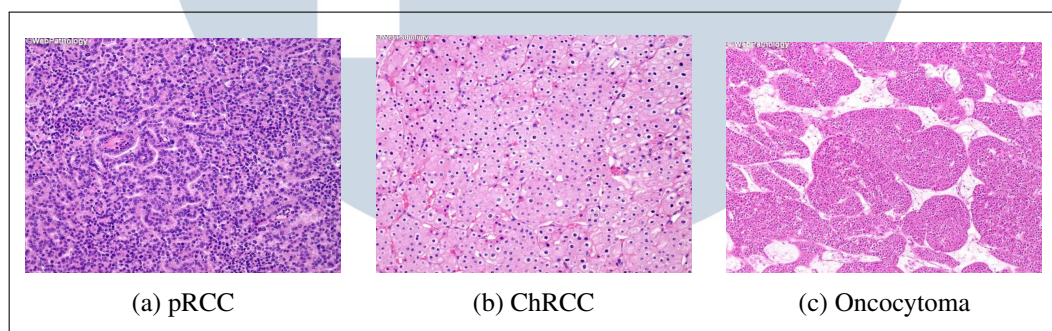


BAB 2

LANDASAN TEORI

2.1 Renal Cell Carcinoma (RCC)

Renal Cell Carcinoma (RCC) merupakan keganasan primer ginjal yang berasal dari sel epitel tubulus proksimal dan memiliki tingkat heterogenitas visual yang sangat tinggi [25]. Dalam konteks radiomik, tantangan utama terletak pada kemiripan karakteristik tekstur antar sub tipe utama, yaitu *Clear Cell* (ccRCC), *Papillary* (pRCC), dan *Chromophobe* (chRCC) pada citra CT scan [26]. Diferensiasi yang akurat secara non-invasif sangat krusial karena setiap sub tipe memiliki respons terapi dan risiko metastasis yang berbeda [27].



Gambar 2.1. Perbandingan gambaran histopatologi sub tipe tumor, (a) Histopatologi pada pRCC, (b) Sitoplasma pucat pada ChRCC, dan (c) Sel onkosit eosinofilik pada Oncocytoma.

Setiap sub tipe tumor ginjal memiliki profil histopatologi unik yang memengaruhi representasi fitur pada citra radiologis:

1. Papillary RCC (pRCC): Ditandai dengan struktur papiler dengan inti fibro-vaskular. Secara radiologis, pRCC cenderung hipovaskular, homogen, dan memiliki nilai *enhancement* yang rendah (10–30 HU), sehingga sering menyerupai kista atau lesi jinak pada CT scan [28].
2. Chromophobe RCC (ChRCC): Memiliki karakteristik sel pucat dengan inti *raisinoid*. ChRCC sering menunjukkan tumpang tindih visual dengan *oncocytoma*, sehingga sulit dibedakan hanya melalui observasi radiologis konvensional [29].
3. Oncocytoma: Merupakan neoplasma jinak yang secara morfologis sangat mirip dengan ChRCC karena sama-sama memiliki sitoplasma eosinofilik [6].

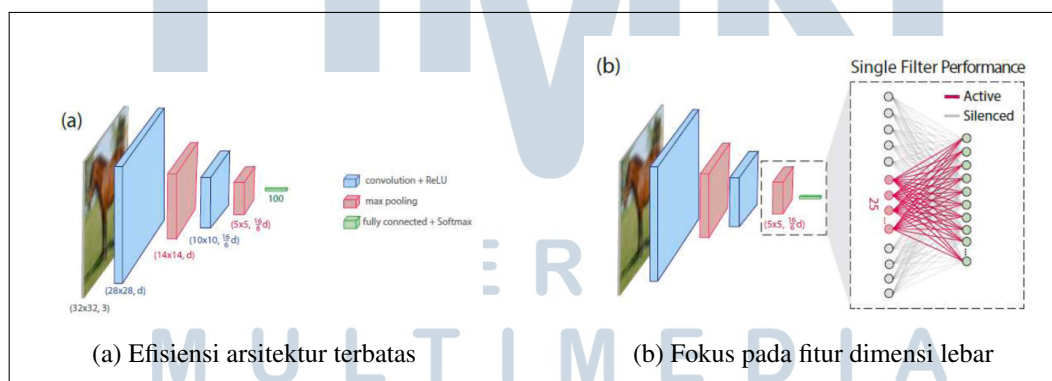
Kemiripan ini menjadi sumber utama *misdiagnosis* dalam klasifikasi massa ginjal [30].

