Scientific Journal of Informatics
Vol. 9, No. 1, May 2022
p-ISSN 2407-7658 http://journal.unnes.ac.id/nju/index.php/sji e-ISSN 2460-0040

Improvement of Image Quality Using Convolutional Neural Networks Method

Arief Kelik Nugroho[1*], Ipung Permadi[2], Muhammad Faturrahim[3]
[1,2,3]Informatics Department, Faculty of Engineering, Universitas Jenderal Soedirman, Indonesia

Abstract.
Purpose: This desire for high resolution stems from two main application areas, namely improving pictorial information for human interpretation and assisting automatic machine perception in representing images or videos. Image resolution describes the detail contained in an image, the higher the resolution, the more detail there is. The resolution of a digital image can be classified into various types, namely pixel resolution, spatial resolution, temporal resolution, and radiometric resolution. In this context, we are interested in spatial resolution.
Methods: Elements of a digital image consist of a collection of small images called pixels. Spatial resolution refers to the pixel density of an image and is measured in pixels per unit area. A quality digital image is determined by the size of the resolution it has. A low resolution or low-resolution is a drawback of a digital image because the information contained in the image means little compared to a high-resolution image.
Result: Therefore, in this study, a digital image processing program was created in the form of Image Super-Resolution with the Convolutional Neural Network method to utilize low-resolution images to produce high-resolution images. With a fairly short training process, namely 6050 datasets with 100 CNN epochs, the average PSNR image is 5% higher.
Novelty: Image quality can be improved by changing the parameters in the CNN method so that image quality can be improved.

Keywords: Image, Resolution, CNN, Size, Pixels
Received June 2021 / **Revised** January 2022 / **Accepted** March 2022

INTRODUCTION

In an application of Digital Imaging, image (pictures or video) with a resolution higher is desirable for processing or analysis. The desire to get a resolution high is originated from two areas of applications principal namely an increase in the information display for the interpretation of man and help the perception of the machine automatically in representing images or videos. Image Resolution (resolution image) describes the detail contained in an image, the higher the resolution, the more a lot of detail there [1]. Digital images that are often encountered have low resolution, which means they store little information and make it difficult for image users, namely humans who see and computer machines that use these digital images for certain purposes. Image Super-Resolution (SR) is a technique to reconstruct the image of a resolution higher than the image resolution of the low was observed [2]. Because it required an application Super Resolution that can transform an image digitally -resolution low into an image with a resolution higher [3]. An image is a combination of points, lines, planes, and colors to create an imitation of an object – usually a physical object or a person. Imagery can be a tangible image (picture) are two -dimensional, like a painting, photograph, and tangible three-dimensional, like a statue [4], [5]. While digital images are images that are displayed on a computer as a set of digital values (resolution) in the form of pixels. The resolution can be low or high and it determines the quality of a digital image [6], [7]. In an application of Digital Imaging, an image (pictures or video) with a resolution higher is desirable for processing or analysis. The desire to get a resolution high originated from two areas of applications principal namely an increase in the information display for interpretation humans and helps automatic machine perception in representing images or videos. Image resolution describes the detail contained in an image, the higher the resolution, the more a lot of detail there [3].

* Corresponding author.
Email addresses: arief.nugroho@unsoed.ac.id (Nugroho), ipung.permadi@unsoed.ac.id (Permadi), muhammad.faturrahim@unsoed.ac.id (Faturrahim)
DOI: 10.15294/sji.v9i1.30892

The resolution of a digital image can be classified into various types, namely pixel resolution, spatial resolution, temporal resolution, and radiometric resolution. In this context, we are interested in spatial resolution [8]. Elements image digitally consists of a collection of images of small are called pixels. Spatial resolution refers to the pixel density in an image and in measuring the pixels per unit area [5].

In getting the image digitally, a spatial resolution that is higher obtained from how good the sensor that is used to capture the image of the good lens optics are used. Super-Resolution is a technique that is used to build a high image resolution (High-Resolution) or image resolution Low (Low-Resolution) which has been observed [1].

LITERATURE REVIEW

Convolutional Neural Networks

Convolutional Neural Networks, or hereinafter referred to as CNNs (Figure 1, Figure 2) are artificial neural networks whose neurons are represented as a matrix (n x n for an example of a 2-dimensional matrix), unlike neural networks in general which are vectors (n x 1) [3], [9], [10].

