

BAB 1 PENDAHULUAN

1.1 Latar Belakang

Kanker payudara dan serviks merupakan jenis kanker tertinggi yang dialami oleh wanita di dunia. Kanker payudara menempati peringkat pertama, kemudian disusul paru-paru, usus besar, dan serviks menempati peringkat keempat pada wanita (WHO, 2020). Kanker payudara bukanlah penyakit menular yang disebabkan oleh virus atau bakteri, kanker payudara disebabkan oleh aktivitas *proliferasi* sel (siklus pembelahan sel) pada payudara serta kelainan yang menurunkan atau menghilangkan regulasi kematian sel (*apoptosis*) (Suarfi et al., 2019). Sekitar setengah dari kanker payudara berkembang pada wanita yang tidak memiliki resiko kanker payudara, yang dapat diidentifikasi selain jenis kelamin dan usia diatas 40 tahun (WHO, 2021). Sedangkan kanker serviks 90% disebabkan oleh infeksi virus *human papillomavirus* (HPV), dan 10% lainnya disebabkan oleh penyakit menular seksual (PMS), pola hidup tidak sehat, penggunaan pil KB, dan keturunan atau *genetical* (Biofarma, 2023).

Berdasarkan perkiraan konservatif, benua dengan kejadian kanker payudara dan serviks terbanyak adalah Asia dengan angka kasus kanker payudara sebanyak 1.026.171 (Merino, 2021), sementara kanker serviks diperkirakan sebanyak 339.675 – 364.192 kasus (WHO, 2022). Perkembangan populasi kanker wanita ini setiap tahunnya meningkat sekitar 14.100 kasus baru. Di Asia Tenggara, Indonesia menjadi negara pertama dengan kasus kanker payudara dan kanker serviks terbanyak dengan jumlah kasus kanker payudara sebanyak 65.858 kasus dan kanker serviks dengan jumlah 36.633 kasus pada tahun 2020 (Globocan, 2020). *World Health Organization* memperkirakan pada tahun 2040 kanker payudara akan meningkat lebih dari tiga juta kasus baru setiap tahunnya atau meningkat 40% dan terdapat lebih dari satu juta kasus kematian per tahun atau meningkat 50% dari tahun 2020. Angka kematian terbesar pada kasus pasien dengan kanker payudara dan serviks terjadi pada negara berkembang atau dalam masa transisi ekonomi (Arnold et al., 2022).

Untuk memperkecil populasi kanker yang terus bertambah, berbagai negara menerapkan program deteksi dini yang dikombinasikan dengan berbagai pengobatan untuk menurunkan angka penyakit tersebut. Pada awal pemeriksaan kanker payudara dapat dilakukan dengan mandiri melalui SADARI (Periksa Payudara Sendiri), jika terdapat kejangalan dalam tubuh, dilakukannya pemeriksaan fisik oleh dokter. Sedangkan pada kanker serviks, dokter akan melakukan pemeriksaan *pap smear* (pengambilan sampel sel serviks). Saat mencurigai adanya sel kanker di tubuh pasien, dokter akan melakukan serangkaian pemeriksaan untuk menegakkan diagnosa kanker. Pemeriksaan yang dilakukan adalah pemeriksaan fisik, pemeriksaan radiologis seperti *rontgen*, *computed tomography scan* (CT scan), *magnetic resonance imaging* (MRI), dan pemeriksaan *Patologi Anatomi* (PA). Pengambilan sampel sel dan jaringan pada pemeriksaan PA ini dapat dilakukan dengan beberapa metode, yaitu biopsi aspirasi jarum halus (menyedotnya dengan jarum suntik), kerokan, biopsi atau operasi. Sampel sel dan jaringan tubuh yang diambil dapat berupa benjolan dalam tubuh (payudara), dan kerokan (serviks) (Nugroho, 2022).

Slide histologi jaringan payudara dan serviks digunakan untuk memberikan evaluasi definitif, interpretasi slide ini dilakukan oleh ahli patologi. Pada kasus penyakit kanker serviks para patolog berusaha mendeteksi *cervical intraepithelial neoplasia* (CIN), yang merupakan kondisi praganas untuk kanker serviks. Sedangkan pada kanker payudara, histopatologi dilakukan untuk membedakan 2 kelas utama yaitu *benign* dan *malignant* dimana *benign* merupakan tumor dengan pertumbuhan lambat dan terlokasi, dan *malignant* merupakan kanker yang melakukan invasi dan merusak struktur yang terhubung dengannya. Metode histopatologi menjadi *gold standard* dalam mendeteksi kanker payudara dan kanker serviks, dimana dalam pemeriksaan ini dapat menentukan jenis pre-kanker ganas maupun jinak (Suarfi et al., 2019).

Pemeriksaan gambar, *scan*, dan foto sangat penting dalam diagnosis medis. Teknologi *computer vision* memiliki klaim untuk tidak hanya menyederhanakan tetapi juga untuk mencegah diagnosis yang salah dan menurunkan biaya pengobatan dengan menganalisis gambar ultrasound, MRI, dan CT scan, yang

semuanya merupakan bagian dari repertoar reguler obat modern (Guggenbühl, 2020). Penggunaan *computer vision* telah digunakan dalam mendeteksi suhu tubuh dan penggunaan masker dengan kamera pada rumah sakit, mall, atau tempat umum lainnya (Ahmad et al., 2021).