2.2 Shallow Learning

Shallow learning dalam konteks penelitian ini merujuk pada algoritma *machine learning* tradisional yang menggunakan hierarki pemrosesan lebih pendek dibandingkan arsitektur jaringan saraf tiruan yang dalam. Landasan teoretis utama dari pemanfaatan model ini didasarkan pada prinsip *Power-Law Scaling* yang menyatakan bahwa akurasi prediktif sebuah model dipengaruhi secara signifikan oleh dimensi informasi pada fitur input dibandingkan dengan kedalaman arsitektur pemrosesan [17]. Tingkat kesalahan model (ϵ) terbukti berkurang seiring dengan bertambahnya dimensi fitur (d_1) yang digunakan, sesuai dengan persamaan berikut:

$$\epsilon(d_1) = A(d_1)^{-p} \quad (2.1)$$

Dalam alur kerja radiomik, citra medis telah diekstraksi menjadi fitur numerik berdimensi tinggi yang padat akan informasi tekstur. Karakteristik fitur yang sudah terstruktur ini memungkinkan algoritma *shallow learning* melakukan pemetaan klasifikasi pRCC secara optimal tanpa perlu melakukan ekstraksi fitur ulang melalui struktur jaringan yang berlapis-lapis [31]. Hal ini memungkinkan model untuk fokus pada generalisasi pola dari representasi fitur yang sudah tersedia.



Gambar 2.2. Representasi mekanisme pemrosesan dalam paradigma *shallow learning*.

Kelebihan utama model dalam paradigma ini terletak pada efisiensi komputasi (*space-time complexity*), di mana waktu pelatihan dan penggunaan memori jauh lebih rendah dibandingkan model kompleks [32]. Struktur algoritma

yang lebih sederhana juga efektif dalam menekan risiko *overfitting* ketika dihadapkan pada dataset dengan jumlah sampel terbatas. Untuk menangani permasalahan ketidakseimbangan kelas antara sampel pRCC dan non-pRCC, diterapkan strategi penyeimbang bobot (*class-weighted strategy*) guna memastikan model tetap memiliki sensitivitas yang stabil dalam mendeteksi kelas minoritas [17, 31].

2.3 Ensemble Learning

Ensemble Learning merupakan paradigma dalam *machine learning* yang mengombinasikan sekumpulan model dasar (*base learners*) untuk meningkatkan stabilitas dan performa prediktif. Melalui integrasi beberapa model dengan karakteristik kesalahan yang berbeda, pendekatan ini efektif dalam meminimalkan bias serta varians guna memperkuat kemampuan generalisasi model dibandingkan penggunaan algoritma tunggal [33, 34]. Secara matematis, prediksi dari model *ensemble* dinyatakan sebagai:

$$H(x) = \phi(h_1(x), h_2(x), \dots, h_M(x)) \quad (2.2)$$

Variabel ϕ merepresentasikan fungsi agregasi seperti *averaging* atau *majority voting*. Strategi arsitektur *ensemble* secara umum terbagi menjadi tiga kategori utama: *bagging* yang bekerja secara paralel untuk mereduksi varians, *boosting* yang bekerja secara sekuensial untuk memperbaiki kesalahan model sebelumnya, serta *stacking* yang mengombinasikan model heterogen melalui *meta-learner*. Optimalisasi kinerja dilakukan dengan meminimalkan fungsi kerugian melalui penambahan fungsi regularisasi $\Omega(H)$ untuk mengontrol kompleksitas model:

$$L = \sum_{i=1}^n l(y_i, H(x_i)) + \Omega(H) \quad (2.3)$$

Pada analisis radiomik *Renal Cell Carcinoma* (RCC), fitur yang diekstraksi memiliki dimensi tinggi dan hubungan non-linear yang kompleks sehingga rentan terhadap *overfitting*. Implementasi arsitektur *ensemble* menjadi solusi teknis untuk menangkap variasi tekstur yang sangat halus antar subtype tumor. Penggunaan regularisasi dalam kerangka ini juga membantu menjaga stabilitas model terhadap

ketidakseimbangan kelas pada dataset klinis [20, 35].

