ADVANCED COMPUTING SESSION

PREPROCESSING IMAGE DATA FOR DEEP LEARNING

Dragana Stojnev*,
Aleksandra Stojnev Ilić

Faculty of Electronic Engineering,
University of Niš,
Niš, Serbia

Abstract:
Neural networks require big amount of input data in order to be properly trained, and the output and its accuracy depend on the quality of the input dataset. Most of the images used to train these networks either contain too much or not enough information, and therefore need to be preprocessed so as to reduce or even remove the noise from them, extract useful information and remove the useless ones, or apply other techniques that improve input quality for a neural network, such as super-resolution. With suitable input provided, it will be possible to create prediction models with higher precision and better accuracy. This paper gives an overview of state-of-the-art techniques for image preprocessing for different convolutional neural networks, and describes an application that demonstrates one of them.

Keywords:
deep neural network, image preparation, super-resolution, noise removal.

INTRODUCTION

Given the fact that the amount of image and video data that is being generated and used on a daily basis is increasing, the need for better and more efficient techniques for image manipulations is also increasing. Recent developments in the field of image processing have led to a renewed interest in neural networks. One subset of neural networks called Convolutional Neural Network is particularly interesting in the area of image processing because it gives a new approach in understanding image data. Since it is very hard to determine all functions that should be performed over the input images so as to extract useful information, the new idea appeared where those functions could be determined by the neural network itself. This cannot be done with a small number of layers in traditional neural networks because of the complexity of the functions that must be represented; there is a need for additional layers in the neural network. That is how the idea of deep neural networks was born.

Correspondence:
Dragana Stojnev

e-mail:
dragana.stojnev@elfak.rs

DOI: 10.15308/Sinteza-2020-312-317

It was shown later that deep neural networks can even outperform regular neural networks in many domains. Nowadays, thanks to the development related to the hardware, both the field of neural and deep neural networks is emerging and new and improved standards in network architectures are being generated continuously. Many of them are tailored to the needs of image processing. However, this is still an open challenge. In order to improve the performance of neural networks, including deep neural networks, additional preprocessing of the input images is required. On one hand, various types of the input data, especially images, contain certain amount of useless data and in order to achieve better performances this information should be removed. On the other hand, instead of removing information, there could be a need to add certain information to the input data. For example, it is possible to combine multiple input images to produce one, more detailed one, as it is done in super-resolution. The scenario when this is appropriate solution is when the frame extraction is performed over an input video and none of the extracted frames contains the crucial information, but some of them together contain some interpretation that could be used. In this manner, it is possible to create a good input from a combination of different data that can be marked as non-complete. Various preprocessing techniques can be used to advance the input data for the neural network, from basic operations such as blur, sharpen, erosion and dilatation, to more complex ones that are tested and developed for other purposes, such as processing of medical and microscopy imagery.

This paper gives a brief overview of some of the state of the art techniques that could be performed in order to provide better input for deep neural networks (DNN). Furthermore, it describes one of the mentioned techniques and presents its implementation as well as the obtained results. The remainder of this paper is organized as follows. Section one gives an overview of the related papers. Section two describes some of the techniques for image preprocessing that can be used on input images for neural networks. Section three describes the developed application and shows the results in form of the resulting images. Section four concludes the paper.

1. RELATED WORK

Image preprocessing often includes morphological operations such as opening and closing. Paper [1] gives an example where these two techniques are used to preprocess astronomical images so as to identify rings, fronts and sunspots. Morphological image processing techniques among which are erosion and dilation are described in [2]. This paper also gives appropriate examples of when any of the techniques is a good choice to make. Different techniques for image reconstruction methods for structured illumination microscopy are presented in [3], such as scaled subtraction, Bayesian estimation and square-law detection. It also shows the results of the algorithms for different evaluating criterias. Paper [4] gives an overview of a super resolution by a factor of two algorithms with the use of spatially structured illumination in a wide-field fluorescence microscope. It describes how to extract high-resolution information from an image that can't be normally processed and gives an overview why is this method better than using conventional or confocal microscopes. The presented algorithm can be modified and applied to the wide range of images so as to enhance them. Deep learning can be used to augment microscopy images. Techniques such as super resolution can reduce the number of raw images required for an input. In [5] it is described how SR-SIM algorithm can be used in deep neural networks with low light images that are forwarded to the network input and by using deep neural networks for augmentation to provide enough input data for super-resolution with structured illumination microscopy. Paper [6] also describes concepts of SR_SIM algorithm but gives an overview of an open-source plugin called fairSIM which uses ImageJ and demonstrates how this algorithm actually works. Super resolution with SIM is most commonly used in terms of medical images, including the imaging of the living brain [7]. The mentioned algorithm can be used in brain imaging in vivo since it can give an insight that can't be given with conventional microscopy. It shows the results of the algorithm performed over the brains of live zebrafish larvae and mice.

