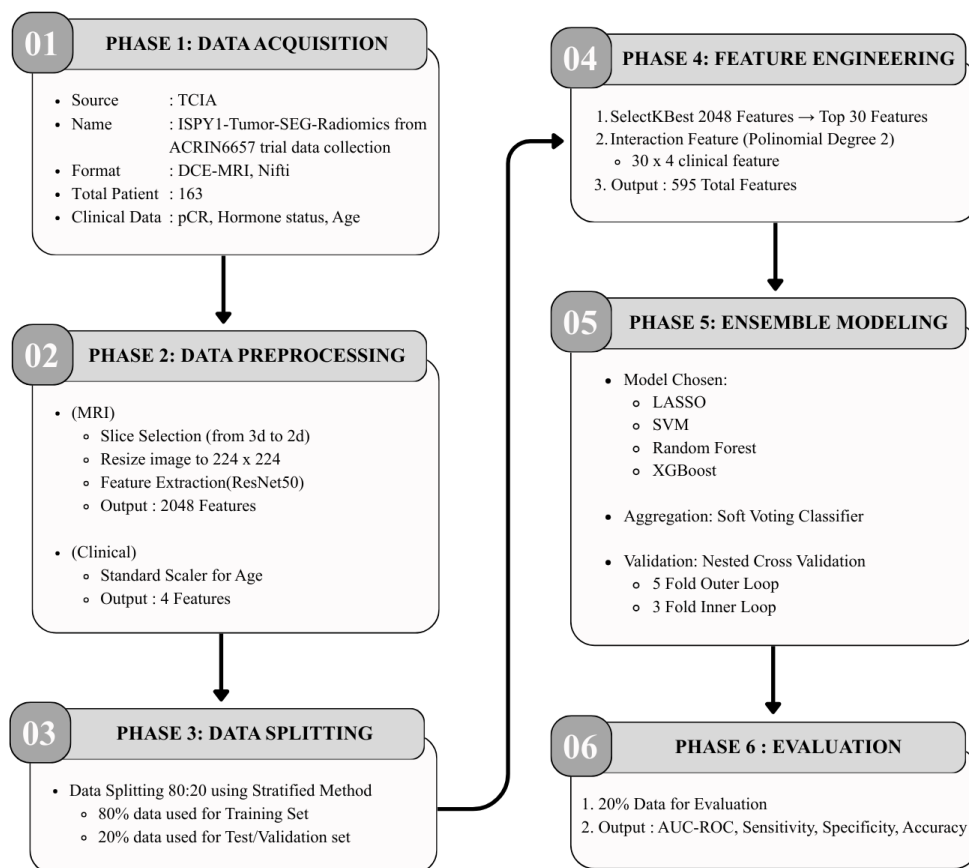


BAB 3

METODOLOGI PENELITIAN

Bab ini menguraikan metodologi penelitian yang dirancang untuk membangun sistem prediksi respons kemoterapi. Tahapan penelitian dimulai dari studi literatur, pengumpulan dan prapemrosesan *dataset* MRI, ekstraksi fitur berbasis *Deep Learning*, hingga pengembangan model *Ensemble* dan evaluasi kinerja. Alur kerja penelitian secara keseluruhan divisualisasikan pada Gambar 3.1.



Gambar 3.1. Diagram alur (pipeline) metode penelitian yang diusulkan.

Seluruh eksperimen komputasi dilakukan menggunakan platform berbasis awan (*cloud computing*) Google Colaboratory dengan spesifikasi lingkungan pengembangan sebagai berikut:

- Sistem Operasi: Linux (Ubuntu 20.04.5 LTS).
- Perangkat Keras: CPU Intel Xeon @ 2.20GHz, RAM 13GB (High-RAM), dan GPU NVIDIA T4 (16GB VRAM).
- Perangkat Lunak: Python 3 dengan pustaka utama *PyTorch* [13] untuk Deep Learning, *Scikit-Learn* [14] untuk Machine Learning, serta *Nibabel* untuk pengolahan citra medis.

3.1 Tinjauan Pustaka

Evaluasi respons pasien terhadap *Neoadjuvant Chemotherapy* (NACT) telah berevolusi dari analisis volumetrik sederhana menuju pendekatan radiomik mendalam. Studi fundamental oleh Hylton et al. (2016) [15] pada *dataset* I-SPY1 membuktikan adanya korelasi kuat antara volume tumor fungsional dengan pCR. Namun, pendekatan ini memiliki keterbatasan karena hanya mengukur perubahan makroskopis dan mengabaikan heterogenitas tekstur internal tumor.

Untuk mengatasi hal tersebut, Braman et al. (2017) [5] mengembangkan model radiomik berbasis tekstur (*texture analysis*) pada area intratumoral dan peritumoral. Meskipun metode ini berhasil meningkatkan akurasi dibandingkan pengukuran volume, kelemahan utamanya terletak pada ketergantungan terhadap fitur buatan tangan (*hand-crafted features*) yang rentan terhadap variabilitas operator.

