

BAB II

TINJAUAN PUSTAKA

2.1 Tinjauan Pustaka

Terdapat beberapa penelitian yang menggunakan metode ekstraksi fitur untuk melakukan pengenalan karakteristik lesi GGO pada citra hasil CT *scan*.

Penelitian yang dilakukan oleh Katsumata dkk (Katsumata et al. 2008) yang mengusulkan sebuah sistem untuk mendeteksi area GGO dengan menghitung empat fitur statistik dari fitur densitas dan satu fitur bentuk. Penelitian ini menggunakan 31 dataset citra thorax abnormal yang diperoleh dari *Multi Detector rowCT* (MDCT). Fitur statistik yang dihitung antara lain *mean*, *standard deviation*, *skewness* dan *kurtosis*. Dari keempat fitur tersebut diperoleh tingkat akurasi 79% dengan tingkat *False Positive* (FP) sebesar 1,3. Kemudian, digunakan fitur bentuk yang dihitung menggunakan *Minimum Directional Difference Filter* (Min-DD) untuk mengurangi tingkat FP. Dari hasil penelitian, disimpulkan bahwa sistem yang diusulkan tersebut dapat digunakan di bidang medis.

Sedangkan penelitian yang dilakukan oleh Bastawrous dkk (Bastawrous et al. 2005), menggunakan beberapa fitur morfologi dan *graylevel*. Fitur morfologi yang dihitung antara lain *area*, *compactness*, dan *irregularity*. Kemudian, fitur *graylevel* terdiri dari *mean intensity value* dan *maximum intensity value*. Penelitian ini menggunakan 715 *slice* citra CT yang berisi 25 nodul GGO dan diklasifikasi menggunakan *template matching* dan diperoleh sensitivitas 92% dengan FP 0,76 per *slice*. Selanjutnya, digunakan metode ANN untuk mengurangi tingkat FP. Setelah menggunakan ANN, FP dapat dikurangi dari 0,76 menjadi 0,25 per *slice* tetapi metode ini mengurangi tingkat sensitivitas menjadi 84%.

Penelitian yang dilakukan Yokota dkk (Yokota et al. 2014), juga melakukan deteksi area GGO pada citra CT berdasarkan fitur statistik dengan menghitung keempat sudut dari metode GLCM antara lain *energy*, *entropy*, *inertia* dan *correlation*. Metode yang diusulkan, diterapkan pada 31 citra CT yang diperoleh dari LIDC. Dengan menggunakan

metode tersebut diperoleh akurasi 93%. Hal ini disebabkan karena area pembuluh darah dihilangkan terlebih dahulu menggunakan *3D line filter*.

Berdasarkan beberapa penelitian terdahulu, proses pengenalan karakteristik GGO dapat dilakukan dengan ekstraksi fitur. Salah satu metode yang sering digunakan adalah ekstraksi fitur berbasis tekstur. Penelitian ini menggunakan metode ekstraksi fitur tekstur orde dua yaitu GLCM yang dianggap lebih baik karena metode ini memperhitungkan hubungan antarpasangan dua piksel pada citra asli, sedangkan pengukuran tekstur pada orde pertama menggunakan perhitungan statistika yang didasarkan pada nilai piksel citra asli, seperti varians, dan tidak memperhatikan hubungan ketetanggaan piksel (Kadir & Adhi 2012). Setelah proses ekstraksi fitur, dibutuhkan proses klasifikasi untuk dapat mengenali karakteristik lesi GGO. Beberapa penelitian yang menggunakan metode-metode klasifikasi untuk citra medis telah dilakukan, seperti penelitian yang dilakukan oleh Kaur dkk (Kaur et al. 2016), yang melakukan pengenalan terhadap jaringan hati yang normal dan jaringan hati yang berlemak menggunakan citra ultrasound. Penelitian tersebut menggunakan fitur tekstur yaitu *Angular Second Moment*, *Contrast*, *Correlation*, *Sum of Squares*, *Sum Average*, *Inverse Difference Moment*, *Sum Variance*, *Sum Entropy*, *Entropy*, *Difference Variance*, *Difference Entropy*, *Information measure of correlation 1* dan *2*. Tahap selanjutnya adalah seleksi fitur menggunakan metode *two-fold* (*Fisher's Discrimination Ratio* dan *Pearsons Correlation Coefficient*) dan diperoleh lima fitur yang signifikan. Tahap terakhir yang dilakukan adalah klasifikasi menggunakan Naïve Bayes dan diperoleh akurasi sebesar 92,3% dengan tingkat sensitivitas sebesar 100%.

