ABSTRACT

Breast Cancer (Carcinoma Mammae) is one of the deadliest diseases in the world, especially in Indonesia itself. With so many cases of death that exist, we need a tool or media that can predict and identify breast cancer, with a detection tool that can provide early treatment for sufferers, it is hoped that it can prevent, treat, and also reduce mortality caused by this disease. The rapid development of Machine Learning has made programs that can detect cancer more effective and efficient over time. Convolution Neural Network (CNN) is one of the models in Machine Learning that is often used for image classification cases. Histopathology images are microscopic images of organ parts taken to be studied using a microscope. Histopathology image is one of the images used to detect and classify breast cancer. Some algorithms that use CNN models are ResNet-50, DenseNet-201, and VGG-16. The research was conducted by finding the best algorithm by comparing the accuracy results of the three algorithms to classify breast cancer histopathology images using the BreakHis Dataset. Hyper-parameters are external variables whose values can be set by the machine or set manually. Determining the value of hyper-parameters can affect the accuracy value. The hyper-parameters whose values are studied are epoch, batch size, learning rate, and dropout. The best algorithm after research is ResNet-50, with a hyper-parameter configuration of epoch 30, batch size 16, learning rate 0.0001, and dropout 0.5.

Keywords—Breast Cancer; Convolution Neural Network (CNN); Histopathology Images; ResNet-50; DenseNet-201; VGG-16; BreakHis Dataset; Hyper-Parameter