TELKOMNIKA Telecommunication, Computing, Electronics and Control
Vol. 18, No. 3, June 2020, pp. 1522~1528
ISSN: 1693-6930, accredited First Grade by Kemenristekdikti, Decree No: 21/E/KPT/2018
DOI: 10.12928/TELKOMNIKA.v18i3.14839

Pulmonary rontgen classification to detect pneumonia disease using convolutional neural networks

Zuherman Rustam, Rivan Pratama Yuda, Hamimah Alatas, Chelvian Aroef
Department of Mahematics, University of Indonesia, Indonesia

Article Info

Article history:

Received Aug 15, 2019
Revised Jan 14, 2020
Accepted Feb 24, 2020

Keywords:

Classification
CNNs
Pneumonia
X-ray

ABSTRACT

Every organism is known to have different structural and biological system, specifically in human immunity. If the immune system weakens, the body is susceptible to disease especially pneumonia disease. Pneumonia disease is caused by the bacterium Streptococcus pneumonia, and according to the World Health Organization (WHO), it is identified as the leading cause of death in children worldwide, which is about 16%, for those under the age of 5. Meanwhile, someone who is predicted to have pneumonia by a doctor is recommended for an X-ray. Convolutional neural networks (CNNs) is an accurate method to help the doctor's predicted correctly. CNNs is divided into two important parts, feature extraction layer (convolutional layer and pooling layer) and fully connected layer. CNNs method is commonly used for image data classification. Therefore, CNNs is suitable to classify pneumonia based on lung X-ray in order to obtain accurate prediction results. And then, the results can be seen based on the graph of the accuracy value and the loss value. When CNNs method applied on the dataset, an accuracy rate of 97% was obtained. Based on accuracy rate, it shows that CNNs can be applied to image data (especially lung X-ray) for classification of pneumonia disease.

Corresponding Author:

Zuherman Rustam,
Department of Mahematics,
University of Indonesia, Depok, 16424, Indonesia.
Email: rustam@ui.ac.id

1. INTRODUCTION

Humans are living creature with the possession of a mind and heart, which serves as the distinguishing factor from other creatures [1]. Meanwhile, the body anatomy indicates the presence of numerous systems, organs and tissues, one of which is that for immunity [1], and while talking about it there is a certain relationship with health. Furthermore, it is better known as a system that exists to work in the direction of protection from outside influences, encompassing viruses or bacteria [1]. Therefore, instances where this defense is unable to fufil its function, there is a higher tendency of being affected by diseases, due to pathogen invasion. Thus, the initiation of abnormal condition in the body or mind, which is capable of causing discomfort, dysfunction or difficulty to the person, is called illness [1]. Furthermore, diseases occur in numerous types, ranging from infectious, non-communicable, mild, moderate and severe, and also, there is a tendency to attack all ages and groups, inluding toddlers, children, adults and elderly. In addition, they are numerous, causing very terrible conversation among the people as well as the biggest and most signigiant initiator of death. Therefore, disease is still a serious problem to be dealt with in every country, including Indonesia. This is indicated by the increasing mortality rate, based on the latest global health index, it is ranked 101 out of 149 countries according to the 2017 report of the legatum prosperity index. However, there is always a continuous strive and experiment to identify the right treatment for cures.

One of the dangerous diseases causing death worldwide, especially in Indonesia is pneumonia [2], commonly known as wet lung, whichi is infectious and attacks the lungs , hence, the air sacs become inflamed and swollen. This possesses the ability to attack all walks of life, although mortality is more dominant in infants and toddlers. According to the World Health Organization (WHO) it is the leading cause of children (infants and toddlers) mortality per week, which is about 16% of deaths at the age of under the age of 5. Based on the statement of the head of sub directorate (Kasubdit) of acute respiratory infection (ISPA), Indonesia is ranked 10th in the world, in cases of pneumonia killer virus attacks, with 15.5% at 2015, or the occurrence of about 554,650 cases, and its cause is often through various means, inluding viruses or bacteria. However, it is generally instigated by Streptococcus pneumonia, which is the major cause in infants, at about 50%, while the remaining is due to viruses and similar bacteria [1, 2].

