ABSTRACT

Cancer based on data from the World Health Organization (WHO) is the deadliest disease in the world. Cervical cancer is the fourth rank that attacks women after breast cancer. Process detecting cervical cancer by Anatomical Pathologists usually uses Pap smear images to detect cervical cancer cells. However, it takes a long time. The development of information technology today has been carried out a lot for computer-based early detection of cervical cancer.

This study aims to improve the accuracy of the Pap smear image classification model for early diagnosis of cervical cancer. One of the problems in making the model is the limited dataset. Generative Adversarial Networks (GAN) can make synthetic data from learning data distribution. Synthetic data can be used to widen the dataset and have similarities with the original image. Tests were carried out 30 times using Pap smear images and the addition of synthetic data. Performance of adding GAN synthetic data using Densely Connected Convolution Network (DenseNet). Making the DeseNet model for detection and classification with five classes, namely, Normal, Adenocarcinoma, Low-grade Squamous intraepithelial lesion (LSIL), High-grade Squamous intraepithelial lesion (HSIL), and Squamous Cell Carcinoma (SCC).

The results obtained with the addition of synthetic data can improve the accuracy of the Pap smear image model which has a limited dataset. DenseNet performance evaluation resulted in an average accuracy of the DenseNet-121 model of 93.97%±2.84, DenseNet-169 of 93.4%±1.63, and DenseNet-201 of 94.4%±2.54. The results of this study show that the addition of synthetic images as data augmentation in the dataset can increase the accuracy of the model.

Keywords: cervical cancer, DCGAN, DenseNet, deep learning, Pap smear

INTISARI

Kanker berdasarkan data dari *World Health Organization* (WHO) pernyakit paling mematikan di dunia. Kanker Serviks merupakan peringkat keempat yang menyerang wanita setelah Kanker Payudara. Proses deteksi kanker serviks yang dilakukan Dokter Patologi Anatomi biasanya menggunakan citra Pap smear untuk mendeteksi sel kanker serviks. Namun, membutuhkan waktu yang lama. Perkembangan teknologi informasi saat ini telah banyak dilakukan untuk deteksi dini kanker serviks berbasis komputer.

Penelitian ini bertujuan meningkatkan akurasi model klasifikasi citra Pap smear untuk diagnosis dini kanker serviks. Salah satu permasalahan dalam pembuatan model klasifikasi yaitu keterbatasan *dataset*. *Generative Adversarial Networks* (GAN) mampu membuat data sintetis dari pembelajaran distribusi data. Data sintetis dapat digunakan untuk memperlebar *dataset* dan memiliki kemiripan dengan citra asli. Pengujian dilakukan sebanyak 30 kali menggunakan citra Pap smear dan penambahan data sintetis. Kinerja penambahan data sintetis GAN menggunakan *Densely Connected Convolution Network* (DenseNet). Pembuatan model DeseNet sebagai deteksi dan klasifikasi dengan kelas lima kelas yaitu, Normal, *Adenocarcinoma, Low-grade Squamous intraepithelial lesion* (LSIL), *High-grade Squamous intraepithelial lesion* (HSIL), dan *Squamous Cell Carcinoma* (SCC).

Hasil yang didapatkan dengan penambahan data sintetis mampu meningkatkan akurasi model citra Pap smear yang memiliki keterbatasan *dataset*. Evaluasi kinerja DenseNet menghasilkan rata-rata akurasi model DenseNet-121 sebesar 93,97%±2,84, DenseNet-169 sebesar 93,4%±1,63 dan DenseNet-201 94,4%±2,54. Hasil penelitian ini menunjukkan bahwa penambahan citra sintetis sebagai data augmentasi di *dataset* mampu meningkatkan akurasi dari model.

Kata kunci – kanker serviks, DCGAN, DenseNet, *deep learing*, Pap smear.