TELKOMNIKA, Vol.17, No.2, April 2019, pp.653~659
ISSN: 1693-6930, accredited First Grade by Kemenristekdikti, Decree No: 21/E/KPT/2018
DOI: 10.12928/TELKOMNIKA.v17i2.8976

Image forgery detection using error level analysis and deep learning

Ida Bagus Kresna Sudiatmika*, Fathur Rahman, Trisno, Suyoto
Magister Teknik Informatika, Universitas Atma Jaya Yogyakarta, Indonesia
*Corresponding author, e-mail: ibkresnasudiatmika@gmail.com

Abstract

Many images are spread in the virtual world of social media. With the many editing software that allows so there is no doubt that many forgery images. By forensic the image using Error Level Analysis to find out the compression ratio between the original image and the fake image, because the original image compression and fake images are different. In addition to knowing whether the image is genuine or fake can analyze the metadata of the image, but the possibility of metadata can be changed. In this case the authors apply Deep Learning to recognize images of manipulations through the dataset of a fake image and original images via Error Level Analysis on each image and supporting parameters for error rate analysis. The result of our experiment is that we get the best accuracy of training 92.2% and 88.46% validation by going through 100 epoch.

Keywords*: convolutional neural network, deep learning, error level analysis, image forensic, image forgery*

1. Introduction

As the rapid growth of technology makes it easier for someone to spread the word, as well as spreading fake images. With so many software that can be used to manipulate the image so as to facilitate the public in manipulating the image. With the spread of fake images on social media that can reap the controversy so that the image forensic to test the truth of the image. Generally, image forensic is a field of the study identifying the origin and verifying the authenticity of the image.

With so many false images which spread across the Internet and social media, hence the need for a tool to help people determine whether the image spread is real or fake pictures. Many methods are used to determine the level of authenticity of the picture, one with determining the quality of the image compression level results. In this research, the methods used to measure the level of compression is using Error Level Analysis (ELA).

Error Level Analysis (ELA) is a forensic technique on the image to analyze images through different levels of compression. This technique is used to find out digitally modified images. To define forgery images and original images, many approaches are done. There are various techniques performed by researchers in this case. Hites C Patel et al. in their research *Forgery Frame Detection From The Video Using Error Level Analysis*. By analyzing the number of frames and comparing the original video frames with the fake ones. Through some attributes that are analyzed like Time length, Frame Rate, no. of a frame, Data Rate, Resolution, Bit Rate Total Bit Rate, Audio Chanel, Audio Sample Rate, Protected, Video quality, Camera Base Editing Video [1]. Meera Mary Isaac et al. doing image forgery detection using Gabor Wavelets dan Local Phase Quantization. By using CASIA TIDE v.1 Dataset [2]. Birajfar et al. using a passive technique method in analyzing false images. In his research summarizes some research that does image forgery [3].

Youseph et al. using the illuminant color Estimation method, by combining the canny detection and HOG edge descriptor to get the edge border of the image. Later on training using SVM with 74% accuracy value [4]. Mohhamad F.H using an efficient and robust method combining un-decimated wavelet transform and scale invariant feature transform and judging from precision, recall, false positive rate [5]. Jie Zhao et al. analyze image forgery using DCT and SVD algorithm analyze based on DAR and FPR [6]. A. Dixit et al. reviewing some of the studies discussing image forgery [7]. In his research, he summarizes some methods of image forgery and its application. Wu-Chih Hu analyze image forgery based on image watermarking

Received February 16, 2018; Revised November 13, 2018; Accepted December 12, 2018

and alpha mattes analysis [8]. Ghulam Muhammad et al. in his research to analyze image forgery using dyadic wavelet transform [9]. Ashwini V Malviya et al. using Auto Color Correlogram on analyzing image forgery [10]. Rani Susan Oommen used Fractal Dimension and singular values in analyzing original image and fake image [11].

In this research, we will implement a new system that can distinguish forgery image and original images with deep learning. Deep learning is a new science in machine learning that recently developed due to the development of GPU acceleration technology. Deep learning methods applied in the introduction of false images and the original image is Convolutional Neural Network (CNN). CNN is the development of a multilayer perceptron (MLP) designed to process two-dimensional data. We choose CNN in this research because CNN does not require image processing before processing by a neural network. We are doing this research to help the community in distinguishing real images and forgery images that are widely circulated in social media. Besides testing against CNN and obtain better results from research that have been studied previously.