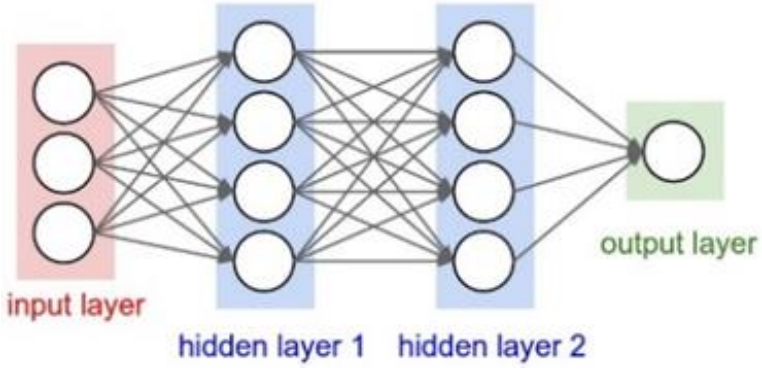

Figure 1. Network Neural Artificial normal

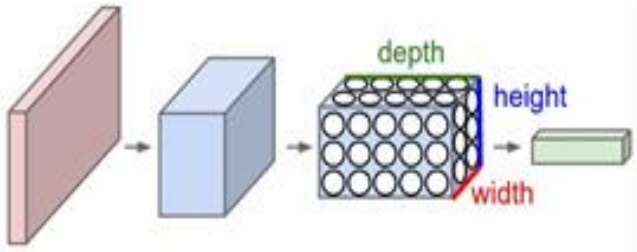

Figure 2. Convolutional Neural Networks

CNNs Architecture

The following is the architecture of a network of CNNs [11].

Input layers

Layer that contains a pair of images (x, y) where x is the original high-resolution image and y is the image x that has been processed into a low resolution. CNN's have a special layer called the convolutional layer. This layer functions as a filter when preprocessing the input neurons. The convolutional layer is represented in a matrix that has the same dimensions as the input layer. For example, if the input is 2D, then the Conv layer is also 2D [12]. The layer size is smaller than the input, which is commonly used as 3x3 for 2D. This layer aims to find out the features contained in an image such as vertical, horizontal edges, gradients, and others. To find out the various features, different filters are needed. In getting the output

from this layer, the dot product is done between the convolutional layer and the input layer as shown in Figure 3.

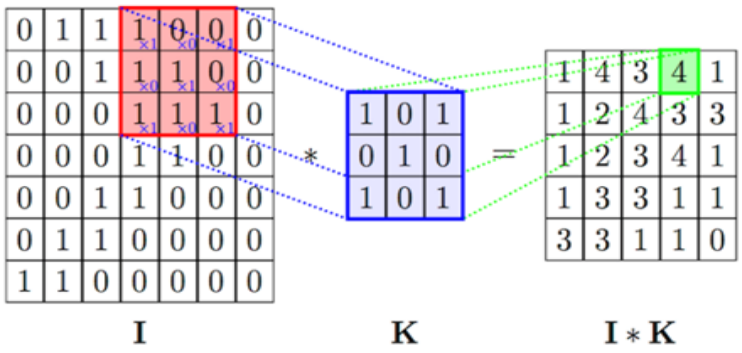

Figure 3. Matrix I is the input layer and matrix K is the filter in the convolutional layer.

ReLU Activation Function

The output layer of the convolutional layer will be subjected to the activation function (Figure 3). The activation function used is the Rectified Linear Unit (ReLU) [1].

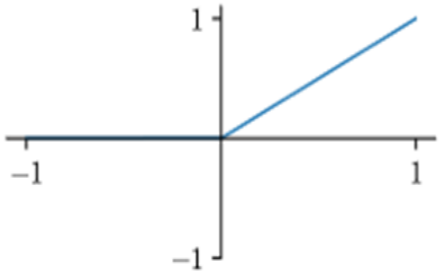

Figure 4. ReLU

The definition of the ReLU function is

$$f(x) = \begin{cases} x \\ 0 \end{cases} \quad (1)$$

Where **x if x > 0, and 0 otherwise.**

If the element is negative then the value is set to 0, with no exponential, multiplication, or division operations. With such characteristics, the advantages of ReLU will appear when dealing with networks that have a lot of neurons so that it can reduce training and testing time significantly.