Computer vision adalah penggabungan antara pengolahan citra digital dan *machine learning*. Untuk menghasilkan ketepatan kelas yang sesuai dalam *computer vision* dibutuhkan klasifikasi. Klasifikasi merupakan satu hal paling penting dan esensial dalam *machine learning* dan data mining. Penelitian-penelitian yang telah dilakukan dengan menerapkan *machine learning* dengan dataset yang berbeda untuk mengklasifikasi kanker serviks dan payudara. Banyak dari penelitian-penelitian tersebut yang menampilkan akurasi klasifikasi yang baik. Mayoritas algoritma yang digunakan oleh para peneliti dalam mendeteksi kanker serviks dan payudara adalah algoritma SVM (*Support Vector Machine*) (He et al., 2023), *Linear Discriminant Analysis* (LDA) (S. Sharma & Kumar, 2022), KNN (*K-Nearest Neighbor*), dan RF (*Random Forest*) (Shallu & Mehra, 2018). Algoritma tersebut banyak digunakan karena termasuk kedalam klasifikasi *supervised learning* dimana label telah diberikan sebelum diolah.

Shallu Sharma dan Rajesh Mehra melakukan perbandingan antara penggunaan arsitektur *neural network* yang disempurnakan dengan algoritma *logistic regression* dan yang tidak. Penelitian ini menggunakan gambar histopatologi kanker payudara yang diklasifikasikan kedalam dua kelas yaitu benign dan malignant. *Neural network* yang digunakan adalah *convolution neural network* (CNN) dan arsitektur yang dibandingkan yaitu *Visual Geometric Group 16* (VGG16), VGG19, dan ResNet50. Pada penelitian ini hanya menggunakan rotasi gambar sebagai teknik penambahan data dengan menggunakan tiga sudut rotasi yaitu 90°, 180°, dan 270°. Hasil penelitian ini yaitu klasifikasi menggunakan arsitektur VGG16 yang disempurnakan dengan *logistic regression* menghasilkan kinerja terbaik dengan akurasi 92,60% dan akurasi *precision score* sebesar 95,95% untuk 90 data training dan 10 data testing. Secara keseluruhan klasifikasi yang disempurnakan dengan *logistic regression* lebih baik daripada yang tidak disempurnakan karena kinerja

jaringan yang dihasilkan lebih kuat bahkan ketika ukuran data training kurang (Shallu & Mehra, 2018).

Perbandingan performa klasifikasi juga dilakukan oleh Siqi He, Bo Xiao, Huajiang Wei, Shenjiao Huang, dan Tongsheng Chen, dimana para peneliti menggunakan data citra histopatologi kanker serviks yang diperoleh dari 20 pasien yang didiagnosa menderita kanker serviks, termasuk 135 gambar slide utuh (WSI). Perbandingan dilakukan dengan menggunakan ekstraksi fitur tekstur dan morfologi, serta *Support Vector Machine* (SVM) sebagai pengklasifikasinya. Hasil dari penelitian ini adalah akurasi klasifikasi SVM menggunakan first-order histogram, K-means clustering, GLCM dan fitur nukleus untuk mengekstraksi fitur masing-masing adalah 87,4%, 90,6%, 91,6%, dan 93,5% (He et al., 2023).

Berdasarkan latar belakang di atas, peneliti melakukan perbandingan terhadap gambar histopatologi kanker serviks dan payudara. Penelitian ini menggunakan ekstraksi fitur GLCM karena gambar histopatologi memiliki keragaman tekstur yang signifikan. Selanjutnya, analisis *Principal Component Analysis* (PCA) dan regresi digunakan untuk memilih fitur GLCM yang paling relevan. Algoritma klasifikasi yang dipilih adalah *Learning Vector Quantization* (LVQ), dipilih karena belum ada penelitian yang mengaplikasikan klasifikasi tersebut terhadap citra histopatologi kanker serviks dan payudara, serta LVQ mampu melakukan klasifikasi *multiclass*. Selain itu, *Convolution Neural Network* (CNN) juga digunakan karena dalam *deep learning*, CNN mencakup semua tahap ekstraksi, seleksi, dan klasifikasi secara menyeluruh.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang di atas, maka dapat dirumuskan permasalahan:

1. Bagaimana melakukan ekstraksi fitur pada citra histopatologi kanker payudara dan serviks dengan GLCM dan CNN?
2. Bagaimana melakukan proses training dengan Algoritma *Learning Vector Quantization* (LVQ) dan *Convolution Neural Network* (CNN)?
3. Bagaimana perbandingan hasil klasifikasi untuk GLCM + LVQ, GLCM + PCA + LVQ, GLCM + regresi + LVQ dan CNN?

1.3 Tujuan Penelitian

Tujuan yang hendak dicapai dalam penelitian ini yaitu:

1. Mengimplementasikan ekstraksi fitur pada citra hispatologi dengan GLCM dan CNN.
2. Mengetahui perbedaan proses training algoritma *machine learning* pada citra histopatologi.
3. Mengetahui perbandingan performa algoritma *machine learning* pada citra histopatologi.

1.4 Manfaat Penelitian

Manfaat dalam penelitian ini yaitu:

1. Memberikan pengetahuan yang lebih mendalam terkait perbandingan algoritma *machine learning* pada gambar hasil histopatologi.
2. Menambah literatur bagi fakultas dan atau universitas dalam memberikan gambaran terkait perbandingan algoritma *machine learning* pada gambar hasil histopatologi.
3. Sebagai referensi dan pedoman dalam membandingkan algoritma *machine learning* pada gambar hasil histopatologi.