Penelitian yang dilakukan oleh Bhuvaneshwari dkk (Bhuvaneshwari et al. 2014), melakukan klasifikasi citra paru-paru secara otomatis untuk mengenali penyakit paru-paru yaitu Efusi Pleura, Emphysema, Bronchitis dan paru-paru normal. Tahap awal yang dilakukan adalah ekstraksi fitur menggunakan *Gabor Filter* dan transformasi *Walsh Hadamard*. Tahap selanjutnya adalah seleksi fitur menggunakan *Correlation based Feature Selection* (CFS) dan *Principal Component Analysis* (PCA). Tahap terakhir yaitu proses klasifikasi dengan membandingkan beberapa metode seperti Naïve Bayes, J48,

K-NN, MLP. Hasil yang diperoleh menunjukkan bahwa PCA dan MLP menghasilkan akurasi tertinggi yaitu 81%. Sementara penelitian yang dilakukan oleh Sergeeva dkk (Sergeeva et al. 2016), melakukan klasifikasi nodul paru-paru dengan melakukan ekstraksi fitur menggunakan fitur tekstur dan morfologi pada citra hasil CT scan yang diperoleh dari *Lung Image Database Consortium* (LIDC). Metode ekstraksi fitur tekstur yang digunakan adalah GLCM, kemudian fitur morfologi yaitu *area*, *perimeter*, *irregularity index*, *equivalent diameter*, *ferret max diameter*, *length*, *convex area*, dan *solidity*. Selain itu, transformasi wavelet diskrit juga digunakan sebagai fitur tambahan. Tahap selanjutnya adalah proses seleksi fitur menggunakan *CFSSubsetEvaluation* kemudian proses klasifikasi menggunakan WEKA dengan *10-fold cross validation* menggunakan beberapa metode yaitu *Neural Network*, *Random Forest*, KNN, Naïve Bayes dan SVM. Hasil yang diperoleh menunjukkan bahwa metode KNN, *Random Forest* dan *Neural Network* memberikan hasil yang terbaik. Metode Naïve Bayes juga digunakan oleh Zhou dkk (Zhou et al. 2015), yang melakukan pengenalan patologi otak menggunakan citra *Magnetic Resonance Imaging* (MRI). Tahap yang dilakukan adalah ekstraksi fitur menggunakan *Discrete Wavelet Transform* (DWT) dan tahap klasifikasi menggunakan Naïve Bayes Classifier. Hasil yang diperoleh dari 64 citra yang digunakan yaitu akurasi 92,60%, sensitivitas 94,50% dan spesifisitas 91,70%.

Sementara penelitian yang dilakukan oleh Yildiz (Yildiz 2017), melakukan klasifikasi dan pengurangan dimensi atau seleksi fitur pada citra serat kain yang diambil menggunakan kamera thermal. Tahap awal yang dilakukan yaitu konversi citra menjadi *gray level*, kemudian pengurangan dimensi menggunakan PCA. Tahap klasifikasi dilakukan menggunakan metode Naïve Bayes dan KNN. Hasil yang diperoleh menunjukkan bahwa KNN memberikan tingkat akurasi yang lebih tinggi karena performa Naïve Bayes akan menurun apabila jumlah dataset yang digunakan terlalu banyak dan tidak efektif pada vector fitur berdimensi tinggi. Berdasarkan asumsi ini, maka perlu adanya pengurangan fitur dan hanya mengambil fitur-fitur yang paling signifikan. Berdasarkan penelitian yang dilakukan oleh Karabulut dkk (Karabulut et al. 2012), yang membandingkan pengaruh seleksi fitur terhadap beberapa metode

klasifikasi seperti Naïve Bayes, MLP, dan J48. Penelitian tersebut menggunakan 15 dataset yang berbeda yang diperoleh dari *Data Mining Repository of University of California Irvine* (UCI). Untuk membandingkan algoritme klasifikasi dengan metode seleksi fitur, WEKA *data mining tools* digunakan dengan *10-fold cross validation*. Hasil penelitian diperoleh bahwa metode Naïve Bayes paling dipengaruhi oleh *Gain Ratio* pada 5 dataset dan terjadi peningkatan akurasi dari 77,68 menjadi 84,93. Sementara, metode Chi-square tidak berpengaruh terhadap Naïve Bayes. Oleh sebab itu, penelitian ini menggunakan algoritme Gain Ratio untuk mendapatkan fitur yang paling signifikan dan meningkatkan hasil klasifikasi.