noise in the image and adaptively improve from the cases encountered [7].

Low light resolution enhancement has also been studied using the autoencoders, which can be used to enhance the lighting of the dull image [8]. A few other methods include inaccurate sub-pixel

matching by applying enhancement ratio on the low-resolution image to obtain gray values of the high-resolution image [9]. The displacement method was also used, which is uniform and applied to the entire image [10]. The sonar images were enhanced by improving their contrast and by learning their characteristics [11]. De-convolutional autoencoders were used to enhance the quality of the fingerprint signal using ConvEnhance, which learns local and global features of the data [12]. Since videos are just a series of images, the image resolution techniques have also been applied to videos. Such a process changes the resolution of the videos, where the low-resolution frame of the video was reconstructed by a regression kernel [13].

The proposed model involves the training of the encoder to create a feature map, which maps the input data with the meaningful features and variations extracted by the encoder. The feature map will become a major building block that could be used to reconstruct the same input data again, which was provided to the encoder. It will be a smaller subset of the input data that hold the necessary information required to develop an identical version of it. In the process of autoencoding of image resolution, the primary step is to solve the image denoising problem. The aim of this study is to strengthen the resolution of the image to obtain clearer images.

2. Methodology

In the proposed model, an autoencoder is developed by adding different layers of the convolutional neural networks. Each layer of the neural network is used to process the input image into different stages and matrices to denoise the image. The model is then trained on a set of low-resolution images and their corresponding high-resolution images to extract the features of those images in a designed encoder. The loaded weights of the encoded images are stored, after the training, as feature vectors. An input test image, which is of low quality, is used to observe the results of the trained model. The model improves the quality of the input image by using the pre-trained autoencoder model. The features used in the model are highlighted below, which will give a detailed understanding of the working and architectural flow of the presented model.

2.1. Dataset

Dataset consists of 9,544 images of different car models, which consists of high-resolution images of the objects within the image.

2.2. Algorithm

The proposed model aims to predict the high-resolution image of the low-resolution image provided as input. It is implemented by applying the following layers to the neural network (shown in Table 1).

- Adding two layers of convolutional two-dimensional with output shape of (256, 256, 64) and an activation layer of rectified linear units (ReLU) function, shown below.

$$y = \max(0, x) \quad (1)$$

These two layers are stored separately in the outer layer 1 and extracts important features from the input image before going down to smaller feature space.

- Max pooling layer, with the output shape of (128, 128, 64), is connected to the previous two-dimensional convolution layer and it reduces the dimensionality into smaller feature space.
- Dropout layer, with the shape of (128, 128, 64), is used to regularize the complex neural network layers running in parallel.
- Again, two layers of two-dimensional convolutional layers are applied one after the other with the same output shape of (128, 128, 128) and with the activation function of ReLU calculated from the formula given above. The layers up to this are stored in outer layer 2 for further processing.

- Max pooling layer, with the output shape of (64, 64, 128), is connected to the previous two-dimensional convolution layer.
- A two-dimensional convolution layer is applied after the previous max pooling layer with the output shape of (64, 64, 256).
- UpSampling layer, with the output shape of (128, 128, 256), is applied to expand the small image matrix into a larger one.
- Again, two layers of two-dimensional convolutional layers are applied one after the other with the same output shape of (128, 128, 128) and with the activation function of ReLU.
- The addition operation is performed on the previously convolutional layer with the outer layer 2, which was stored for further processing with the shape of (128, 128, 128).
- UpSampling layer, with the output shape of (256, 256, 128), is applied to expand the small image matrix into a larger one.
- Two layers of two-dimensional convolutional layers are applied one after the other with the same output shape of (256, 256, 64) and with the activation function of ReLU.
- Another operation is performed on the previously convolutional layer with the outer layer 1, which was stored for further processing with the shape of (256, 256, 64).
- The final layer of the model is applied to the two-dimensional convolutional layer with the output shape of (256, 256, 3) to obtain the resultant autoencoder output.

Table 1. Autoencoder summary.