2.4 Stacked Ensemble Learning

Stacked Ensemble Learning atau *stacking* mengintegrasikan beberapa model dasar melalui model tingkat kedua (*meta-learner*) untuk menghasilkan keputusan klasifikasi yang optimal. Berbeda dengan strategi lainnya, *stacking* memungkinkan penggunaan algoritma yang heterogen sehingga model dapat mempelajari pola data dari berbagai perspektif arsitektural yang berbeda [20, 34].

Jika terdapat M model dasar yang menghasilkan prediksi terhadap input x , maka keluaran tersebut digunakan sebagai fitur baru untuk melatih model tingkat kedua $g(\cdot)$. Prediksi akhir dari arsitektur ini dinyatakan sebagai:

$$H(x) = g(h_1(x), h_2(x), \dots, h_M(x)) \quad (2.4)$$

Model $g(\cdot)$ bertindak sebagai pengambil keputusan akhir yang mempelajari bobot optimal dari setiap *base learner*. Dalam pengolahannya, *stacking* sangat relevan untuk data radiomik ginjal yang memiliki korelasi fitur kompleks. Kombinasi algoritma seperti *Random Forest* dan *XGBoost* yang dikoordinasikan oleh *meta-learner* linear seperti *Logistic Regression* mampu mengompensasi kelemahan individual tiap model serta meningkatkan sensitivitas terhadap subtype pRCC [33, 35].

2.5 Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting (XGBoost) merupakan implementasi efisien dari algoritma *gradient boosting machine* yang beroperasi berdasarkan prinsip *additive training*. Model ini membangun ansambel secara bertahap dengan menambahkan pohon keputusan (*decision tree*) baru guna meminimalkan residual prediksi model sebelumnya melalui evaluasi gradien pada fungsi kerugian [10].

2.5.1 Fungsi Objektif dan Regularisasi

XGBoost mengoptimalkan model dengan meminimalkan fungsi objektif gabungan yang terdiri atas fungsi kerugian (*loss function*) dan suku regularisasi untuk menjaga keseimbangan antara akurasi dan kompleksitas:

$$\text{Obj} = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (2.5)$$

Pada Persamaan 2.5, $\Omega(f_k)$ berfungsi mengontrol kompleksitas setiap pohon f_k guna menekan risiko *overfitting* pada dataset radiomik yang memiliki dimensionalitas tinggi. Prediksi akhir diperoleh melalui akumulasi skor dari seluruh K pohon keputusan:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), \quad f_k \in \mathcal{F} \quad (2.6)$$

2.5.2 Optimisasi Berbasis Gradien Orde Dua

XGBoost menerapkan *second-order Taylor expansion* pada fungsi kerugian untuk mempercepat konvergensi selama proses pelatihan. Ekspansi pada iterasi ke- t dirumuskan sebagai berikut:

$$\text{Obj}^{(t)} \approx \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \quad (2.7)$$

Dalam persamaan tersebut, g_i merupakan gradien (turunan pertama) dan h_i merupakan *Hessian* (turunan kedua). Pemanfaatan informasi turunan orde dua ini memungkinkan penentuan bobot optimal di setiap simpul menjadi lebih presisi dibandingkan pendekatan *gradient boosting* konvensional [10].

2.5.3 Regularisasi L1 dan L2

Guna mencegah terjadinya *overfitting* pada dataset medis dengan jumlah sampel terbatas, XGBoost mengintegrasikan mekanisme penalti L1 (*Lasso*) dan L2 (*Ridge*) ke dalam fungsi regularisasinya:

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 + \alpha \sum_{j=1}^T |w_j| \quad (2.8)$$

Mekanisme regularisasi ini, dikombinasikan dengan teknik seleksi fitur internal melalui parameter seperti `colsample_bytree`, menjadikan XGBoost

sangat tangguh dalam menangani data radiomik pRCC yang bersifat kompleks dan *sparse* [7].

2.6 Random Forest

Random Forest (RF) merupakan algoritma *ensemble learning* yang mengombinasikan sekumpulan pohon keputusan (*decision trees*) secara independen melalui teknik *bootstrap aggregating* atau *bagging*. Algoritma ini bekerja dengan membangun beberapa pohon pada subset data yang berbeda dan menentukan prediksi akhir berdasarkan mekanisme pemungutan suara mayoritas (*majority voting*):

$$\hat{y} = \text{mode}\{h_1(x), h_2(x), \dots, h_T(x)\} \quad (2.9)$$

Di mana $h_t(x)$ merepresentasikan hasil prediksi dari pohon keputusan ke- t . Penggabungan hasil dari banyak pohon ini efektif dalam meningkatkan stabilitas model dan mereduksi varians tanpa meningkatkan bias secara signifikan [19].