2. IMAGE DATA PREPROCESSING

Improving input data quality for neural networks can highly contribute to better prediction results and higher accuracy, which are the main goals when using any kind of neural networks. Some of the techniques that can be used for image data preprocessing for the purpose of improving inputs for neural networks are erosion, dilation, opening, closing and super-resolution. Erosion and dilation are two fundamental morphological operations. Firstly, they were used only over binary images and later on grayscale images. Binary image is actually a matrix containing only 1's and 0's.

Erosion operation has two input arguments – one is the picture over which the erosion will be applied and the second is structuring element, or kernel. Kernel is a set of pixels that define what will be the effect of the operation over given image. Erosion calculates local minimum in each area affected by the kernel. Since dark pixels are represented with 0, if the region overlapping with kernel contains at least one 0 it will be forwarded to the output. Simulation of erosion could be as follows: With input image I and kernel C, resulting image will be actually the result of logical operator AND performed over I and C in each iteration given that matrix traversal is with constant offset either row wise or column wise. The purpose of this is to remove thin links between components and object bulges. This can be used when processing medical images which show a plenty of connected cells and almost always is necessary to detect which are actually connected and which only appear to be connected. In order to show how different operations affect image, we use a basic image drawn in Paint (Fig.1.). The effect of erosion is shown in Fig. 2a, where it can be seen that operation removed the link between two squares and the edges of a square became thinner. The used kernel is ellipse with dimensions (5,5). Unlike erosion which detects local minimum, dilation detects local maximum. It has the same input arguments as erosion, but here the output is the result of logical operator OR performed over input image and kernel. The main usage of dilation is when there is a need for filling existing holes in image subset and object enlargement. When kernel is overlapped with image partition, the output is pixel with largest value and this way all pixels that are by chance set to 0 are annulled. For example, it can be used when input image is processed somehow that resulted with irregular object shapes and all object shapes need to be of the regular shape. The result of dilation of previously shown input is given in Fig. 2b. It can be seen that the edges after the dilation are thicker.

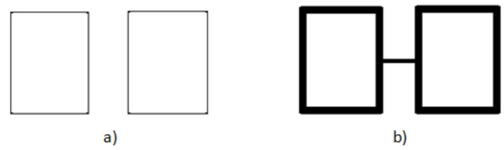

Fig. 2. Resulting image after a) erosion, b) dilation

Combining previously described erosion and dilation it is possible to define some of the most significant operations such as opening and closing. Opening represents erosion followed by dilation with the same structural element. Opening is idempotent operation, reapplying always gives the same result. If input object is some polygon and structuring element circle for example, opening could be simulated as circle "rolling around" the edges on the inside of the given polygon. It is used for removing noise since annuls small points that don't carry any useful information. The result of this morphological operation is shown in Fig. 3a. Since opening is equal to firsrly performing erosion followed by dilation expected output is firstly removed edges and after that resulted edges were highlighted. Closing is reverse operation to opening. First, dilation is applied and after that erosion with the same structural element. This is also idempotent operation. It is useful in removing small holes from the inside of the object which is also categorized as noise. This operation can be used when segmentation is needed. Furthermore, when converting image to a binary format, thresholding operation is performed, and based on the thresholding ratio some pixels can be excluded from the result and closing solves this problem. The effect of this operation over given image is shown in Fig. 3b.

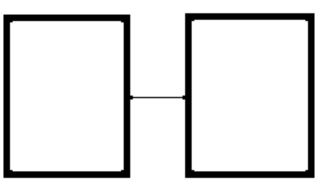

Fig. 1. Input image

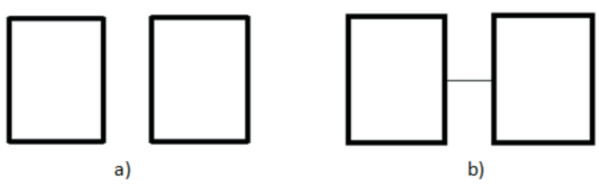

Fig. 3. Resulting image after a) opening, b) closing

Previously described techniques are morphological ones. There are various more complex operations used for image preprocessing that are not morphological. One of them is super-resolution whose main aim is to improve image quality and if possible – extract more information from given input image. A special type of super-resolution called super-resolution with structured illumination microscopy or shortly, SR-SIM is described in [8]. The main idea with this algorithm is to take several low quality images and use them in order to create a new image with better quality and higher resolution. The use of SR-SIM is actually a combination of Moire effects or overlapping large-scale interference patterns. Basic input to the SR-SIM algorithm is an image with specific grid that is overlapped with it and that grid is actually the described pattern. One of the parameters for this algorithm is number of directions which defines how many axes will be used when generating output image. Each direction has 3 images – central image, which is same for every direction and two more that are symmetric to the central one but have different patterns. The output image resolution can be improved maximum twice and that is the case when generated images are sequentially one next to another on given axis. This is done in spectral domain using Fourier transformation. Input images are represented as Fourier specters, their combination is created also in Fourier domain and output image is achieved with inverse Fourier transformation. The actual effects of super-resolution on an image cannot be seen using simple shapes. For that manner, example images for super resolution will be provided in the following section.