Based on this, it possible to say that pneumonia is not an indiscriminate disease, and it also requires appropriate treatment to reduce the amount [2]. Therefore, on instances where there is an occurrence of symptoms, including shortness of breath, cough with phlegm or fever that does not heal, it is recommended to perform a health check, involving the doctor. Subsequently, there is a conduction of physical and other supporting examinations, to determine the patient's condition, and one of which is to perform an X-ray of the lungs. This promotes the doctor to provide the right diagnosis as to whether or not pneumonia is present, while clarification and assistance in classifying healthy and infected lungs, providing precise explanations and predictions, requires an accurate method. The accurate method that can be used is (CNNs), because this method is commonly used for image data classification [3]. Thus CNNs is suitable to classify pneumonia disease based on lung X-ray in order t o obtain accurate prediction results. This is evidenced by the previous studies using CNNs such as ImageNet Classification [4], Document Preprocessing [5], Face Detection [6], Image Segmentation [7], etc. Not only that, research about pneumonia using chest radiographs or X-ray has been done before by using several machine learning methods such as k-nearest neighbor (KNN), support vector machines (SVM) and Naïve Bayes [8]. Therefore, in this study CNNs method will be used to classify lung X-ray to detect pneumonia disease.

2. RESEARCH METHOD

2.1. Dataset

Dataset from patient lung X-ray were used in this experiment, obtained from Women and Children's Medical Center, Guangzhou. The number of image used were 330, where 150 images as healthy category, 112 images as bacterial pneumonia, and 68 images as viral pneumonia. Figures 1-3 are some examples of the patient's X-ray test from each category. Based on the figures, to find out the difference, X-ray affected by pneumonia more fade the color of the lung than normal. And to try dataset into program, in each category 80% of all images were used as training data and 20% for testing.

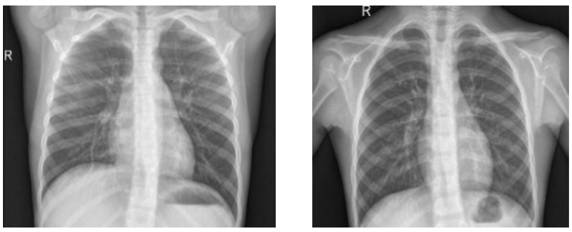

Figure 1. Normal lungs

Figure 2. Bacterial pneumonia

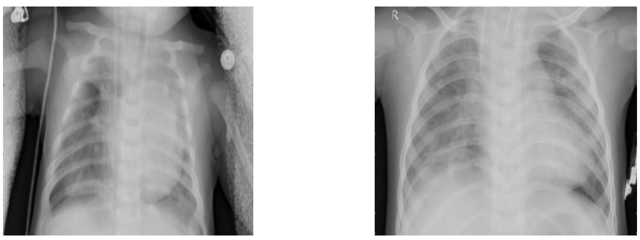

Figure 3. Viral pneumonia

2.2. Convolutional neural networks

The method used in this experiment is (CNNs), which is one type of Deep Neural Networks that is a result of the development of the multilayer perceptron (MLP) [3, 7]. The distinguish between CNNs and MLP is can retaining the capability of being used in the detection and recognition of objects in an image that are not just in the middle but MLP cannot be like that. CNNs will get better results than neural networks (NNs) because of the addition of one layer to CNNs after the input is called convolutional layer. Therefore, its development is CNNs, which are forms of neural network commonly used in image data. The component in CNNs are not much different from NNs, usually consist of neurons with weight, bias and activation function [3]. CNNs divided into 2 important parts, the first part is feature extraction (convolutional layer and pooling layer) and the second is fully-connected layer [3, 7, 9]. Illustration of CNNs can be seen in Figure 4.