Layer (Type)	Output Shape
Input Layer	(256, 256, 3)
Conv (2D)	(256, 256, 64)
Conv (2D)	(256, 256, 64)
MaxPooling (2D)	(128, 128, 64)
Conv (2D)	(128, 128, 128)
Conv (2D)	(128, 128, 128)
MaxPooling (2D)	(64, 64, 128)
Conv (2D)	(64, 64, 256)
Add	(128, 128, 128)
UpSampling (2D)	(128, 128, 256)
Conv (2D)	(128, 128, 128)
Conv (2D)	(128, 128, 128)
Add	(128, 128, 128)
UpSampling (2D)	(256, 256, 128)
Conv (2D)	(256, 256, 64)
Conv (2D)	(256, 256, 64)
Add	(256, 256, 64)
Output Layer	(256, 256, 3)

2.3. Encoder

An encoder is used to encode all the necessary information or features stored within the image. It is used to represent the original image in smaller representations in the form of feature vectors, which stores the characteristic of the image and can be decoded whenever the corresponding decoder

is applied to the feature vector. As shown in Figure 1, the original image is stored as a compressed representation after the encoder is applied to it.

2.4. Decoder

The decoder uses the feature vector of a compressed image and is used to reconstruct the original image. It is implemented with the encoder so that important features are not lost at the time of decoding the image. As shown in Figure 1, the compressed image is converted back into its original state when a decoder is applied to it.

2.5. Autoencoder

Autoencoder use unsupervised learning techniques to denoise the image. It performs convolution and deconvolution procedures on the input image using the same input and output matrix [11]. The autoencoders can also be used for image enhancement that can strengthen the process of feature extraction. An autoencoder, while encoding and decoding the image, tends to perform a lossless transformation in an ideal scenario. Practically, an ideal autoencoder is difficult to implement, and hence a distance function is maintained, which calculates the loss encountered when an image is transformed from the encoder to the decoder.

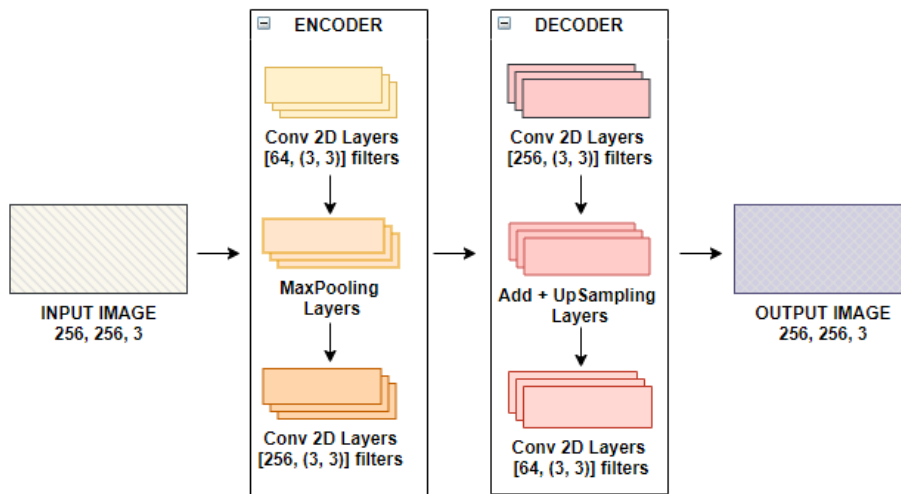


Figure 1. Schematic of an autoencoder.

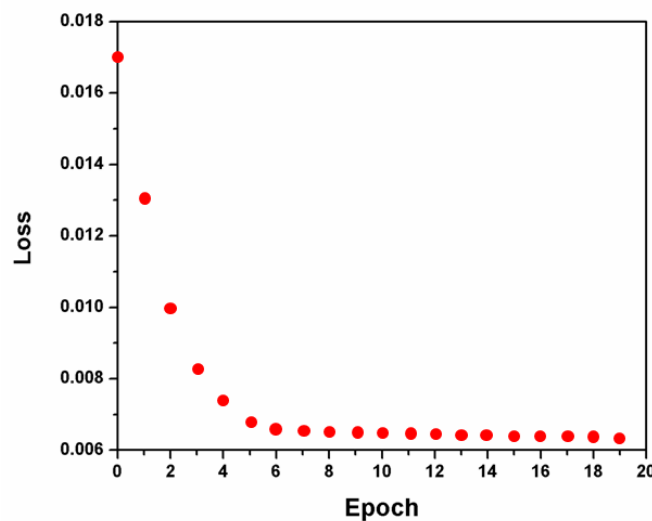


Figure 2. Loss function.