Choosing the right enhancement method for an image is not a trivial task – one must take into account the origin of the image, the content and available features, similarity with other images, different types of noise, both spectral and spatial domain, but also the purpose of the enhancement. Each method, if applied correctly and when needed, can significantly improve the outcome of the algorithms that follow, whether they are a neural network or different computer vision algorithms.

3. IMPLEMENTATION

For the purpose of demonstration of the described super-resolution technique, a demo application is created. Entire application was written in python programming language using numpy and OpenCV libraries, and it is a console application as graphical user interface was not needed.

The developed application has five arguments including input image, number of directions, initial phase, radius and center. Input image represents an URI for location of the image that will be processed. Image can be in .jpg or .png format. Number of directions is parameter that, along with the initial phase, defines which parts of the Fourier domain will be taken to create input dataset. Radius and center parameters define a circle that will be taken from the image.

First part of the algorithm creates the input dataset for super resolution algorithm while the second one performs actual resolution enhancements. The first step in part one is to load the input image in grayscale and transform it to Fourier domain. The result will be used for generating multiple images which will be used as an input dataset for the algorithm itself. After switching the input image to Fourier domain, input dataset is generated by specific masking of the input image, meaning that some parts of the spectrum are used and others are not. Which parts of the domain are used is defined with input parameter that determines how many directions will be used. Since all slices are not centered, the next step is to return them to the center of the image. This must be done in order to correctly perform the rest of the algorithm. After the image centering, the next step is to create their combination by putting all the slices to the exact same position from where they were cut. This step is very important because with image centering we achieve the correct shape of grid that basically simulates the moiré effects and by putting them back together we achieve the right results. Finally, it is important to perform inverse Fourier transformation so as to retrieve the output image in spatial domain. Resulting image is a grayscale image.

We tested the application on random image downloaded from the web. The input image and its spectral domain representation are shown in Fig. 4. The number of directions parameter is set to three, phase to zero, radius to 100 while center is equal to the image center (image width/2, image height/2). In this example, there are two slices for each direction and one central slice, as shown in Fig. 5. Not whole specter is necessary since the central part contains most of the information and resulting image can be reconstructed with this tiny part (as brightened the pixel is – the more information contains). Centralized slices are shown in Fig. 6. The result of the inverse Fourier transformation performed on centralized slices is shown in Fig. 7. This step is performed so as to see the actual results of these transformations. Since inverse Fourier transformation returns two-channel signal it couldn't be saved locally as it is, both real and imaginary parts were used to create a magnitude which has only one channel and in this form it is possible to save it directly to hard drive. This was achieved with the use of built in function from OpenCV. Each of the next steps uses the appropriate two-channel inputs. The next phase is to combine slices, again in Fourier domain. The result is shown in Fig. 8. The resulting image in spatial domain is shown in Fig. 9.

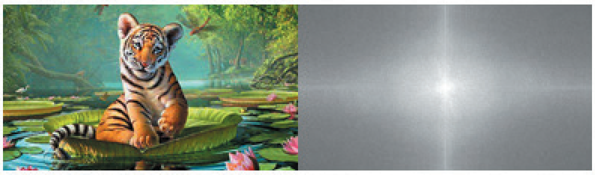

Fig. 4. The original image, found on https://wallpapersafari.com/w/jYg9w6 and the same image in spectral domain