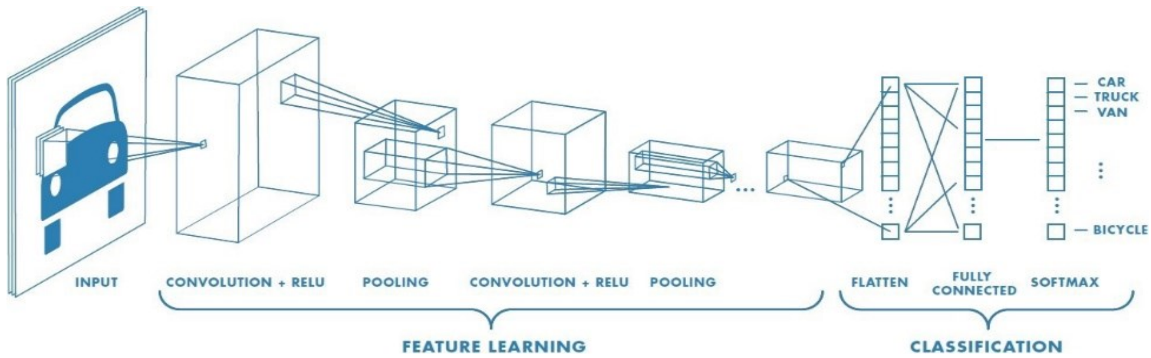

Figure 4. Illustration of convolutional neural network

2.2.1. Feature extraction layer

Feature extraction layer "encodes" an image into features that are in the form of representative numbers (feature extraction) [3]. Hence, CNNs is technically an architecture encompassing several stages, and each input and output consists of numerous arrays, termed feature maps, while the extraction layer individually comprises of two important parts, including 1) convolutional layer, which is major and most important for use, closely followed by 2) the pooling layer, combination used to extract the average or maximum value on the pixel, and each input possesses a different volume, and is represented based on depth, height and width [4, 9-11]. Therefore, the results obtained include the varied amount due to the fact that it is influenced by the filtering of the previous layers, and also the amount of filters used.

a. Convolutional layer

This is a stage that uses a convolution process through the utilization of filters inside, which has similar size as that of an image in general, including the length and height (pixels), as well as width [12, 13], as illustrated in Figure 5. Based on Figure 5, all parts of the image are shifted by the filters used, and from each position, numbers are generated, which are the results of "dot product" between the image inputs and filter values [14-16]. Therefore, the results are termed activation/feature maps as in the purple part.

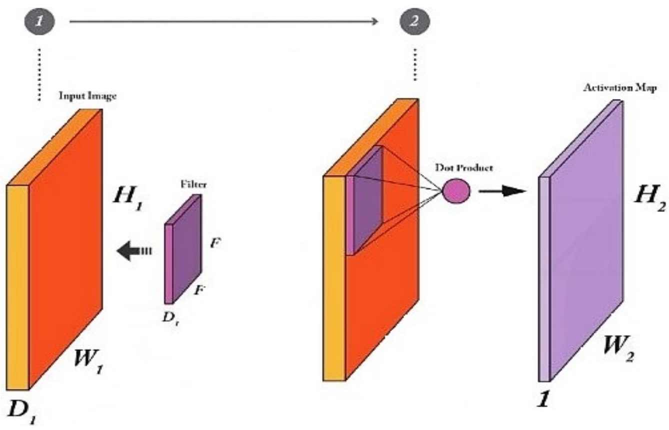

Figure 5. Illustration of convolutional layer

b. Stride

Stride is a parameter that determines the amount of shifts in the entire image [3, 17], and instances where the value used was 1, is an indication that the convolutiol filter shifted horizontally first by 1 pixel, followed by 1 for the vertical. However, the selection of step numbers or stride has no rules, and more information is obtained from an input, only while using a smaller stride. Meanwhile, if the selected one becomes smaller, the dot product calculation between the input and the value of the filter is performed more than the choice of a larger size while a small stride does not necessarily lead to a good performance [18].

c. Zero padding

Zero Padding is a parameter that is capable of determining the number of pixels that contain the value 0, added on each side of the image input [18, 19], which is used with an aim of manipulating the output dimensions of the convolutional layer (feature map).

d. An overview of feature extraction layer

In accordance with Figure 6, it is seen that the first layer in the feature extraction is an input image, with a size of 5x5x3, including 2 filters, and for example the use of 2 strides and a single zero padding. Therefore, the 5x5x3 means 5 pixels in length, 5 pixels in height and 3 pieces in thickness, and the illustration as shown in Figure 6 indicates that the dimensions of the actual input are 5x5, on instances where convolution is conducted with a 3x3 filter, and a stride of 2, thus, a feature map with 2x2 size is obtained. Meanwhile, if zero padding is added as much as 1, a 3x3 size is obtained, and more information is further generated. Consider the illustration as shown in Figure 6. To calculate the dimensions of a feature map, the (1) is used [3, 20]:

$$Output = \frac{W - N + 2P}{2S} + 1 \tag{1}$$

where,
W= long or high of input
N= long or high of filter
P= as a padding
S= as a stride
for example, based on Figure 3, with W = 5 N = 3 P = 1 S = 2, the output dimension obtained is 3.