Pool Layers

Layer that is inserted after Conv layer + ReLU (figure 5) serves to reduce the size of the representation (matrix) to reduce the number of parameters (each neuron is a parameter) and computations in the neural network. There are two types of pooling, namely Max-pooling to choose the highest value or average-pooling to find the average [1].

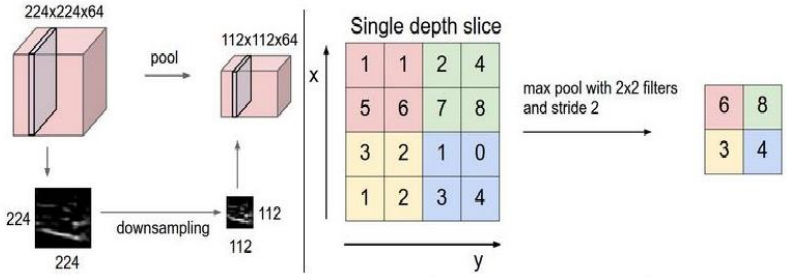

Figure 5. Pooling layer process with max-pooling filter 2x2

Fully Connected Layer

The neurons in the screen FC are fully connected to all activations in the previous layer as shown in Figure 6.

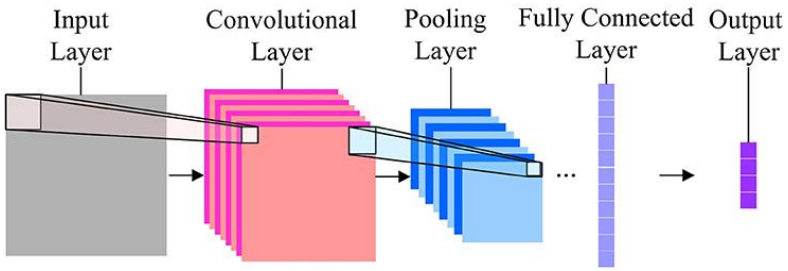

Figure 6. Example of CNN architecture

METHODS

Based on this architecture, the CNN training procedure begins with preparing data in the form of High Resolution (HR) and Low Resolution (LR) images on the same type of image. The LR image was obtained by downsampling the original HR image according to Figure 7.