Proses pemisahan fitur pada setiap simpul (*node*) di dalam pohon ditentukan berdasarkan penurunan tingkat ketidakmurnian (*impurity*). Dalam penelitian ini, kriteria yang digunakan adalah *Gini Index*, yang secara matematis dirumuskan sebagai berikut:

$$Gini = 1 - \sum_{i=1}^C p_i^2 \quad (2.10)$$

Variabel p_i merupakan proporsi sampel yang tergolong dalam kelas i pada suatu simpul tertentu. Nilai *Gini Index* yang lebih rendah menunjukkan tingkat kemurnian simpul yang lebih tinggi. Dalam analisis radiomik, RF secara intrinsik mampu menghitung tingkat kepentingan fitur (*feature importance*) melalui akumulasi penurunan *Gini impurity* di seluruh pohon guna mengidentifikasi fitur tekstur atau intensitas citra yang paling diskriminatif dalam membedakan subtype pRCC [19].

2.7 Logistic Regression

Logistic Regression merupakan algoritma klasifikasi yang memodelkan probabilitas kejadian kelas kategorikal melalui transformasi fungsi non-linear *sigmoid*. Kombinasi linear dari fitur input, yang dinyatakan sebagai z , dipetakan ke dalam rentang nilai $(0, 1)$ guna menentukan probabilitas klasifikasi sub tipe pRCC:

$$p = \sigma(z) = \frac{1}{1 + e^{-z}} \quad (2.11)$$

Nilai p merepresentasikan keyakinan model terhadap sampel sebagai kelas positif. Dalam konteks medis, pemetaan probabilitas ini memungkinkan klinisi melakukan penyesuaian ambang batas (*threshold*) keputusan guna menyeimbangkan nilai sensitivitas dan spesifisitas.

Guna menangani dimensionalitas fitur radiomik yang tinggi serta meminimalkan risiko *overfitting*, diterapkan teknik regularisasi L2 (*Ridge*). Metode ini bekerja dengan menambahkan penalti berupa kuadrat dari koefisien β ke dalam fungsi biaya (*cost function*) untuk mencegah bobot fitur menjadi terlalu ekstrem:

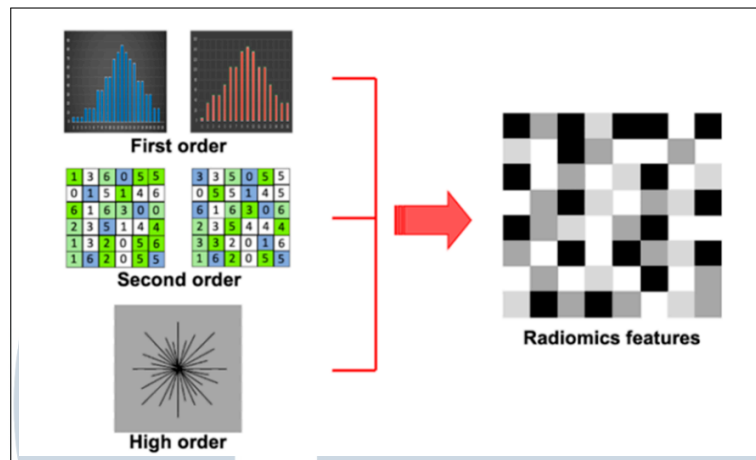
$$J(\beta) = -\frac{1}{n} \sum_{i=1}^n [y_i \ln(p_i) + (1 - y_i) \ln(1 - p_i)] + \lambda \sum_{j=1}^n \beta_j^2 \quad (2.12)$$

Kekuatan penalti dikontrol oleh parameter regularisasi λ . Dalam implementasi praktis menggunakan pustaka pemrograman, parameter ini sering direpresentasikan sebagai C (di mana $C = 1/\lambda$). *Logistic Regression* dipilih sebagai komponen *meta-learner* dalam arsitektur *stacked ensemble* penelitian ini karena sifatnya yang stabil dan mampu menghasilkan generalisasi yang baik pada dataset biomedis dengan ukuran sampel terbatas [18].