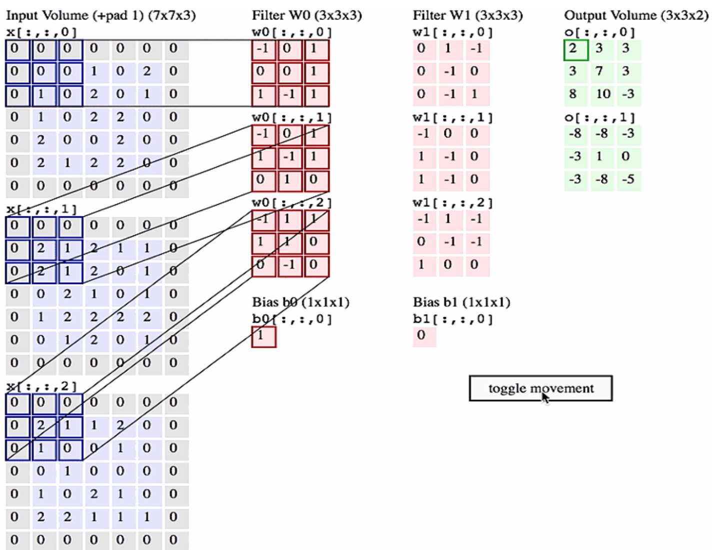

Figure 6. Illustration of feature map

e. Pooling layer

The pooling layer is usually located after the convolutional [21, 22], and the point is that it consists of a certain sized filter, and strides that shifts throughout the feature map area. Meanwhile, the pooling process commonly used includes (1) max pooling, where the use of 2x2 with stride 2 leads to the maximum value of 2x2 pixel area selected for each shift, (2) average pooling, which chooses its mean value, and an illustration of pooling layer (especially max pooling) are shown in Figure 7. From the illustration with a 2x2 filter used, the filter will move in the input area with a strides of 2 and take the maximum value of each filter movement towards the input. The output dimension of the pooling layer also uses the same formula as the convolutional, and the reason for use is to reduce the dimensions of the feature map (down sampling), thus, speeding up computation, due to the presence of fewer parameters the require an updated and overfitting issues.

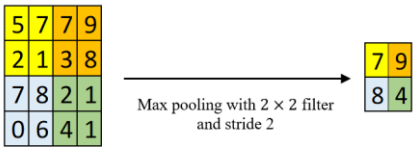

Figure 7. Illustration of max pooling

2.2.2. Fully connected layer

The feature map is created based on the feature extraction layer, which is still a multidimensional array, hence, there is need to flatten or reshape it in the vector feature map, which is used as input from the fully-connected layer [23, 24]. In addition, all that activate neurons from the previous are linked in the next layer, as seen in neural networks. Therefore, in order to connect properly, individual activation of the previous ought to be first converted into one-dimensional data. These usually use a method termed multi-perceptron layer, which has a target in processing data, in order for it to carry out proper classification [25]. Meanwhile, the difference against convolution layers is that the neurons in the latter are connected to a specific input area, while the fully-connected occurs in almost all parts. However, both continue to perform “dot product” operations, thus, their functions are not significantly different.

2.3. Model

The method applied is CNNs within Tensorflow, which is a framework of programs or packages that assist the systems to process image data. Meanwhile, the chosen model indicates the use of 4 convolutional layers, including:

– First layer use 16-3 x 3 filters
– Second layer use 32-3 x 3 filters
– Third layer use 64-3 x 3 filters
– Fourth layer use 128-3 x 3 filters

Each of these involve max-pooling of 2x2, with amount of stride being 2 and input has a size of 28×28 pixels with a thickness of 1. Therefore, because this study will be classified into three classes, the categorical cross-entropy loss function will be used and Adam (Adaptive Momentum) as an optimizer with a learning rate of 0.0001. The model of CNNs is explained in the form of a flowchart as shown in Figure 8.