2. Research Method

The dataset we get is through CASIA version 2.0. Inside there are 7491 original images and 5123 tampered images. The size of the dataset is changed to 224x224 pixels. In this experiment, we divide the dataset into two namely training set and test set. In the range 50-90% for the training set and the rest is used for test data. In compiling the dataset, we divide the data train and test data each of which there are 2 categories, namely the category of fake images and the original image.

The first step we took was to divide the dataset from Casia V.2 into 2 categories: original and fake images. We normalize the image by processing the image to a size of 224x224 pixels. Then our next step is to perform analysis on the level of compression error image, from the compression result then we use the VGG 16 architecture for CNN in recognizing the original image and fake images according to the ELA. Our next step is to summarize the results of the training. Our proposed method described on flowchart as shown in Figure 1.

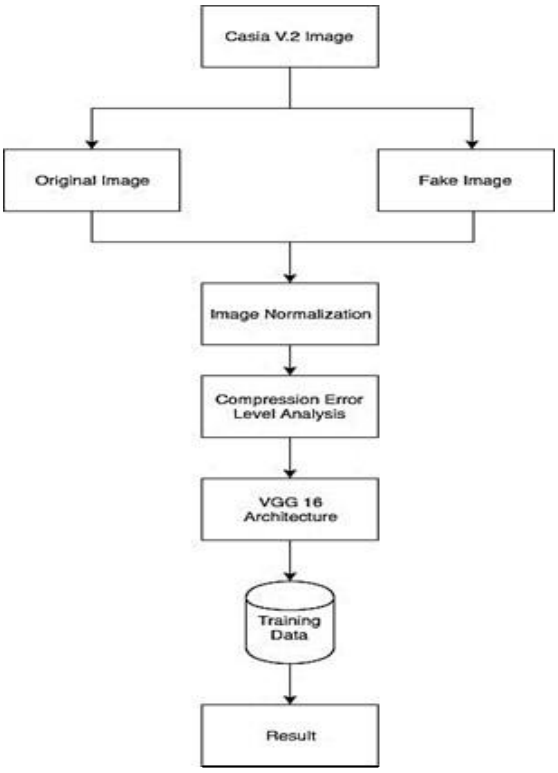

Figure 1. Proposed method

2.1. Error Level Analysis

Error level analysis is one technique for knowing images that have been manipulated by storing images at a certain quality level and then calculating the difference from the compression level [12]. When JPEG was first saved, then it will compress the image the first time, most editing software like adobe photoshop, gimp, and adobe lightroom support JPEG compressing operation. If the image is rescheduled using image editing software, then compressed again.

So it shows that the original image when the first image is taken using a digital camera has been compressed twice, first use the camera and the second is editing software. When viewed with the naked eye the image looks the same, but by using this method it will look the difference between a forgery image with the original image. Calculation for the average difference of the quantization table Y (luminance) and CrCb (Chrominance). The digital camera does not optimize the image for a specified camera quality level (high, medium, low, etc.). Original images from digital cameras should have high ELA values. Each subsequent resave will decrease the potential error rate. Original images from photography have high ELA values shown through white on the ELA image, as shown in Figure 2. When the image is resaved, using ordinary human vision does not show a significant degree of difference, but ELA shows the dominant black and dark colors. If this image is resaved again it will decrease the image quality. If the original image is then modified, ELA will show the modified area has a color with a higher ELA level. The Figure 2 describes how the output of ELA on the condition of the image.

