

BAB I PENDAHULUAN

1.1 Latar Belakang

Kanker paru termasuk salah satu jenis kanker dengan angka mortalitas yang tinggi, baik skala nasional maupun dunia. Berdasarkan angka jumlah kasus baru, kanker paru di Indonesia berada pada peringkat ketiga setelah kanker payudara dan kanker serviks. Kanker paru juga menjadi jenis kanker dengan angka kematian paling tinggi dengan persentase 13,2% dari total 30.843 kasus (Kementerian Kesehatan, 2023). Proses diagnosis yang sering terlambat karena gejala awal penyakit yang cenderung tidak khas dan sulit dikenali, menjadi salah satu faktor utama tingginya angka kematian pada kanker paru (Gatsonis dkk., 2011). Mayoritas pasien kanker paru diketahui mengidap penyakit tersebut ketika telah memasuki stadium lanjut, padahal tingkat kelangsungan hidup pasien sangat bergantung pada saat diagnosis ditegakkan (Goldstraw dkk., 2016). Diagnosis dini terbukti meningkatkan efektivitas terapi dan prognosis pasien secara signifikan. Oleh karena itu, diperlukan metode diagnosis yang akurat dan mampu mendeteksi kanker paru sejak tahap awal.

Metode biopsi masih menjadi *gold standard* dalam diagnosis kanker paru karena memberikan informasi histopatologis yang akurat, tetapi prosedur ini bersifat invasif dan memiliki risiko komplikasi (Pandey & Yadav, 2025). Keterbatasan tersebut mendorong pengembangan metode non-invasif yang tetap memiliki tingkat akurasi tinggi, salah satunya melalui pencitraan medis seperti *Computed Tomography* (CT). CT scan mampu menghasilkan citra resolusi tinggi yang menampilkan detail anatomi dan karakteristik lesi secara jelas. Seiring perkembangan teknologi analisis data, citra CT tidak lagi hanya berfungsi sebagai media visualisasi, tetapi juga sebagai sumber informasi kuantitatif yang dapat dianalisis lebih lanjut menggunakan teknik pengolahan citra dan *machine learning* (Aerts dkk., 2014; Gillies dkk., 2016).

Perkembangan *machine learning* dalam pencitraan medis memungkinkan ekstraksi fitur statistik dan tekstur dari citra CT untuk membedakan jaringan normal dan patologis (Litjens dkk., 2017). Gillies dkk. (2016) menegaskan bahwa citra medis telah berkembang menjadi instrumen analisis kuantitatif, sedangkan Aerts dkk. (2014) menunjukkan bahwa fitur tekstur citra CT mampu merepresentasikan karakteristik biologis tumor. Representasi tersebut dapat diperoleh dengan teknik ekstraksi fitur contohnya melalui teknik *Gray Level Co-occurrence Matrix* (GLCM), *Local Binary Pattern* (LBP), serta Gabor filter. Fitur-fitur ini mengubah pola visual menjadi representasi numerik yang dapat digunakan sebagai input algoritma klasifikasi (Haralick dkk., 1973; Ojala dkk., 2002). Dengan demikian, kualitas dan kombinasi fitur yang digunakan sangat menentukan performa model dalam membedakan jenis kanker paru.

Berbagai penelitian menunjukkan bahwa kombinasi fitur tekstur atau fitur hibrida dapat meningkatkan kinerja klasifikasi dibandingkan fitur tunggal. Li dkk. (2024) mengombinasikan GLCM, fitur *Haralick*, *Autoencoder* bersama algoritma klasifikasi *Support Vector Machine* yang berhasil dengan hasil akurasi mencapai 100%. Nair dkk. (2024) serta Kumari dkk. (2025) juga mengombinasikan berbagai fitur tekstur dan frekuensi menggunakan *Random Forest* dengan performa yang sangat baik. Kedua penelitian tersebut masih berfokus pada klasifikasi biner antara paru normal dan kanker paru. Tantangan yang lebih kompleks terdapat pada proses klasifikasi multi-kelas, terutama pada kasus kanker paru tipe *Non-Small Cell Lung Cancer* (NSCLC) yang terdiri atas beberapa subtype dengan karakteristik tekstur yang saling beririsan.

Upaya klasifikasi multi-kelas NSCLC telah dilakukan, seperti oleh Pahlevi dkk. (2025) yang menggunakan GLCM dan *Extreme Learning Machine* dengan akurasi 91,06%, serta Zhang dkk. (2025) yang mengekstraksi lebih dari 1.800 fitur radiomik dengan performa tinggi pada validasi silang namun menurun pada data independen. Hasil tersebut menunjukkan masih adanya keterbatasan dalam representasi fitur dan generalisasi model. Berdasarkan pertimbangan celah penelitian yang masih ada, maka diusulkan penelitian dengan teknik ekstraksi fitur

tekstur hibrida yang menggabungkan GLCM, LBP, dan Gabor filter pada citra CT paru. Teknik ekstraksi tersebut akan digabungkan juga dengan algoritma klasifikasi *machine learning* seperti *Random Forest* dan *K-Nearest Neighbor* (KNN). Pendekatan ini diharapkan mampu meningkatkan kemampuan diferensiasi antar subtype NSCLC, sehingga memberikan kemajuan terhadap peningkatan kualitas sistem klasifikasi kanker paru dengan tingkat akurasi lebih baik dan berpotensi mendukung diagnosis klinis non-invasif.

1.2 Tujuan

Berlandaskan uraian latar belakang yang sudah dijelaskan, penelitian ini memiliki tujuan sebagai berikut:

- 1 Mendapatkan representasi fitur tekstur citra CT paru melalui teknik ekstraksi fitur tekstur hibrida dengan menggabungkan fitur tekstur seperti GLCM, LBP, dan Gabor filter.
- 2 Mendapatkan hasil klasifikasi citra CT paru ke dalam empat kelas, yaitu normal serta tiga subtype kanker NSCLC dengan menggunakan pendekatan algoritma klasifikasi seperti *Random Forest* serta KNN.
- 3 Mendapatkan kombinasi fitur tekstur hibrida dan algoritma klasifikasi yang memberikan performa terbaik berdasarkan metrik kinerja seperti akurasi, spesifisitas, sensitivitas, dan *Area Under the Curve* (AUC).

1.3 Manfaat

Berlandaskan tujuan yang telah dipaparkan, penelitian ini diharapkan memberikan sejumlah manfaat sebagai berikut:

- 1 Berkontribusi secara ilmiah dalam bidang pengolahan citra medis dengan mengusulkan pendekatan teknik ekstraksi fitur tekstur hibrida untuk klasifikasi kanker paru.
- 2 Menyediakan kajian komparatif kinerja algoritma *Random Forest* dan KNN pada kasus klasifikasi multi-kelas kanker paru, yang sebelumnya masih jarang dilakukan.