2.8 Feature Extraction

Feature extraction merupakan tahapan untuk mentransformasi *region of interest* (ROI) hasil segmentasi citra menjadi representasi numerik yang padat informasi. Proses ini menerjemahkan karakteristik visual jaringan menjadi parameter kuantitatif yang merepresentasikan heterogenitas tumor secara objektif sebagaimana diilustrasikan pada Gambar 2.3 [36, 37]. Melalui ekstraksi fitur, pola morfologi dan distribusi intensitas yang tidak terdeteksi secara visual dapat

dikuantifikasi sebagai fondasi pembangunan model prediktif radiomik [37].



Gambar 2.3. Ilustrasi mekanisme ekstraksi fitur radiomik pada citra medis.

Penelitian ini menggunakan *PyRadiomics* yang mengacu pada standar *Image Biomarker Standardization Initiative* (IBSI) guna menjamin reproduisibilitas dan homogenitas fitur yang dihasilkan [8]. Standarisasi pada tahap ini krusial untuk meminimalkan bias akibat variasi protokol akuisisi citra sehingga fitur yang diperoleh bersifat *robust* [37]. Fitur radiomik yang diekstraksi dikategorikan menjadi kelompok *shape features*, *first-order features*, serta *texture features* yang mencakup fitur *second-order* maupun *high-order* [37].

2.8.1 First-Order Features

First-order features atau fitur berbasis histogram mengekstraksi distribusi intensitas voxel dalam ROI tanpa mempertimbangkan hubungan spasial antar voxel. Fitur ini merepresentasikan karakteristik radiometrik global guna menilai kecerahan, homogenitas, dan rentang nilai keabuan pada jaringan tumor [38, 12]. Beberapa parameter statistik utama yang digunakan dalam klasifikasi ini meliputi:

1. *Mean Intensity*: Mengukur rata-rata nilai voxel dalam ROI yang mencerminkan tingkat intensitas rata-rata jaringan:

$$\text{Mean} = \frac{1}{N_p} \sum_{i=1}^{N_p} X(i) \quad (2.13)$$

2. *Variance*: Menggambarkan tingkat penyebaran intensitas voxel di sekitar nilai rata-rata guna menilai variasi tekstur global:

$$\text{Variance} = \frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) - \bar{X})^2 \quad (2.14)$$

3. *Skewness*: Mengukur derajat asimetri distribusi nilai keabuan di sekitar nilai rata-rata:

$$\text{Skewness} = \frac{\frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) - \bar{X})^3}{\left(\sqrt{\frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) - \bar{X})^2} \right)^3} \quad (2.15)$$

4. *Kurtosis*: Menilai ketajaman puncak distribusi histogram intensitas untuk mengevaluasi konsentrasi nilai keabuan:

$$\text{Kurtosis} = \frac{\frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) - \bar{X})^4}{\left(\frac{1}{N_p} \sum_{i=1}^{N_p} (X(i) - \bar{X})^2 \right)^2} \quad (2.16)$$

5. *Energy*: Mengukur magnitudo atau kepadatan nilai intensitas dalam citra:

$$\text{Energy} = \sum_{i=1}^N p(i)^2 \quad (2.17)$$

6. *Entropy*: Menganalisis tingkat keacakan atau kompleksitas informasi pada distribusi nilai piksel:

$$\text{Entropy} = - \sum_{i=1}^N p(i) \log_2 p(i) \quad (2.18)$$

Fitur-fitur ini memberikan landasan statistik dasar untuk membedakan karakteristik jaringan melalui distribusi nilai keabuan yang unik dalam ROI [38, 39].

2.8.2 Second-Order Texture Features

Second-order features atau *texture features* merepresentasikan heterogenitas internal jaringan tumor dengan menangkap keteraturan dan variasi spasial antar voxel. Fitur ini mampu menguantifikasi karakteristik histopatologis seperti kepadatan sel dan vaskularisasi melalui distribusi tekstur pada citra [40]. Kelompok matriks tekstur utama yang digunakan meliputi:

1. Gray Level Co-occurrence Matrix (GLCM)

GLCM mengukur probabilitas hubungan spasial antar pasangan voxel berdasarkan nilai intensitas, jarak, dan arah tertentu guna menilai kompleksitas struktur jaringan [39, 9].