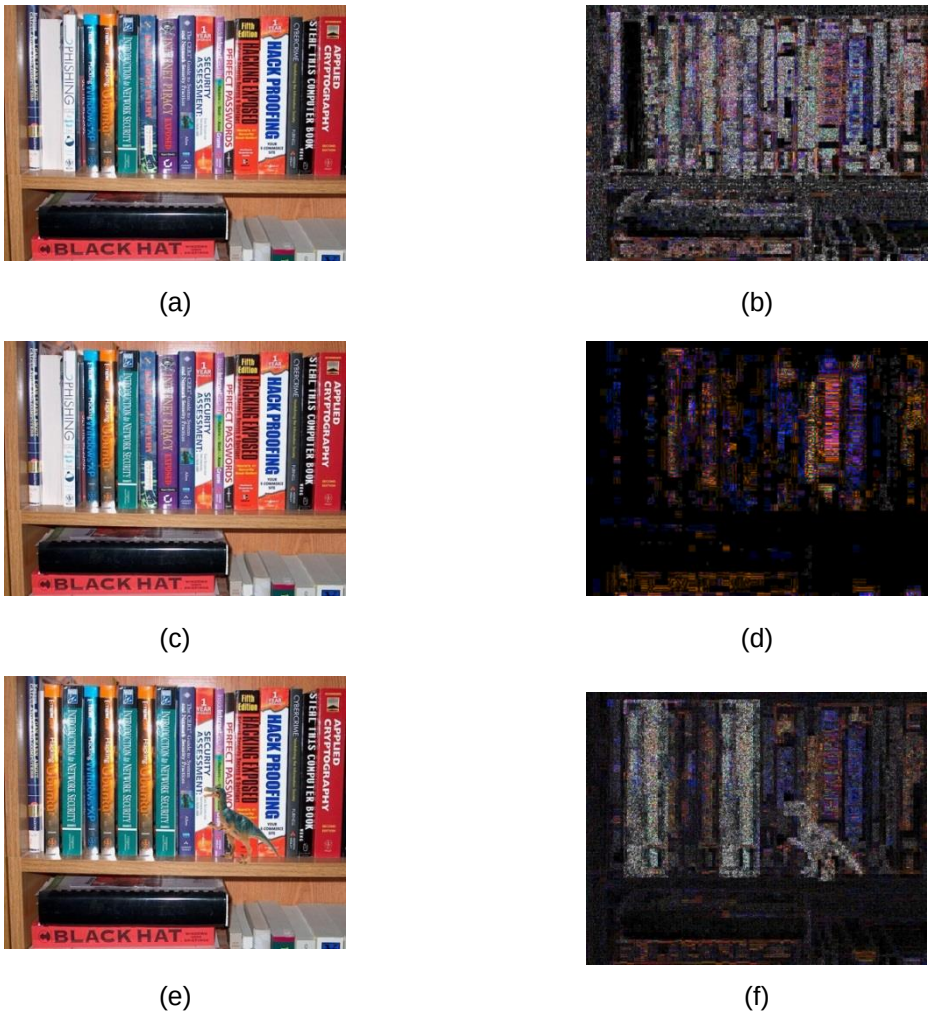

Figure 2. Error level analysis compression: (a) original image, (b) ELA original Image, (c) resave image, (d) ELA resave image, (e) tampered image, (d) ELA tampered image

Due to the insufficient number of datasets, the performance of CNN will not work optimally. Given this problem, we use VGG16 architecture to do the training. VGG16 has been proven in pattern recognition with a small number of datasets that have been demonstrated in the imageNet competition.

2.2. Convolutional Neural Network

The convolutional neural network was originally proposed by LeCun et al. for handwritten recognition has been successful in image identification, detection, and segmentation of the image [13]. CNN has a high ability in large-scale image classification. Cnn consists of three layers: convolutional layer, pooling layers, and fully connection layers [14]. A Convolutional layer and pooling layer is the most important layer on CNN. Convolutional layer is used for extract feature by combining the image area with many filters. Pooling layer reduces the size of the output map of the convolution layer and prevents overfitting. Through these two layers the number of neurons, parameters, and connections is much less than there is a CNN model. This makes CNN more efficient compared with BP networks with similar layers. The final formula of the single output image channel of the convolution layer as (1):

$$conv(I,K)_{xy} = \sigma\left(b + \sum_{i=1}^{h}\sum_{j=1}^{w}\sum_{k=1}^{d} K_{ijk} \cdot I_{x+i-1,y+j-1,k}\right) \quad (1)$$

The layers are determined by specific kernels, K along with the bias value (b) on each kernel. It then operates by calculating the output image of the previous layer with each of the kernels. Convolution is a mathematical term that means applying a function to the output of another function repeatedly. The kernel moves from the top left corner to the bottom right. So the result of the convolution of the image can be seen in the picture on the right. The goal of convolution in image data is to extract features from the input image. The convolution will produce a linear transformation of the input data according to the spatial information in the data. A very popular approach to downsampling is to use pooling layers. Pooling layer usually deciphers the image (like 2x2) in the aggregation into a single unit. The most popular scheme for aggregation is the incorporation of the maximum value (max-pooling).