3. Results

The presented model is trained on a dataset of 9544 images. A powerful open-source library Keras with GPU acceleration is used for training the autoencoder. The graph showing the variation of the loss function and epoch is given in Figure 2. After the training, the model is used to give a high-resolution image from the input low-resolution image. Figure 3 shows the low-resolution images, given as input to the model, and the high-resolution output images.

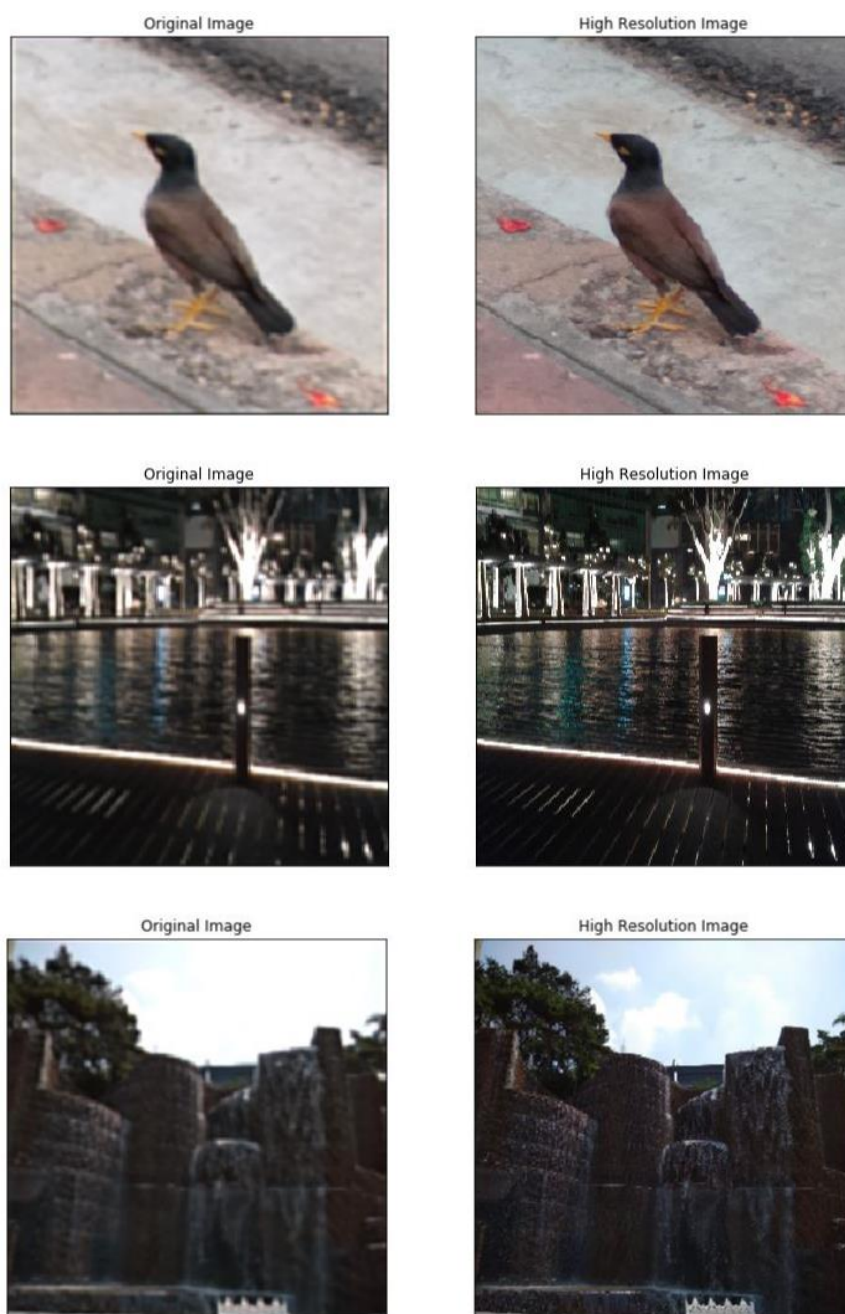


Figure 3. Comparison of the original low-resolution and predicted high-resolution images.

