1. Abstract

Images are often manipulated with the intent and purpose of benefiting one party. In fact, images are often considered as evidence of a fact or reality, therefore, fake news or any form of publication that uses images that have been manipulated in such a way has greater capability and potential to mislead. To detect image forgeries, large amounts of image data are needed, and a model that can process every pixel in the image. Apart from that, efficiency and flexibility in training data are also needed to support its use in everyday life. The concept of big data and deep learning is the right solution to this problem. Therefore, with a Convolutional Neural Network (CNN) architecture that utilizes Error Level Analysis (ELA), image forgery detection can reach 91.83% and convergence in only 9 epochs.

2. Literature Review

According to the EU High Level Expert Group (2018), fake news is defined as disinformation, namely any form of inaccurate, false or misleading information that is presented, promoted or designed. Behind fake news, there are several reasons for these publications. One of them is to gain economic benefits, whether through increasing the number of news clicks or creating news that is not supposed to benefit one party.

In addition, fake news can also affect stock prices, which can provide an advantage to the party who released the news. Another reason is to gain support or bring down other parties socially or politically.

Based on statistics courtesy of the Telematics Society Indonesia (MASTEL) in 2017, the types of fake news most frequently received were socio-political, SARA (ethnicity, religion and race), health, food and drink, financial fraud, and science and technology. As many as 84.5% of all respondents stated that they felt disturbed by fake news, and more than 70% agreed that fake news disturbs social harmony and hinders development.

Apart from written form, around 40% of respondents stated that the spread of fake news was also often accompanied by images. Images are used by humans to reproduce reality, and are often used as evidence of news, publications, or facts. Fake news that has supporting images tends to be accepted and trusted by the public.

In general, humans find it easier to remember images than writing. According to the Social Science Research Network, as many as 65% of humans are people who enjoy learning through visuals. In marketing and visual science, it is said that images have a very big influence on an article.

People tend to respond when there are images rather than just writing. According to an infographic with a theme about the influence of images in the world of marketing, images can increase the number of respondents to an article by up to 94%. Therefore, an image is a strong element in disseminating information.

To determine whether an image is real or fake, it is very difficult to see with the naked eye, special techniques and certain precision are needed in order to know for sure whether an image is a real image or has undergone modification. For ordinary people, this may be difficult to do. For this reason, image forgery detection technology needs to be developed, so that it can be used as a means to help people determine the authenticity of an image.

This technology requires a lot of image data, and each image has many pixels that make up it. With regular machine learning, this technology would be difficult to develop. So, big data and deep learning is the right solution to solve this problem of image forgery detection.

3. Methodology

Data mining in the form of image forgery detection has two main objectives as follows. Convolutional Neural Network (CNN) By having a reference for the public to find out whether an image is real or not, it will certainly reduce the anxiety caused by fake images:

1. Proposing a new method using deep learning to classify images as original images and images that have undergone modifications with a simpler architecture, so that computational costs can be reduced.
2. Involving the use of ELA in machine learning as an effort to increase efficiency.

There is some impetus behind these two main goals. As is generally known, there have been several previous studies that also aim to detect image forgery. However, most of this research requires quite large computational costs (which can be seen from the number of epochs and layers required), so that the flexibility of the proposed method is reduced and is difficult to apply in everyday life due to computational costs. In fact, there is a need for image forgery detection methods to be able to adapt to the addition of original and modified image data over time.

Therefore, in this report, an image forgery detection method is proposed that is relatively more efficient and has increased scalability that is directly proportional to the increase in data.

3.1 BENEFITS

Data mining in the form of image forgery detection can be used for the following things:

1. Increase comfort in obtaining information that is in accordance with the facts.
2. The public gets consideration in determining whether an image is real or fake.

By having a reference for the public to find out whether an image is real or not, it will certainly reduce the anxiety caused by fake images.

3.2 LIMITATIONS

There are several limitations that apply to this image forgery detection data mining, namely that the raw data must be an image with lossy compression (for example .jpg), and also not a computer-generated image (CGI).

4. Implementation

There are two main methods used in data mining, namely Error Level Analysis (ELA) and machine learning with deep learning techniques in the form of Convolutional Neural Network (CNN).

4.1 Error Level Analysis (ELA)

Error Level Analysis is a technique used to detect image manipulation by re-saving the image at a certain quality level and calculating the comparison between the compression levels. In general, this technique is carried out on images that have a lossy format (lossy compression). The image type used in this data mining is JPEG.

In JPEG images, compression is performed independently for each 8x8 pixel in the image. If an image is not manipulated, every 8x8 pixel in the image must have the same error rate.

4.2 Convolutional Neural Network (CNN)

CNN is a type of network based on feedforward, where the flow of information is only one way, namely from input to output. Although there are several types of CNN architecture, in general, CNN has several convolutional layers and pooling layers. Then, followed by one or more fully connected layers. In image classification, the input to the CNN is in the form of an image, so that each pixel can be processed.

In short, convolutional layers are used as feature extractors that learn feature representations from images that are input to the CNN.