Subsampling is the process of reducing the size of the image data. In image processing, subsampling also aims to increase the position invariance of features. In most CNN, the subsampling method used is max pooling. Max pooling divides the output from the convolution layer into several small grids and then takes the maximum value of each grid to construct a reduced image matrix. The red, green, yellow and blue grids are the grid to be selected for maximum value. So the results of the process can be seen on the grid set on the right. The process ensures that the features obtained will be the same even if the image object is translating (shifting). The formula of max-pooling is as follows (2):

$$Yi^{l+1}, j^{l+1}, d = \frac{max}{0 \le i < H, 0 \le j < W} \mathcal{X}^{l}_{i^{l+1} \times H + i, j^{l+1} \times W + j, d} \quad (2)$$

Layer is a layer that is usually used in the application of MLP and aims to transform the dimensions of data so that data can be classified in a linear [15]. Each neuron in the convolution layer needs to be transformed into one dimensional data first before it can be inserted into a fully connected layer. Because it causes data to lose spatial information and not reversible, fully connected layer can only be implemented at the end of the network. Applying CNN for fake image classification and original image converted into error level form on image. We know through the previous literature that CNN can achieve competitive performance and even better than humans in some visual problems, and we wanted to test CNN's ability to classify forgery image and original images via Error Level Analysis.

The description of VGG is based on the CNN VGG-Very-Deep-16 architecture. VGG16 which was first tested on imageNet 2012 [16]. Because the structure is very deep, VGG16 has gained recognition precision is very promising, therefore we use VGG16 as a pre-training model for original image and forgery images recognition. A network consisting of convolutional, pool, and fully connected (FC) [17]. Convolutional layer uses 3-dimensional filter. While layer pool layer does subsampling with factor 2. The VGG16 architecture is shown in Figure 3. Table 1 described VGG16 architecture parameters.

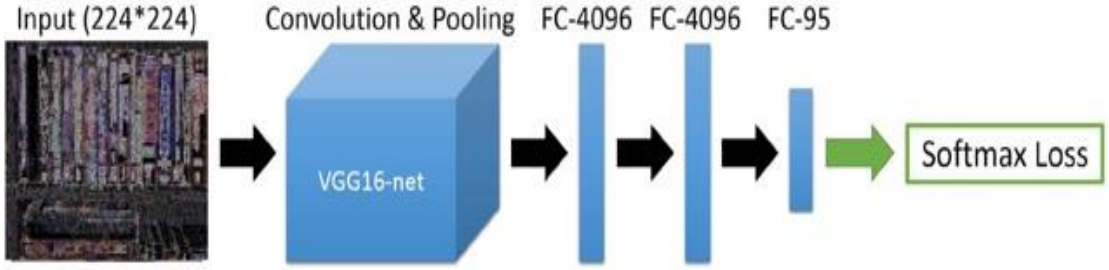

Figure 1. VGG architecture

Table 1. VGG-16 Architecture Parameter

Layer	Volume	Parameter	Layer	Volume	Parameter
Input	224×224×3	0	CONV3-512	28×28×512	(3*3*256)*512=1,179,648
CONV3-64	224×224×64	(3*3*3)*64=1,728	CONV3-512	28×28×512	(3*3*512)*512=2,359,296
CONV3-64	224×224×64	(3*3*64)*64=36,864	CONV3-512	28×28×512	(3*3*512)*512=2,359,296
POOL2	112×112×64	0	POOL2	14×14×512	0
CONV3-128	112×112×128	(3*3*64)*128=73,728	CONV3-512	14×14×512	(3*3*512)*512=2,359,296
CONV3-128	112×112×128	(3*3*128)*128=147,456	CONV3-512	14×14×512	(3 3*512)*512=2,359,296
POOL2	56×56×128	0	CONV3-512	14×14×512	(3 *3*512)*512=2,359,296
CONV3-256	56×56×256	(3*3*128)*256=294,912	POOL2	7×7×512	0
CONV3-256	56×56×256	(3*3*256)*256=589,824	FC-1	1×1×4096	7*7*512*4096=102,760,448
CONV3-256	56×56×256	(3*3*256)*256=589,824	FC-2	1×1×4096	4096 * 4096 = 16,777,216
POOL2	28×28×256	0	FC-3	1×1×2622	4096 * 2622 = 10,739,712

Here we describe the approach on recognition of forgery images and original images through the ELA on the picture in terms of pre-processing images, feature extraction using CNN, and functional block classification. Image Processing is about the normalization of the next image in terms of size. The feature extraction here utilizes the convolution layer on CNN in getting the feature on the image. As shown by the picture, filter sliding over on the whole picture. With this technique, it will take the maximum value as a characteristic of the pixel and written on (1, 1) in the output layer [18]. When the stride is worth 1, this means the filter will move one pixel to the right and perform the same operation as described previously. After performing the operation in one line, then the filter will move on the next line to process the whole image. In the research method, we use the architecture of VGG-16 as our classification technique. It starts with an image input with dimensions 224x224x3 (224 stating height and width, 3 stating the depth (RGB)) [19, 20].