4. Conclusions

The existing denoising filter techniques have some limitations, as they may not be able to enhance the low light images or the naturally blurred images. The presented model aims to remove the noise from the image using autoencoders, where the model is trained on the low-resolution and their corresponding high-resolution images. The autoencoder aims to understand the essential

features of the images and create feature vectors that are later used to improve the quality of the input low-resolution images. So, the chances of quality enhancement get increased to a much higher level in comparison to the traditional approaches. The model performs self-learning, and, hence it learns from the previous noises that it recognized or corrected. Therefore, as the dataset increases, the quality of the image resolution enhancement rises at a faster pace. The proposed model could be useful for CCTV camera image enhancement for police and security departments, number plate recognition at traffic signals, medical image diagnosis for skin pigmentation, detection of any tissue injury, and other applications.

Author Contributions: Conceptualization, P.K. and N.G.; methodology, P.K.; software, P.K.; validation, N.G.; formal analysis, P.K.; investigation, N.G.; resources, P.K.; data curation, P.K. and N.G.; writing—original draft preparation, P.K. and N.G.; writing—review and editing, P.K.; visualization, N.G.; supervision, P.K.; project administration, P.K. and N.G.; funding acquisition, P.K. and N.G. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rundo, L.; Tangherloni, A.; Nobile, M.S.; Militello, C.; Besozzi, D.; Mauri, G.; Cazzaniga, P. MedGA: A novel evolutionary method for image enhancement in medical imaging systems. *Expert Systems with Applications* **2019**, *119*, 387–399.
2. Ancuti, C.O.; Ancuti, C.; De Vleeschouwer, C.; Bekaert, P. Color balance and fusion for underwater image enhancement. *IEEE Transactions on Image Processing* **2017**, *27*, 379–393.
3. Liu, Z.S.; Siu, W.C.; Wang, L.W.; Li, C.T.; Cani, M.P.; Chan, Y.L. Unsupervised real image super-resolution via generative variational autoencoder. In 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 14–19 June 2020, pp. 1788–1797.
4. Hoshyar, A.N.; Al-Jumaily, A.; Hoshyar, A.N. Comparing the performance of various filters on skin cancer images. *Procedia Computer Science* **2014**, *42*, 32–37.
5. Tekalp, A.M.; Ozkan, M.K.; Sezan, M.I. High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration. In 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), San Francisco, CA, USA, 23–26 March 1992, pp. 169–172.
6. Demirel, H.; Anbarjafari, G. Satellite image resolution enhancement using complex wavelet transform. *IEEE Geoscience and Remote Sensing Letters* **2010**, *7*, 123–126.
7. Vincent, P.; Larochelle, H.; Lajoie, I.; Bengio, Y.; Manzagol, P.A.; Bottou, L. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of Machine Learning Research* **2010**, *11*, 3371–340.
8. Lore, K.G.; Akintayo, A.; Sarkar, S. LLNet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognition* **2017**, *61*, 650–662.
9. Li, P.; Shen, H.; Zhang, L. A method of image resolution enhancement based on the matching technique. Available online: <http://www.isprs.org/proceedings/xxxv/congress/comm3/papers/356.pdf> (accessed on 10 October 2018).
10. Irani, M.; Peleg, S. Improving resolution by image registration. *CVGIP: Graphical Models and Image Processing* **1991**, *53*, 231–239.
11. Kim, J.; Song, S.; Yu, S.C. Denoising auto-encoder based image enhancement for high resolution sonar image. In 2017 IEEE Underwater Technology (UT), Busan, 21–24 February 2017, pp. 1–5.
12. Schuch, P.; Schulz, S.; Busch, C. De-convolutional autoencoder for enhancement of fingerprint samples. In 2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA), Oulu, 12–15 December 2016.
13. Islam, M.M.; Asari, V.K.; Islam, M.N.; Karim, M.A. Super-resolution enhancement technique for low resolution video. *IEEE Transactions on Consumer Electronics* **2010**, *56*, 919–924.