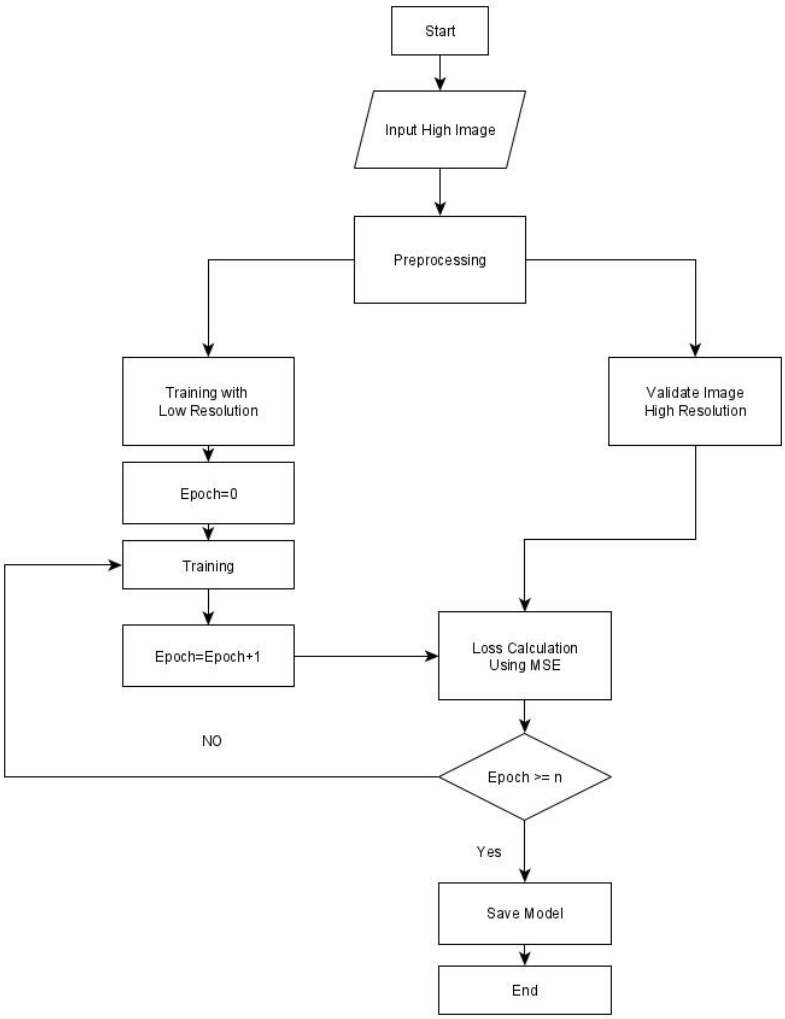

Figure 7. CNN Training Flow

To get an optimal network, the error rate or error must be as small as possible. Therefore we need a function called Loss Function [9]. The loss function that will be used is the mean squared error, or MSE, an error estimation method commonly used in signal processing [13]. MSE is a pixel-wise comparison with the formula 2

$$L_{\text{MSE}}(y, \hat{y}) = \frac{1}{N}||y - \hat{y}||_2^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2, \quad (2)$$

where N is the number of pixel channels in the image [6]. For example, for an image with a pixel size of RGB color. From MSE we can also calculate the peak signal-to-noise ratio, or PSNR, which is used to evaluate the existing neural network [1], [13].

$$\text{PSNR} = 10\log_{10}\left(\frac{L^2}{L_{\text{MSE}}(y,\hat{y})}\right) \quad (3)$$

where L is the dynamic range of each image [6]. Examples for color 8-bit integer, L = 255. Unit PSNR measured in decibels, the higher the value the network is getting good. The image that is identical having *i* value of MSE = 0 and PSNR = ∞ dB [14].

RESULT AND DISCUSSION

Data Acquisition

In developing this Super Resolution application, there are 2 types of datasets used on the CNN network. The first dataset is used as training data and the second dataset is used as test data for 6050 HR images from the "Kou photo collection" which is free to download by visitors. The dataset consists of various types of images such as buildings, furniture, and food [13].

Figure 8. Sample Dataset Kou. Photo Collection

Set14 is a dataset that is often used to test a network [15]. This dataset consists of 14 images, but only 9 images will be used in the test.

Results

The model training was carried out with a configuration of x2 scaling and epoch 100. The following are the results of testing using the Set14 dataset which consists of 9 images, namely by comparing the MSE and PSNR values of LR images that have been processed with Lanczos filtering and CNN [16]. The results of calculations with upscaling 2x obtained the values presented in Table 1

Table 1. MSE and PSNR values upscaling 2 times

Image Name	Lanczos upscaling(x2)	
	MSE	PSNR
Baboon	356.8932	22.6054
Barbara	146.2169	26.4808
Coastguard	105.4399	27.9008
Face	45.7540	31.5265
Foreman	47.3821	31.3747
Lenna	36.8088	32.4713
man	102.8257	28.0098
Monarch	44.0056	31.6957
Pepper	55.6836	30.6735
Average	104.5566	29.1932

Table 2 shows the test results using CNN, the results obtained have significantly changed the MSE and PSNR values.

Table 2. MSE and PSNR values with CNN

Image Name	CNN(x2)	
	MSE	PSNR
Baboon	297.6390	23.3939
Barbara	125.8139	27.1335
Coastguard	75.1858	29.3694
Face	39.3101	32.1858
Foreman	37.7148	32.3657
Lenna	25.8514	34.0060
man	65.2992	29.9817
Monarch	12.7152	37.0876
Pepper	49.1228	31.2180
Average	80.9614	30.7491

Table 3 training changes using upscaling 4x, this process is used to test the changes in the value of the training image.

Table 3. MSE And PSNR value upscaling 4 times

Image Name	Lanczos upscaling(x4)	
	MSE	PSNR
Baboon	607.0631	20.2985
Barbara	283.5630	23.6043
Coastguard	258.1030	24.0129
Face	82.9035	28.9451
Foreman	156.5030	26.1856
Lenna	94.5116	28.3760
man	234.0768	24.4372
Monarch	152.9175	26.2862
Pepper	121.6539	27.2795
Average	221.2550	25.4917

Table 4 training changes using 4x upscaling, this process is used to test the value changes in the training image using CNN.