This architecture uses 3x3 filter on convolutional layer and 2x2 on pooling layer, then continued with 3 fully-connected layers. In each convolutional layer forward it to the activation function. The Recified Linear Unit (ReLU) is used as an activation function that removes non-linear values from process data, as in (3).

$$f(x) = max(x, 0) \tag{3}$$

Data is divided into training data and data testing (validation). The Feature that has been recognized in training data will be comparable with feature on test data. Then, the fully-connected layer will do pattern recognition training and provide a prediction of per class based on the input layer. The softmax function squashes the outputs of each unit to be between 0 and 1, just like a sigmoid function. Mathematically, softmax activation is shown as (4) [21-22].

$$\sigma(\mathcal{Z})_j = \frac{e^{\mathcal{Z}_j}}{\sum_{k=1}^{K} e^{\mathcal{Z}_k}} \tag{4}$$

In forensic photos, reading metadata is very important to do to know the level of authenticity of image data [23, 24], we do this analysis to compare whether the results of the test data on CNN are consistent with the metadata. In this case, we analyze the picture if found tags like Photoshop, Gimp, Adobe etc. most likely the image tampered [25, 26].

3. Results and Discussion

In this section, we will describe the experimental results from the recognition of the original image and the fake image. We analyze the accuracy percentage of the drawing training. Percentage of training varies from 60-90%. This shows that the method we use is able to study the data despite the small amount of data. From the training we have done, then we get results based Figure 4.

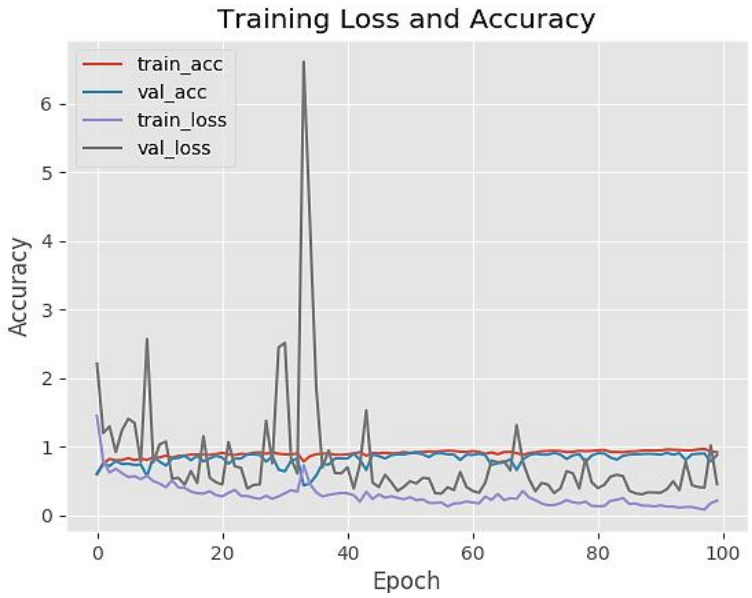

Figure 4. Model accuracy vs model loss image forgery detection

From the picture above that known with training accuracy of the model achieved up to 92.2% and for validation 88.46% using 100 epoch. Thus, by using the deep learning architecture of VGG 16 in analyzing error level image analysis for image forgery can be applied and get good results on recognition.

4. Conclusion

In this paper, we have solved the problem of distinguishing real images and forgery images using deep learning. We propose a new system from combination Error Level Analysis and Convolutional Neural Network in machine learning and computer vision to solve the problems above. First, we divide the dataset into tampered images and original images, then we determine the architecture that will be used to train the recognition. We chose to use VGG 16 in this training because VGG is perfect for training with minimal datasets. The result of our experiment is that we get the best accuracy of training 92.2% and 88.46% validation by going through 100 epoch. In our next study, we will conduct a CNN architecture variant to get the best accuracy and do other approaches in processing image processing to recognize the original image and forgery image.