(a) *Contrast*: Mengukur variasi lokal intensitas dalam citra.

$$\text{Contrast} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i-j)^2 p(i,j) \quad (2.19)$$

(b) *Correlation*: Menilai ketergantungan linear nilai keabuan antar pasangan voxel.

$$\text{Correlation} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j)ij - \mu_x \mu_y}{\sigma_x(i) \sigma_y(j)} \quad (2.20)$$

(c) *Joint Energy*: Menguantifikasi tingkat homogenitas tekstur.

$$\text{Joint Energy} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (p(i,j))^2 \quad (2.21)$$

(d) *Joint Entropy*: Mengukur derajat ketidakteraturan informasi pada tekstur citra.

$$\text{Joint Entropy} = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \log_2(p(i,j) + \epsilon) \quad (2.22)$$

2. Gray Level Run Length Matrix (GLRLM)

GLRLM menguantifikasi panjang deretan (*run*) voxel dengan intensitas identik dalam arah tertentu untuk membedakan tekstur halus (*short runs*) dan kasar (*long runs*) [9].

- (a) *Short Run Emphasis* (SRE): Memberikan bobot lebih pada tekstur yang lebih halus.

$$\text{SRE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{\mathbf{P}(i,j|\theta)}{j^2}}{N_r(\theta)} \quad (2.23)$$

- (b) *Long Run Emphasis* (LRE): Menekankan keberadaan tekstur yang lebih kasar atau homogen.

$$\text{LRE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \mathbf{P}(i,j|\theta) j^2}{N_r(\theta)} \quad (2.24)$$

3. Gray Level Size Zone Matrix (GLSZM)

GLSZM mengukur ukuran zona voxel yang saling terhubung dengan intensitas identik tanpa mempertimbangkan arah, guna mengidentifikasi tingkat heterogenitas area tumor [39].

- (a) *Small Area Emphasis* (SAE): Menguantifikasi dominasi zona-zona kecil dalam citra.

$$\text{SAE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{\mathbf{P}(i,j)}{j^2}}{N_z} \quad (2.25)$$

- (b) *Large Area Emphasis* (LAE): Menilai dominasi area homogen yang lebih luas.

$$\text{LAE} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \mathbf{P}(i,j) j^2}{N_z} \quad (2.26)$$

2.8.3 High-Order Features

High-order feature extraction melibatkan penerapan transformasi matematis pada citra asli untuk menonjolkan karakteristik tekstur dan heterogenitas jaringan

yang tidak terlihat secara langsung. Transformasi ini mengubah representasi intensitas citra pada berbagai skala spasial guna meningkatkan sensitivitas fitur terhadap variasi morfologi tumor [34].

A Laplacian of Gaussian (LoG)

Laplacian of Gaussian (LoG) mengombinasikan *Gaussian smoothing* untuk reduksi *noise* dengan operator Laplacian untuk deteksi perubahan intensitas. LoG efektif dalam menonjolkan struktur mikro dan batas lesi pada berbagai skala spasial yang dikontrol oleh nilai σ [41]. Operasi konvolusi LoG dirumuskan sebagai:

$$I_{LoG}(x, y) = \nabla^2 \left(\left[\frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \right] * I(x, y) \right) \quad (2.27)$$

B Wavelet Decomposition

Transformasi *wavelet* melakukan dekomposisi multiresolusi yang memisahkan citra ke dalam komponen frekuensi rendah (*approximations*) dan frekuensi tinggi (*details*) pada berbagai arah spasial (LLH, LHL, LHH, dsb.). Pada data volumetrik 3D, transformasi ini menangkap variasi tekstur multiskala yang merepresentasikan struktur global maupun detail lokal jaringan tumor [4].

$$I_{\text{wavelet}} = W_{3D}(I(x, y, z)) \quad (2.28)$$

C Filter Non-Linear dan Gradien

Selain transformasi multiskala, digunakan beberapa filter intensitas dan spasial untuk memodifikasi distribusi nilai keabuan voxel guna memperkuat pola heterogenitas tertentu:

1. *Square Root dan Logarithm*: Digunakan untuk mengompresi rentang intensitas tinggi dan menstabilkan variansi guna menonjolkan variasi tekstur halus pada area berintensitas rendah [42].

$$I_{\log} = \log(I + \varepsilon); \quad I_{\text{sqr}} = \sqrt{I + \varepsilon} \quad (2.29)$$

2. *Square dan Exponential*: Memperbesar perbedaan intensitas pada nilai tinggi guna memperkuat pola heterogenitas yang berkaitan dengan distribusi intensitas ekstrem pada jaringan [34].