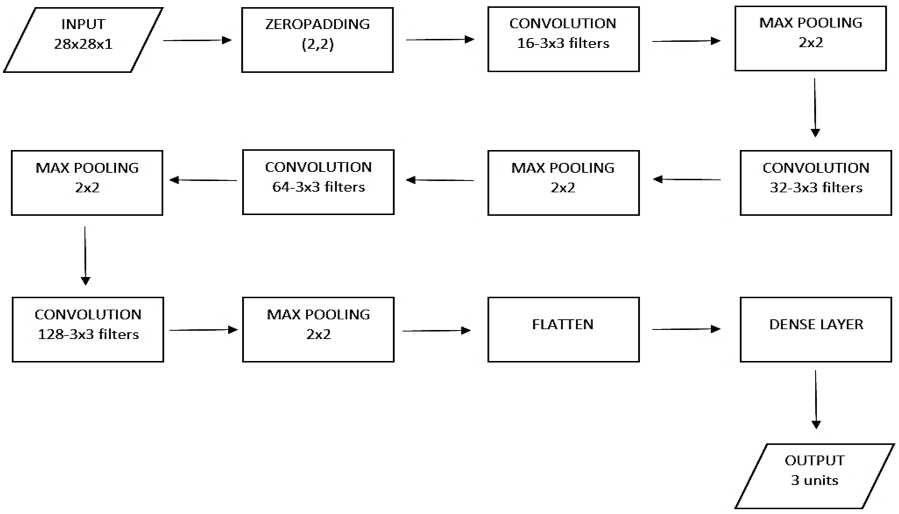

Figure 8. Model of CNNs

3. RESULTS AND DISCUSSIONS

This study used software Python version 2.1.9 for convolutional neural networks. After running the program, accuracy value and loss value will be generated in graphical form. The image data that has been obtained as explained in the dataset section, 330 were classified into 3 classes, with the output as accuracy value as shown in Figure 9 and loss value as shown in Figure 10. The accuracy value and the loss value are used to see how capable the CNNs model that has been formed to classify pneumonia disease.

Based on Figure 9, the accuracy value has increased in epochs to 0-25, which means the model is trying to memorize data. But after 25-50 epochs the accuracy value starts to stagnant with training accuracy converge to 1.00 and test accuracy converge to 0.91. Different with Figure 10, loss training decreases and converges to zero while the loss test continues to increase. Furthermore, test accuracy and test lost shows from Figure 9 and Figure 10, a sign of overfitting. But based on the results of previous studies the accuracy value obtained in detection of pneumonia disease based on X-ray images using the SVM method is 77%, KNN method is 70% and Naïve Bayes is 68% [8]. And when compared with the results of this study or using the CNNs method produces better accuracy values than three methods, which is 91%.

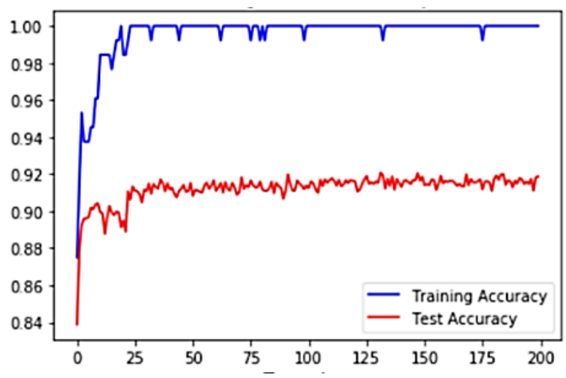

Figure 9. Accuracy value

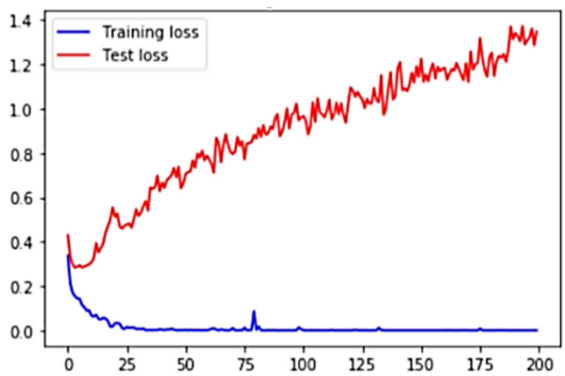

Figure 10. Loss value

4. CONCLUSION

Based on results and discussions, the convolutional neural networks (CNNs) method is able to classify X-ray images into healthy lungs and pneumonia, due to bacteria and viruses. So that this method is able to assist doctors in providing the appropriate beliefs and predictions to patients. In addition, it is evidenced in the evaluation obtained after epochs 25, for training accuracy is 0.97 and for test accuracy is 0.91. This evidenced shows the capability for the program from CNNs method to accurately identify the patient's X-ray test images. However, the accuracy rate at some epochs is identified as unstable, due to the fact that the attempt of CNNs to memorize the data set image. The limitation of this study lies in the image dataset because that is not taken in large quantities and the program of CNNs doesn't use a good processor. So, for the next researcher can modify or adding high processor to speed up the program of CNNs.

ACKNOWLEDGEMENTS

This research was full supported financially by University of Indonesia , with a DRPM PUTI KI Q2 2020 research grant scheme.