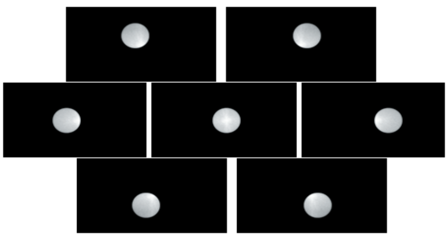

Fig. 5. Slices of the image in spectral domain

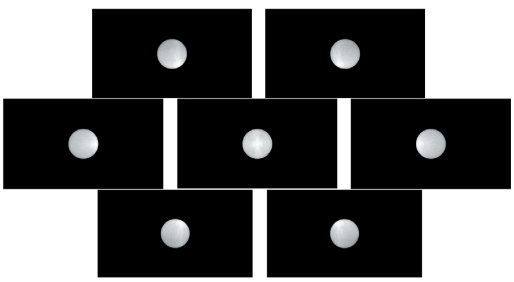

Fig. 6. Centralized slices of the image in spectral domain

Fig. 7. Centralized slices of the image in spatial domain

As it can be seen from the provided figures, implemented application demonstrates each step in SR_SIM algorithm, and it shows its effect on a random image. It is a technique that imitates the process in which a grid pattern is generated through interference of diffraction orders and superimposed on the specimen while capturing images. The grid pattern originally is shifted or rotated in steps between the capture of each image set. The artificially created grid allows us to extract information from the raw image by creating the resulting image that has a lateral resolution approximately twice that of the original image. Maximal enhancement that can be achieved with this algorithm related to the output image resolution is 2x. The main reason for this is the fact that maximal distance the slices can be moved while creating the grid is equal to the half of the radius of the chosen circle.

However, as it can be seen from resulting image, the brightness and contrast are not the best. Although this would not affect the behavior of neural networks when trained with similar input, in order to present the actual improvements that the super resolution filter can bring, it is necessary to perform a postprocessing filtering. The basic method for improving contrast in images is called histogram equalization which increases the global contrast of the image. Improved method, named adaptive histogram equalization computes different histograms for different sections of the image so as to redistribute the lightness values of the image. However, if an image consists of relatively homogenous regions, it can over-amplify existing noise. This can be solved by limiting the actual amplification. This variant of the algorithm is called contrast limited adaptive histogram equalization. We have chosen this algorithm as a postprocessing step which is performed after SR_SI algorithm. The resulting image is shown in Fig.10.

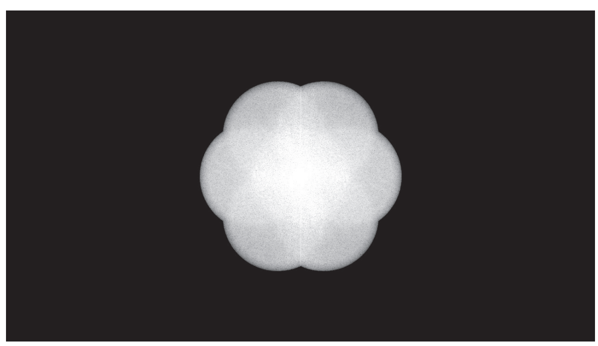

Fig. 8. Resulting image in spectral domain

Fig. 9. Resulting image in spatial domain

Fig. 10. Resulting image after postprocessing step

4. CONCLUSION

The main aim of this paper is to provide an overview of different techniques for image data preprocessing that is to be used as an input data for convolutional neural networks. Furthermore, one of the techniques, super resolution, is implemented, and obtained results are presented in form of images for each step in the algorithm. Super resolution was chosen because it goes beyond the light diffraction limit and collects high frequency information that can't be detected on input images. The super resolution algorithm is demonstrated for number of directions equal to three, and initial phase equal to zero. In our future work we will focus on the effects of image enhancing and preprocessing steps to the outcome of different neural networks. In that way, we will be able to compare different techniques and their effect on the resulting performance of neural networks.

REFERENCES

[1] M. Lybanon, S.M. Lea, and S.M. Himes, "Segmentation of diverse image types using opening and closing," Proceedings of 12th International Conference on Pattern Recognition, pp. 347-351, 9-13 Oct. 1994.

[2] G. Megha, "Morphological image proccesing," IJCST, vol. 2, issue. 4, pp. 161-165, 2011.

[3] L. Tomas, H.M. Guy, K. Pavel, S. Zdenek, F. Karel, and K. Milos, "Comparison of image reconstruction methods for structured illumination microscopy," Proceedings of SPIE - The International Society for Optical Engineering, vol. 9129, 2014.

[4] M. G. L. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," Journal of Microscopy, vol. 198, Pt. 2, pp. 82-87, 2000.

[5] L. Jin, B. Liu, F. Zhao, S. Hahn, B. Dong, R. Song, T. Elston, Y. Xu, and K. M. Hahn, "Deep learning enables structured illumination microscopy with low light levels and enhanced speed," Nature Communications, vol. 11, no.1, 2020.

[6] M. Müller, M. Viola, H. Simon, H. Wolfgang, and H. Thomas, "Open-source image reconstruction of super-resolution structured illumination microscopy data in ImageJ," Nature Communications, vol. 7, pp. 1-6, 2016.

[7] T. Raphaël, L.Yajie, T. Masashi, Z. Qinrong, L. Ziwei, K. Minoru, B. Eric, and L. Na, "Dynamic super-resolution structured illumination imaging in the living brain," Proceedings of the National Academy of Sciences, vol. 116, no. 19, pp. 9586-9591, 2019.

[8] H. Rainer, H. Thomas, "Super-resolution structured illumination microscopy," Chemical Reviews, vol. 117, no. 23, pp. 13890-13908, 2017.