Perkembangan terkini beralih pada penerapan *Deep Learning* untuk ekstraksi fitur otomatis. Duanmu et al. (2020) [4] mendemonstrasikan bahwa integrasi fitur *Deep Learning* dengan data klinis dan demografi mampu menghasilkan prediksi yang lebih akurat (AUC 0.80) dibandingkan penggunaan data citra saja. Temuan ini memvalidasi pentingnya fusi data multimodal dalam onkologi presisi.

Meskipun demikian, penelitian Duanmu et al. masih mengandalkan arsitektur klasifikasi tunggal. Pendekatan model tunggal menghadapi risiko varians tinggi (*high variance*) dan ketidakstabilan prediksi pada *dataset* medis yang terbatas. Oleh karena itu, penelitian ini mengisi kesenjangan tersebut

dengan mengusulkan pendekatan *Ensemble Learning*. Metode ini tidak hanya memanfaatkan ekstraksi fitur otomatis ResNet50, tetapi juga mengintegrasikan kekuatan empat algoritma berbeda (Voting) untuk menghasilkan prediksi yang lebih stabil, akurat (AUC 0.88), dan dapat digeneralisasi.

Rangkuman perbandingan posisi penelitian ini terhadap penelitian terdahulu disajikan pada Tabel 3.1.

Tabel 3.1. Perbandingan Penelitian Terdahulu

Peneliti	Metode	Hasil Utama	Celah / Perbedaan Pendekatan
Hylton et al. (2016) [15]	Analisis Volume Tumor (<i>Functional Tumor Volume</i>).	Membuktikan volume tumor berkorelasi dengan pCR, namun performa prediksi moderat (AUC 0.70).	Hanya menggunakan fitur volumetrik, belum mengeksplorasi fitur tekstur atau <i>Deep Learning</i> .
Braman et al. (2017) [5]	Radiomik Intratumoral & Peritumoral.	Menunjukkan bahwa analisis tekstur sekitar tumor meningkatkan akurasi (AUC 0.74).	Masih menggunakan fitur manual (<i>hand-crafted</i>) yang bergantung pada segmentasi manual yang presisi.
Duanmu et al. (2020) [4]	Deep Learning + Data Klinis & Demografi.	Model fusi multimodal terbukti lebih unggul (AUC 0.80) dibanding model citra tunggal.	Menggunakan pendekatan klasifikasi tunggal, belum menerapkan <i>Ensemble Voting</i> untuk stabilitas.
Penelitian Ini (Proposed)	ResNet50 (2.5D) + Ensemble Learning.	AUC 0.88 pada Lockbox Set.	Mengusulkan arsitektur <i>Ensemble Soft-Voting</i> untuk mengatasi varians pada <i>dataset</i> kecil.

3.2 Pengumpulan Data

Dataset yang digunakan dalam penelitian ini adalah diperoleh dari TCIA dengan judul *ISPY1-Tumor-SEG-Radiomics* [16]. *Dataset* merupakan bagian dari uji klinis publik *I-SPY1 TRIAL* (ACRIN 6657) yang dideskripsikan oleh Hylton et

al. [15]. Data diunduh melalui website resmi *The Cancer Imaging Archive* (TCIA) [17].

Pemilihan *dataset* ini didasarkan pada ketersediaan anotasi tumor oleh pakar dan standarisasi kualitas yang telah dilakukan oleh penyedia *dataset*. *Dataset* mencakup 156 pasien dengan karakteristik sebagai berikut:

- Data Citra Terstandarisasi: Citra MRI telah dikonversi ke format NIfTI (.nii.gz) dan telah melalui proses prapemrosesan meliputi koreksi medan bias (*bias field correction*), *resampling 1mm*, dan normalisasi Z-score.
- Segmentasi Tumor: Tersedia segmentasi tumor 3D yang dibuat secara semi-otomatis dan telah diverifikasi oleh radiolog ahli, membatasi *Region of Interest* (ROI) pada area tumor aktif.
- Data Klinis: Variabel klinis pasien meliputi Usia serta status biomarker (ER, PR, HER2).

3.3 Prapemrosesan Data

Prapemrosesan dilakukan untuk menyesuaikan dimensi input arsitektur ResNet50.

3.3.1 Prapemrosesan Data Klinis

Data klinis diproses agar dapat digabungkan dengan fitur citra:

- Normalisasi: Variabel usia dinormalisasi menggunakan teknik *Standard Scaler*.
- Encoding Kategorikal: Status reseptor (ER, PR, HER2) dikonversi menjadi representasi biner.

3.3.2 Prapemrosesan Data Radiomik

Data citra MRI mengalami transformasi dimensi dan intensitas untuk mempersiapkan tensor input yang valid bagi model:

1. Konversi Input 2.5D (*Stacking*): Mengidentifikasi slice tengah tumor (z) berdasarkan area terluas, lalu mengambil irisan tetangga ($z - 1$ dan $z + 1$).