Meanwhile, the pooling layer is responsible for reducing the spatial resolution of the feature maps. Generally, before the fully connected layer, there is a stack of several convolutional and pooling layers which function to extract more abstract feature representations.

After that, the fully connected layer will interpret these features and perform functions that require high-level reasoning. Classification at the end of the CNN will use the softmax function.

4.3 DESIGN AND IMPLEMENTATION

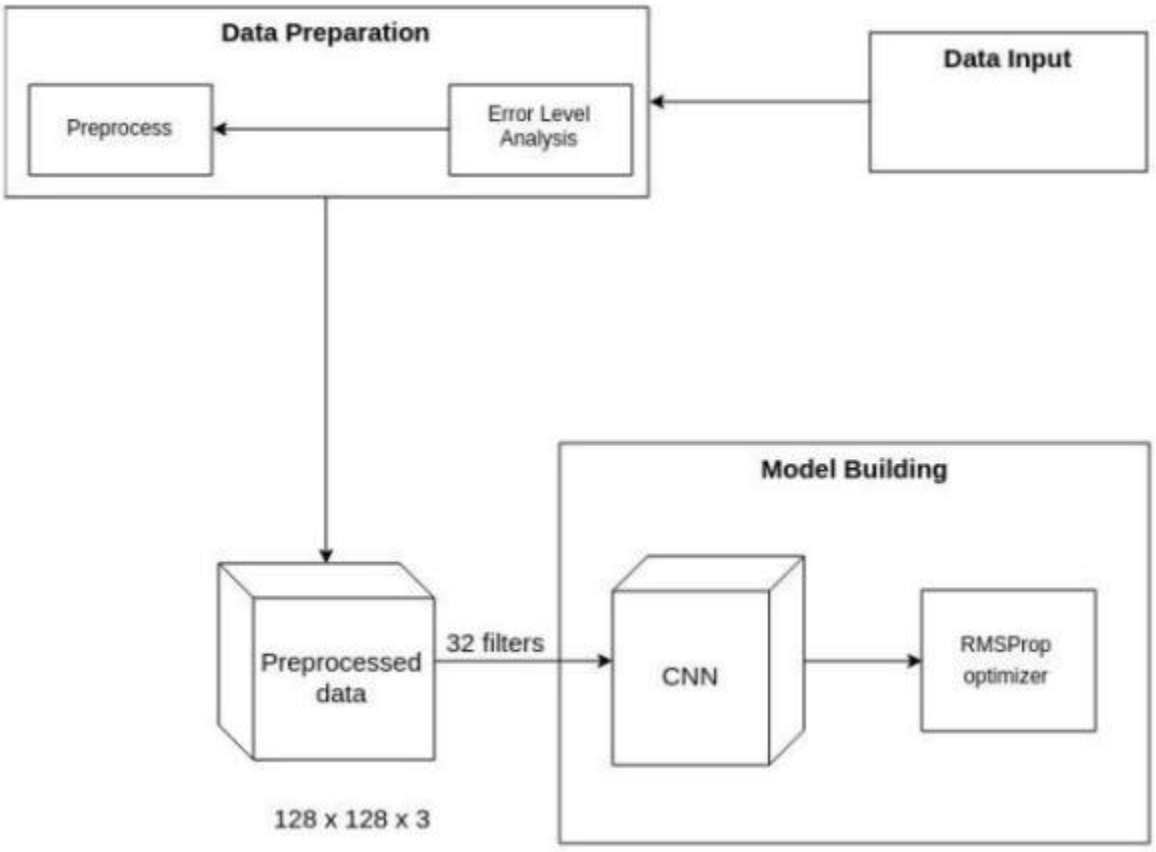

Figure 1 CNN architecture in general

In general, architectural design is divided into two large parts, namely data preparation and model building. In the initial stage, input data consisting of images in ".jpg" format, with the following details: 1771 images with tampered labels and 2940 images with real labels, were entered into the data preparation stage. The data preparation stage is the stage where each image which is input data is first converted into an image resulting from Error Level Analysis. Then, the ELA image will be resized into an image with a size of 128 x 128.

Figure 2 (a) Example of an original image of a lizard and (b) example of a modified image

Converting raw data to ELA images is a method used to increase the training efficiency of the CNN model. This efficiency can be achieved because the resulting ELA image contains information that is not excessive like the original image. The features produced by ELA images are focused on parts of the image that have an error level above the limit. In addition, pixels in ELA images tend to have similar colors or even very contrast to nearby pixels , so training the CNN model becomes more efficient.

Figure 3 (a) ELA image results from Figure 2 (a) and (b) ELA image results from Figure 2 (b)

After that, change the image size. The next step is to carry out normalization by dividing each RGB value by the number 255.0 to carry out normalization, so that the CNN converges faster (reaches the global minimum of the loss value belonging to the validation data) because the value of each RGB value only ranges between 0 and 1.

The next step is to change the label on the data, where 1 represents tampered and 0 represents real to categorical value. After that, the training data and validation data were divided using a division of 80% for training data and 20% for validation data.

The next step is to use training data and validation data to train a deep learning model using CNN. The optimization applied during training is the RMSProp optimizer, which is one of the adaptive learning rate methods.