$$I_{\text{square}} = I^2; \quad I_{\text{exp}} = \exp(I) \quad (2.30)$$

3. *Gradient Filter*: Mengukur perubahan intensitas spasial melalui turunan parsial terhadap sumbu x, y, z untuk menonjolkan batas lesi dan kompleksitas morfologi [43].

$$I_{\text{grad}} = \sqrt{I_x^2 + I_y^2 + I_z^2} \quad (2.31)$$

2.9 Feature Selection

Tahap *feature selection* berperan kritis dalam menentukan subset fitur paling relevan dari ribuan kandidat fitur radiomik guna menekan risiko *overfitting* akibat *curse of dimensionality*. Selain mereduksi dimensi, proses ini berfungsi sebagai penyaring kualitas untuk mengeliminasi fitur yang redundan dan tidak stabil akibat variasi segmentasi maupun perbedaan protokol akuisisi citra [44, 37]. Pemilihan fitur yang *robust* memastikan model memiliki kemampuan generalisasi yang tinggi pada data klinis yang heterogen. Dalam penelitian ini, digunakan tiga pendekatan seleksi fitur yang bersifat komplementer untuk menangkap berbagai aspek relevansi fitur:

2.9.1 Recursive Feature Elimination (RFE)

RFE merupakan metode seleksi berbasis model yang bekerja secara iteratif dengan menghapus fitur yang memiliki kontribusi terendah terhadap performa prediktif model [12]. Prosedur eliminasi dilakukan melalui tahapan berikut:

1. Melatih model menggunakan seluruh set fitur yang tersedia.
2. Melakukan pemeringkatan fitur berdasarkan bobot atau koefisien kepentingan.
3. Mengeliminasi fitur dengan nilai kepentingan terkecil secara bertahap.
4. Mengulangi proses hingga diperoleh jumlah fitur optimal yang menghasilkan performa puncak.

2.9.2 Mutual Information (MI)

Mutual Information mengukur ketergantungan statistik antara fitur dan label target melalui konsep entropi. Pendekatan non-parametrik ini mampu menangkap hubungan non-linear yang kompleks, yang sering ditemukan pada dataset medis dengan distribusi data tidak beraturan [44]. Persamaannya dinyatakan sebagai:

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \quad (2.32)$$

2.9.3 Analysis of Variance (ANOVA)

ANOVA mengevaluasi kemampuan diskriminatif fitur dengan membandingkan variansi nilai antar kelas terhadap variansi di dalam kelas melalui statistik F [15]. Fitur dengan nilai F yang tinggi menunjukkan perbedaan signifikan secara statistik dalam membedakan subtype tumor ginjal:

$$F = \frac{\text{Between-class variance}}{\text{Within-class variance}} \quad (2.33)$$

2.10 Metrik Evaluasi

Evaluasi performa dilakukan untuk mengukur akurasi dan konsistensi model dalam mengidentifikasi subtype pRCC. Penelitian ini menggunakan metrik berbasis *confusion matrix* guna memberikan penilaian yang komprehensif, terutama dalam menghadapi tantangan ketidakseimbangan kelas pada dataset klinis [45, 46].

2.10.1 Confusion Matrix

Confusion matrix menyediakan perincian tabular mengenai perbandingan antara prediksi model dengan label aktual melalui empat komponen dasar:

1. *True Positive (TP)*: Sampel pRCC yang terdeteksi dengan benar.
2. *True Negative (TN)*: Sampel non-pRCC yang terdeteksi dengan benar.

3. *False Positive* (FP): Sampel non-pRCC yang salah diidentifikasi sebagai pRCC.

4. *False Negative* (FN): Sampel pRCC yang salah diidentifikasi sebagai non-pRCC.

2.10.2 Accuracy

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.34)$$

2.10.3 Precision

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2.35)$$

Metrik ini merepresentasikan reliabilitas prediksi positif model guna meminimalkan risiko tindakan medis yang tidak diperlukan akibat kesalahan identifikasi tumor (*false positive*) [46].

2.10.4 Recall

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2.36)$$

Recall atau sensitivitas merupakan metrik kritis dalam diagnosis kanker untuk meminimalkan *false negative*, yang berpotensi menyebabkan keterlambatan penanganan medis bagi pasien [47].

2.10.5 F1-Score

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (2.37)$$

F1-Score memberikan keseimbangan antara *precision* dan *recall*, menjadikannya indikator performa yang lebih stabil pada dataset dengan distribusi kelas yang tidak seimbang [45].