Table 4. MSE And PSNR value upscaling 4 times

Image Name	CNN(x4)	
	MSE	PSNR
Baboon	562,9126	20.6264
Barbara	243.8837	24.2590
Coastguard	226.9385	24.5717
Face	68.8065	29.7545
Foreman	131.6927	26.9352
Lenna	63.8341	30.0803
man	167.8824	25.8808
Monarch	63.1497	30.1271
Pepper	90.2292	28.5773
Average	179.9255	26.7569

According to Table 2, Table 3, and Table 4, it can be seen that the CNN, MSE value is smaller than Lanczos in all tested images, while the PSNR value is higher. The comparative performance of upscaling with the CNN method can produce better image improvements than with the usual method. Learning by using low-resolution images, image quality can be improved by changing the parameters in the CNN method so that image quality can be improved. the result of the comparison is presented in Figure 9.

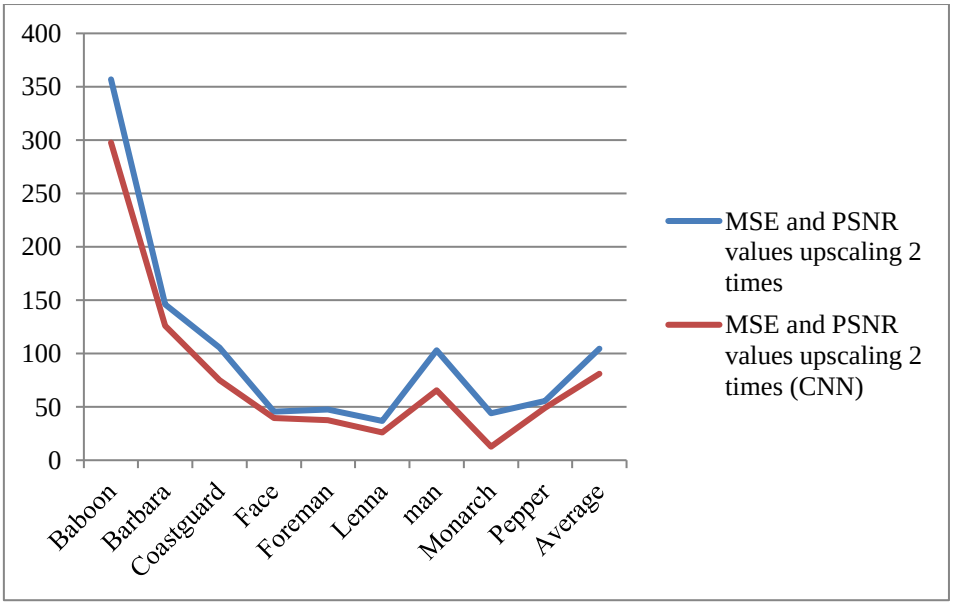

Figure 9. upscaling (2x) MSE and PSNR value change images using CNN and Not using CNN

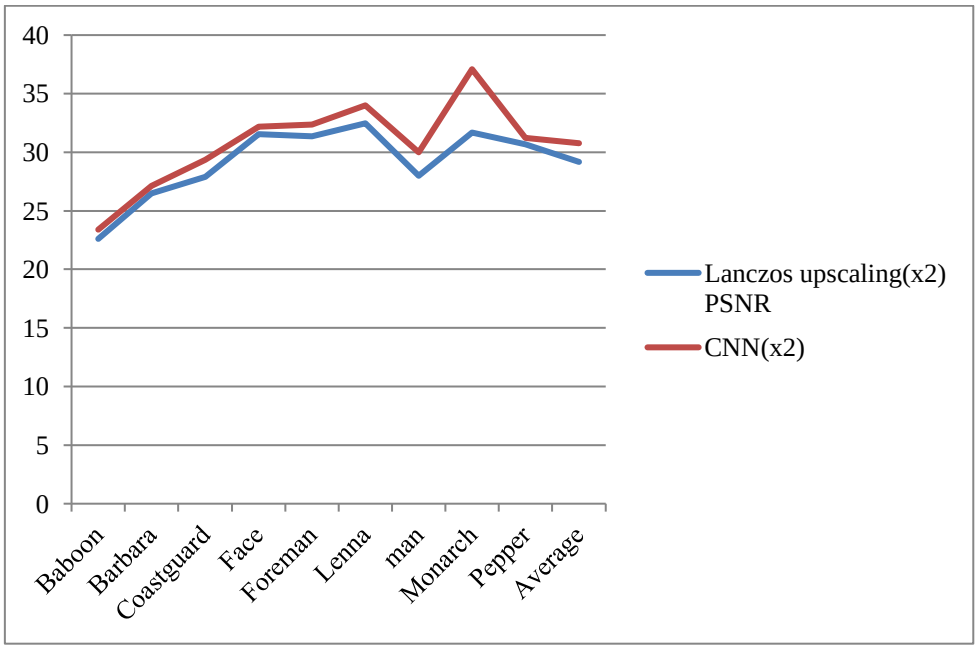

Figure 10. upscaling (4x) MSE and PSNR value change images using CNN and Not using CNN