References

[1] Patel HC, Patel MM. Forgery Frame Detection From The Video Using Error Level Analysis. *IJAERD*. 2015; (6): 242–247.
[2] Isaac MM, Wilscy M. Image Forgery Detection Based on Gabor Wavelets and Local Phase Quantization. *Procedia Computer Science*. 2015; 58: 76–83.
[3] Birajdar GK, Mankar VH. Digital image forgery detection using passive techniques: A survey. *Digital Investigation*. 2013; 10(3): 226–45.
[4] Youseph SN, Cherian RR. Pixel and Edge Based Illuminant Color Estimation for Image Forgery Detection. *Procedia Computer Science*. 2015; 46: 1635–42.

[5] Hashmi MF, Anand V, Keskar AG. Copy-move Image Forgery Detection Using an Efficient and Robust Method Combining Un-decimated Wavelet Transform and Scale Invariant Feature Transform. *AASRI Procedia*. 2014; 9: 84–91.
[6] Zhao J, Guo J. Passive forensics for copy-move image forgery using a method based on DCT and SVD. *Forensic Science International*. 2013; 233(1–3): 158–66.
[7] Hashmi MF, Anand V, Keskar AG. Copy-move Image Forgery Detection Using an Efficient and Robust Method Combining Un-decimated Wavelet Transform and Scale Invariant Feature Transform. *AASRI Procedia*. 2014; 9: 84–91.
[8] Hu WC, Chen WH, Huang DY, Yang CY. Effective image forgery detection of tampered foreground or background image based on image watermarking and alpha mattes. *Multimedia Tools and Applications*. 2015; 75(6): 3495–516.
[9] Muhammad G, Hussain M, Bebis G. Passive copy move image forgery detection using undecimated dyadic wavelet transform. *Digital Investigation*. 2012; 9(1): 49–57.
[10] Malviya AV, Ladhake SA. Pixel Based Image Forensic Technique for Copy-move Forgery Detection Using Auto Color Correlogram. *Procedia Computer Science*. 2016; 79: 383–90.
[11] Oommen RS, Jayamohan M, Sruthy S. Using Fractal Dimension and Singular Values for Image Forgery Detection and Localization. *Procedia Technology*. 2016; 24: 1452–9.
[12] Jeronymo DC, Borges YCC, Coelho L dos S. Image forgery detection by semi-automatic wavelet soft-Thresholding with error level analysis. *Expert Systems with Applications*. 2017; 85: 348–56.
[13] LeCun, Y. et al. *Comparison of learning algorithms for handwritten digit recognition*. International conference on Artificial Neural networks. France. 1995: 53–60.
[14] Kuo C-CJ. Understanding convolutional neural networks with a mathematical model. Journal of *Visual Communication and Image Representation*. 2016; 41: 406–13.
[15] Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, et al. Recent advances in convolutional neural networks. *Pattern Recognition*. 2018; 77: 354–77.
[16] Kim J, Sangjun O, Kim Y, Lee M. Convolutional Neural Network with Biologically Inspired Retinal Structure. *Procedia Computer Science*. 2016; 88: 145–54.
[17] Liu S, Deng W. *Very deep convolutional neural network based image classification using small training sample size*. 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR). 2015.
[18] Kumar PR, Srikanth C, Sailaja KL. Location Identification of the Individual based on Image Metadata. *Procedia Computer Science*. 2016; 85: 451–4.
[19] Pratt H, Coenen F, Broadbent DM, Harding SP, Zheng Y. Convolutional Neural Networks for Diabetic Retinopathy. *Procedia Computer Science*. 2016; 90: 200–5.
[20] Schmidhuber J. Deep learning in neural networks: An overview. *Neural Networks*. 2015; 61: 85–117.
[21] Teddy SG, Siti AMH, Mira K, Ismail N, Nor FZ, Anis NN. Development of photo forensics Algorithm by detecting photoshop manipulation using Error Level Analisis. *IJEECS*. 2017; 7: 131-137.
[22] Seo, Y, Shin, K,. Hierarchical convolutional neural networks for fashion image classification. *Expert Systems with Applications*. 2019: 328–339.
[23] Ou, J, Li, Y. Vector-kernel convolutional neural networks. *Neurocomputing.* 2019; 330: 253–258.
[24] Fu, Y, Aldrich, C. Flotation froth image recognition with convolutional neural networks. *Minerals Engineering.* 2019; 132: 183–190.
[25] Olivas-Padilla, BE, Chacon-Murguia, MI. Classification of multiple motor imagery using deep convolutional neural networks and spatial filters. *Applied Soft Computing*. 2019; 75: 461–472.
[26] Han, H, Li, Y, Zhu, X. Convolutional neural network learning for generic data classification. *Information Sciences*. 2019; 447: 448–465.