2. **Resize & Crop:** Citra diubah ukurannya menjadi dimensi tetap 224×224 piksel menggunakan interpolasi bilinear untuk memenuhi persyaratan input spasial ResNet.
3. **Konversi Tensor:** Citra dikonversi dari array NumPy menjadi Tensor PyTorch dengan rentang nilai $[0, 1]$.
4. **Normalisasi ImageNet:** Citra dinormalisasi ulang menggunakan statistik distribusi warna *dataset* ImageNet ($\mu = [0.485, 0.456, 0.406]$, $\sigma = [0.229, 0.224, 0.225]$) agar fitur yang diekstraksi selaras dengan bobot *pre-trained* model [18].

3.4 Ekstraksi Fitur (Deep Learning)

Setelah melalui tahap prapemrosesan, proses ekstraksi fitur visual dilakukan menggunakan arsitektur *Convolutional Neural Network* (CNN) ResNet50 yang telah dilatih sebelumnya (*pre-trained*) pada *dataset* ImageNet.

3.4.1 Mekanisme Feature Extraction

Model ResNet50 digunakan untuk mengekstraksi fitur murni (*feature extractor*) dengan memodifikasi lapisan akhir (*Fully Connected Layer*) menjadi lapisan identitas. Tensor citra input ($3 \times 224 \times 224$) dipropagasi melewati lapisan konvolusi, dan output diambil dari lapisan *Global Average Pooling*. Proses ini mentransformasi informasi visual menjadi vektor fitur kontinu berdimensi 2048 (1×2048) untuk setiap pasien, yang merepresentasikan karakteristik abstrak tumor.

3.5 Seleksi dan Rekayasa Fitur

Untuk menangani dimensi fitur yang tinggi (2048 fitur radiomik) dan jumlah sampel terbatas, diterapkan strategi reduksi dimensi bertingkat.

3.5.1 Seleksi Fitur (Univariate Selection)

Seleksi fitur dilakukan menggunakan metode *SelectKBest* dengan fungsi skor ANOVA F-value (*f_classif*) [14]. Hanya 30 fitur radiomik dengan skor korelasi tertinggi terhadap label target (pCR) yang dipertahankan. Penting dicatat bahwa

proses *fitting* seleksi fitur hanya dilakukan pada data latih (*Discovery Set*) untuk mencegah *data leakage*.

3.5.2 Interaksi Fitur (Interaction Engineering)

Fitur radiomik terpilih (30 fitur) dan fitur klinis (4 fitur) digabungkan, kemudian dilakukan ekspansi fitur menggunakan *Polynomial Features* derajat 2 dengan mode *interaction only*. Proses ini menghasilkan fitur-fitur baru yang merupakan hasil perkalian silang antar variabel (misalnya: $FiturMRI_A \times Usia$), yang bertujuan menangkap interaksi biologis non-linier antara fenotipe citra dan data klinis [11].

3.6 Proses Pemilihan Model

Model prediksi dibangun menggunakan arsitektur *Ensemble Learning* heterogen yang menggabungkan empat algoritma klasifikasi berbeda:

1. LASSO Logistic Regression: Dengan regularisasi L1 ($C=0.1$) untuk seleksi fitur implisit [19].
2. Support Vector Machine (SVM): Menggunakan kernel linear dan kalibrasi probabilitas ($probability=True$) [20].
3. Random Forest (RF): Dengan 100 estimator dan kedalaman pohon maksimum 5 [21].
4. XGBoost Classifier: Dengan *learning rate* 0.05 dan kedalaman pohon 3 [22].

Keempat model tersebut digabungkan dalam wadah *Voting Classifier* dengan metode *Soft Voting*, di mana prediksi akhir ditentukan berdasarkan rata-rata probabilitas terbobot dari setiap model dasar. Penyetelan *hyperparameter* dilakukan menggunakan *Grid Search* untuk mencari kombinasi parameter optimal.

3.7 Evaluasi Model

3.7.1 Skema Lockbox Split

Dataset dibagi menjadi dua bagian independen menggunakan metode *Stratified Sampling*:

- Discovery Set (80%): Digunakan sepenuhnya untuk pelatihan model, seleksi fitur, dan validasi silang (*cross-validation*).
- Lockbox Set (20%): Disimpan sebagai data uji "buta" yang tidak pernah dilihat oleh model selama proses pelatihan. Data ini hanya dibuka satu kali pada tahap akhir evaluasi.

3.7.2 Nested Cross-Validation

Pada tahap pelatihan (Discovery Set), diterapkan *Nested Cross-Validation*:

- *Outer Loop* (5-Fold): Untuk mengevaluasi performa generalisasi model.
- *Inner Loop* (3-Fold): Untuk pencarian *hyperparameter* terbaik (*Grid Search*) di dalam setiap lipatan pelatihan.

Metrik evaluasi utama yang digunakan adalah *AUC-ROC* (*Area Under Curve*), dilengkapi dengan Akurasi, Sensitivitas, dan Spesifisitas.