CONCLUSION

This study aims to develop a Super Resolution application using the Convolutional Neural Networks method. With a fairly short training process, namely 6050 datasets with 100 epochs with a training duration of 16 hours, CNN can outperform the already popular upscaling method, namely Lanczos in the Super Resolution application with an average PSNR of 5% higher.

REFERENCES

[1] P. Liu, Y. Hong, and Y. Liu, "Deep differential convolutional network for single image super-resolution," IEEE Access, vol. 7, no. 1, pp. 37555–37564, 2019.

[2] K. H. and X. T. C. Dong, C. C. Loy, "Image Super-Resolution Using Deep Convolutional Networks," IEEE Trans. Pattern Anal. Mach. Intell., vol. 38, no. 2, pp. 295–307, 2016.

[3] S. Ayas and M. Ekinci, "Single image super resolution using dictionary learning and sparse coding with multi-scale and multi-directional Gabor feature representation," Inf. Sci. (Ny)., vol. 512, pp. 1264–1278, 2020.

[4] C. Feng, M. & Zhao, S. & Xing, "Image quality assessment using PSNR based on visual feature," J. Nanjing Univ. Posts Telecommun. (Natural Sci., vol. 35, no. 4, pp. 33–38, 2015.

[5] C. K. and Z. D. O. Keleş, M. A. Yılmaz, A. M. Tekalp, "On the Computation of PSNR for a Set of Images or Video," 2021.

[6] R. Fan, S. Li, G. Lei, and G. Yue, "Shallow and deep convolutional networks for image super-resolution," Proc. - Int. Conf. Image Process. ICIP, vol. 2017-September, pp. 1847–1851, 2018.

[7] X. Ji, Y. Lu, and L. Guo, "Image super-resolution with deep convolutional neural network," Proc. - 2016 IEEE 1st Int. Conf. Data Sci. Cyberspace, DSC 2016, pp. 626–630, 2017.

[8] X. Shang et al., "A New Combined PSNR For Objective Video Quality Assessment School of Communication and Information Engineering , Shanghai University , China School of Engineering Science , Simon Fraser University , Canada," no. 61271212, pp. 811–816, 2017.

[9] G. Q., Huynh-Thu; M., "Scope of validity of PSNR in image/video quality assessment," Electron. Lett., vol. 44, no. 13, 2008.

[10] B. Nugroho and E. Y. Puspaningrum, "Kinerja Metode CNN untuk Klasifikasi Pneumonia dengan Variasi Ukuran Citra Input," J. Teknol. Inf. dan Ilmu Komput., vol. 8, no. 3, p. 533, 2021.

[11] B. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Commun. ACM, vol. 60, no. 6, pp. 84–90, 2017.

[12] N. Fadlia and R. Kosasih, "Klasifikasi Jenis Kendaraan Menggunakan Metode Convolutional Neural Network (Cnn)," J. Ilm. Teknol. dan Rekayasa, vol. 24, no. 3, pp. 207–215, 2019.

[13] M. Sara, U., Akter, M., & Uddin, "Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A Comparative Study," J. Comput. Commun., 2019.

[14] A. Sharma, "Confusion Matrix in Machine Learning," 2021. https://www.geeksforgeeks.org/confusion-matrix-machine-learning/

[15] P. Mane, Vanita & Jadhav, Suchit & Lal, "Image Super-Resolution for MRI Images using 3D Faster Super-Resolution Convolutional Neural Network architecture," ITM Web Conf., vol. 32, p. 03044, 2020.

[16] Z. Zhang, S. Shan, Y. Fang, and L. Shao, "Deep Learning for Pattern Recognition," Pattern Recognit. Lett., vol. 119, pp. 1–2, 2019.