REFERENCES

[1] Urry L. A., et al, "Campbell biology," Eleventh Edition New York: Pearson, pp. 975-998, 2016.
[2] Kermany D., Zhang K,. and Goldbaum M., "Labeled optical coherence tomography (OCT) and chest X-Ray images for classification," *Mendeley Data,* vol. 2, 2018.
[3] O'Shea K., and Nash R., "An introduction to convolutional neural networks," *arXiv*, Preprint arxiv:1511.08458v2, 2015.
[4] Krizhevsky A., Sutskever I. and Hinton G. E., "ImageNet classification with deep convolutional neural networks," *Advances Neural Information Processing Systems,* pp. 1907-1105, 2012.
[5] Chellapilla K., Puri S., and Simard P. Y., "High-Performance convolutional neural networks for document Processing," *Proc. of the 10th Int. Workshop on Frontiers in Handwriting Recognition,* 2006.
[6] Li H., et al., "A convolutional neural network cascade for face detection," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* pp. 5325-5334, 2015.
[7] Turaga S. C., et al., "Convolutional networks can learn to generate affinity graphs for image segmentation," *Neural Computation,* vol. 22, no. 2, pp. 511–538, 2010.
[8] Sousa R. T., et al., "Comparative performance analysis of machine learning classifiers in detection of childhood pneumonia using chest radiographs," *Procedia Computer Science,* pp. 2579-2582, 2013.
[9] Ciresan D., Meier U., and Schmidhuber J., "Multi-column deep neural networks for image classification," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* pp. 3642–3649, 2012.
[10] Ciresan D. C., et al., "High performance convolutional neural networks for image classification," *IJCAI Proceedings-International Joint Conference on Artificial Intelligence,* vol. 22, pp. 1237, 2011.
[11] Ciresan D. C., et al., "Convolutional neural network committees for handwritten character classification," *IEEE Conference on Document Analysis and Recognition,* pp. 1135–1139, 2011.
[12] Egmont-Petersen M., de. Ridder D., and Handels H., "Image processing with neural networks a review," *Pattern recognition,* vol. 35, no. 10, pp. 2279–2301, 2002.
[13] Ji S., et al., "3D convolutional neural networks for human action recognition," *Pattern Analysis and Machine Intelligence,* vol. 35, no. 1, pp. 221–231, 2013.
[14] Farabet C., et al., "Hardware accelerated convolutional neural networks for synthetic vision systems," *Proceedings of 2010 IEEE International Symposium on Circuits and Systems,* pp. 257–260, 2010.
[15] Hinton G. E., et al., "Improving neural networks by preventing co-adaptation of feature detectors," *arXiv*, Preprint arXiv:1207.0580, 2012.
[16] Karpathy A, et al., "Large scale video classification with convolutional neural networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* pp. 1725–1732, 2014.
[17] Tuti P., Imania A. A., and Pertiwi B. U., "Image classification of golek puppet images using convolutional neural networks algorithm," *International Journal of Advances,* vol. 11, pp. 34-45, 2019.
[18] Muhamad Y., Budhi I., and Casi S., "Application of transfer learning using convolutional neural network method for early detection of terry's nail," *Journal of Physics: Conference Series,* vol. 1201, 2019.
[19] Simard P. Y., Steinkraus D. and Platt J., "Best practice for convolutional neural networks applied to visual document analysis," *International Conference on Document Analysis and Recognition,* pp. 958-962, 2003.
[20] LeCun Y., et al., "Gradient-Based learning applied to document recognition," *Proceedings of the IEEE,* pp. 2278-2324, 1998.
[21] Steinkraus D., Buck I. and Simard P. Y., "GPUs for machine learning algorithms," *Eighth International Conference on Document Analysis and Recognition,* pp. 1115-1120, 2005.
[22] Wang L., *et al.,* "Training deeper convolutional networks with deep supervision," arXiv Preprint arXiv:1505.02496, 2015.
[23] Szarvas M., et al., "Pedestrian detection with convolutional neural networks," *Intelligent Vehicles Symposium, IEEE,* pp. 224–229, 2005.
[24] Szegedy C., Toshev A., and Erhan D., "Deep neural networks for object detection," *Advances in Neural Information Processing Systems,* pp. 2553–2561, 2013.
[25] Tivive F. H. C., and Bouzerdoum A., "A new class of convolutional neural networks (siconnets) and their application of face detection," *Proceedings of the International Joint Conference,* pp. 2157–2162, 2003.