The complete architecture used in the model building section can be seen in the image below which is a complete architectural image.

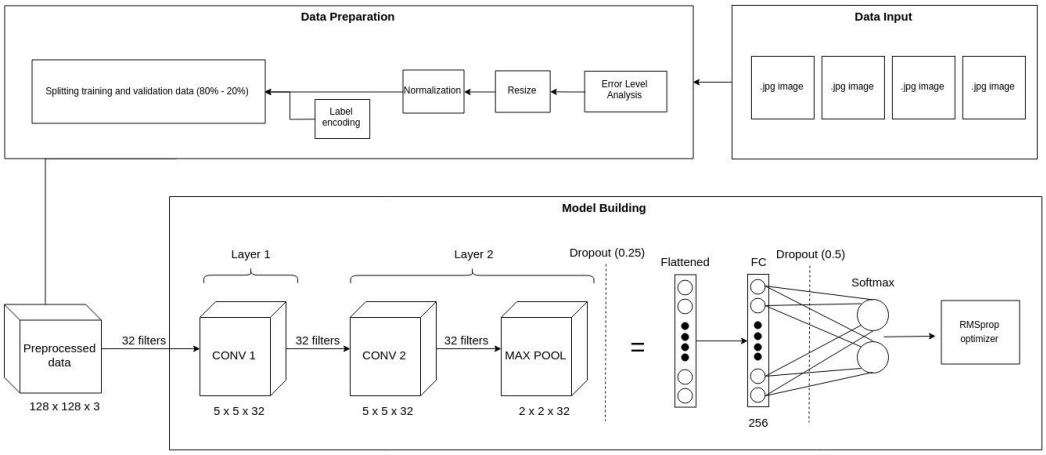

Figure 4 CNN model building architecture

In the deep learning model used, the first CNN layer consists of a convolutional layer with a kernel size of 5x5 and a number of filters of 32. The second layer of the CNN consists of a convolutional layer with a kernel size of 5x5 and a number of filters of 32, and a Max Pooling layer with a size of 2x2. The two convolutional layers used use a uniform glorot initializer kernel, and a ReLU activation function to make neurons in the convolutional layer carry out selection so that they can receive useful signals from the input data.

After that, the MaxPooling layer added a dropout of 0.25 to prevent overfitting. Next layer is a fully connected layer with a total of 256 neurons and a ReLU activation function. After the layer is fully connected, a dropout of 0.5 will be added to prevent overfitting. The output layer used has a softmax activation function.

In the architecture used, only two convolutional layers are needed, because the results produced from the conversion process into an ELA image can highlight important features to determine whether an image is original or has been properly modified.

5. Result

The results obtained from the proposed method have a maximum accuracy of 91.83%. Images of the accuracy curve and loss curve can be seen in the image below.

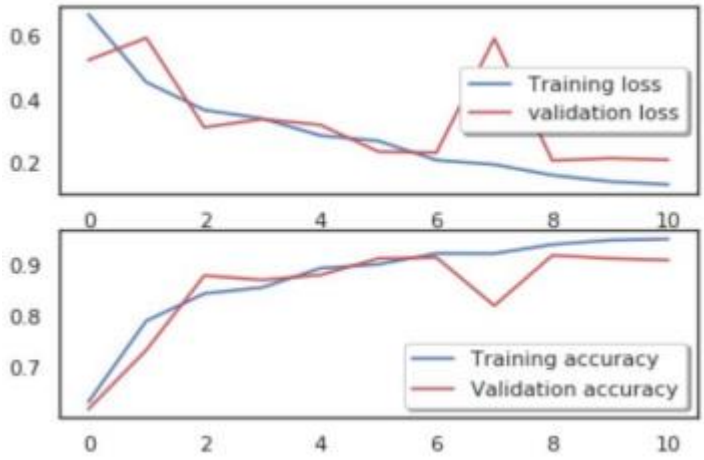

Figure 5 Accuracy curve and loss curve for training data

It can be seen in the picture above that the best accuracy was obtained at the 9th epoch . The validation loss value after the 9th epoch starts to plateau and eventually increases, which is a sign of overfitting. A good method for identifying the number of epochs to use during training is early stopping. With this method, training will be stopped when the validation accuracy value starts to decrease or the validation loss value starts to increase.

The number of training epochs required is small to achieve convergence, because the use of ELA converted image features makes model training much more efficient, and the normalization carried out on the RGB values for each pixel also speeds up the convergence of the CNN model.

The accuracy results obtained by the model in carrying out classification can be said to be relatively high. This is an indication that the feature is an ELA image successfully used to classify whether an image is an original image or has been modified.

6. CONCLUSION

In this report, there are several things that can be concluded from the results of machine learning using ELA and CNN.

1. CNN uses two convolutional layers, one MaxPooling layer, one fully connected layer, and one output layer with softmax to achieve 91.83% accuracy.
2. The use of ELA can increase efficiency and reduce computational costs of the training process.
   This can be seen from the reduction in the number of layers from the previous method and the number of epochs required. In the proposed model, the number of epochs required to achieve convergence is only 9.