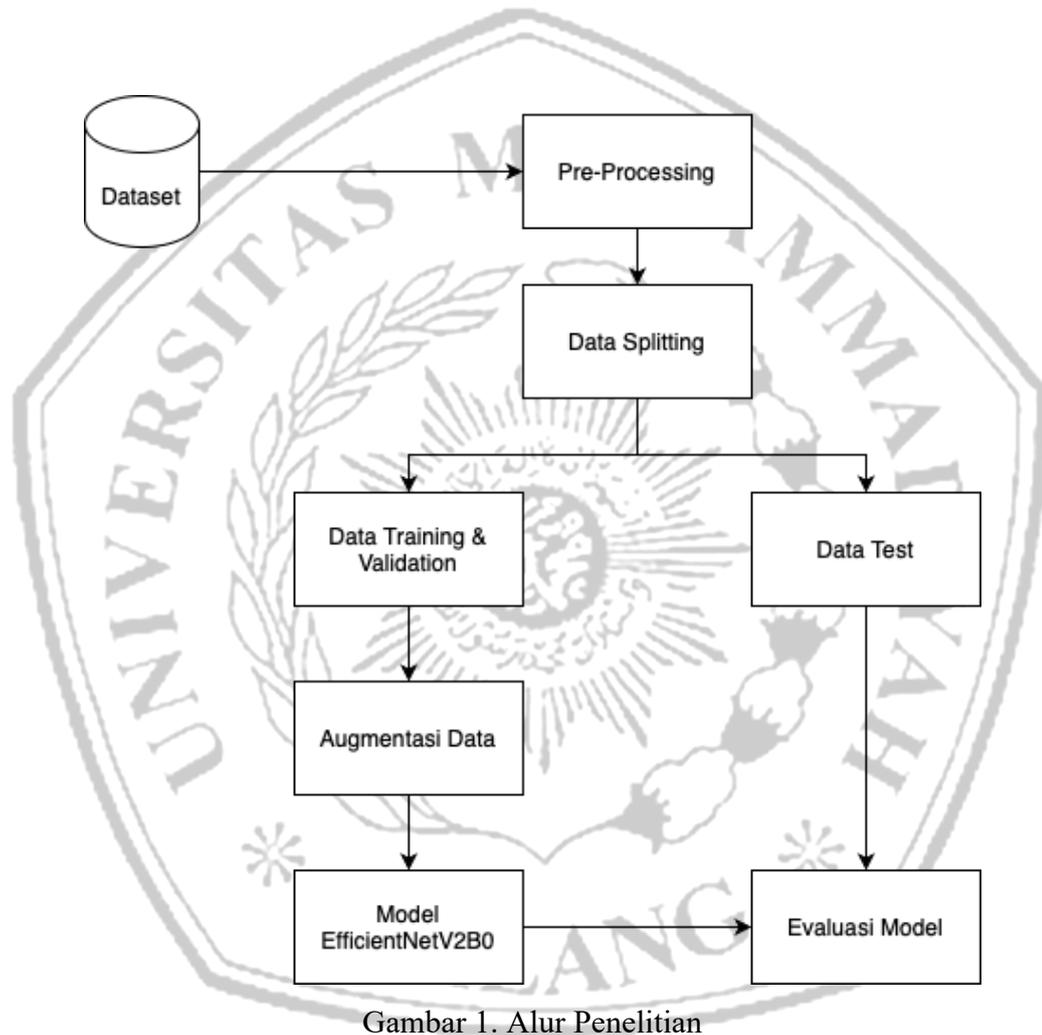


BAB III

METODOLOGI PENELITIAN

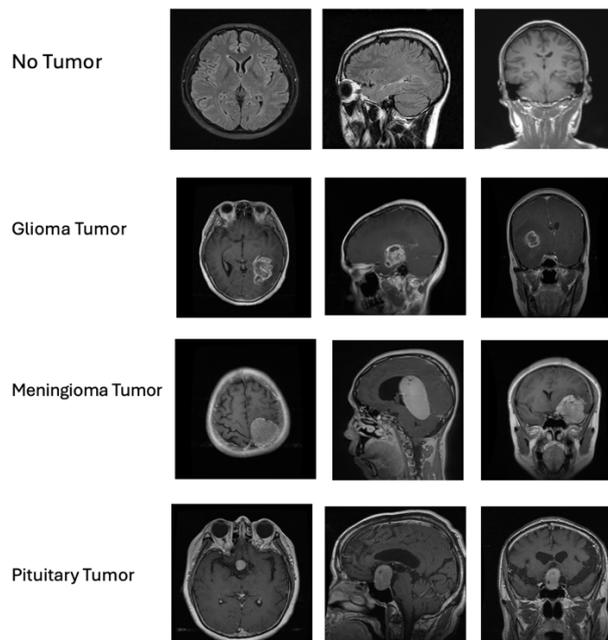
Berikut adalah langkah-langkah yang dilakukan dalam penelitian ini, yang ditampilkan dengan diagram alir. Diagram alir penelitian memberikan gambaran tentang proses dan tahapan-tahapan yang diikuti selama penelitian berlangsung. Pada penelitian ini, diagram alir tersebut dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitian

3.1. Dataset

Dataset yang digunakan yaitu “Brain Tumor Classification (MRI)” yang didapat dari situs web Kaggle. Dataset ini terdiri dari 4 kelas yaitu terdiri dari no tumor (500 gambar), glioma tumor (926 gambar), meningioma tumor (937 gambar), dan pituitary tumor (901 gambar) dengan total jumlah data sebanyak 3264 gambar.



Gambar 2. Gambar pada setiap label otak

3.2. Pre-Processing

Pada tahap pre-processing data, dilakukan perubahan ukuran gambar dengan tujuan mengurangi dimensi data gambar tanpa menghilangkan fitur-fitur penting. Langkah ini dilakukan untuk meningkatkan efisiensi komputasi selama proses pelatihan data [16]. Ukuran gambar pada dataset yang digunakan semula yaitu 512 x 512 pixel lalu disesuaikan menjadi 150 x 150 piksel. Penggunaan gambar yang lebih kecil memungkinkan pelatihan yang lebih cepat dan bahkan dapat menghasilkan hasil yang lebih baik dalam generalisasi model, karena jumlah data yang lebih sedikit yang perlu disesuaikan.

3.3. Data Splitting

Pada penelitian ini data splitting dilakukan untuk membagi data menjadi 3 bagian yaitu data latih, data validasi dan data uji. Dataset pada penelitian ini dibagi menjadi 80% data latih, 10% data validasi dan 10% data uji dari total keseluruhan data yaitu 3264 data. Data validasi digunakan untuk memvalidasi melalui pengujian berdasarkan dataset yang telah dipelajari pada tahap pelatihan.

3.4. Augmentasi Data

Augmentasi data adalah teknik yang digunakan untuk menambah jumlah data dengan memodifikasi gambar asli sehingga memiliki tampilan yang berbeda. Teknik ini membantu meningkatkan kinerja klasifikasi dengan menyediakan lebih banyak variasi data. Dengan menyediakan lebih banyak variasi data, augmentasi data dapat membantu model untuk lebih baik mengenali dan mengklasifikasikan objek dalam gambar. Dalam penelitian ini, berbagai teknik augmentasi data diterapkan untuk menciptakan variasi dalam dataset gambar. Teknik-teknik augmentasi yang digunakan yaitu $rotation = 20$, $zoom_range = 0.2$, $width_shift_range = 0.1$, $height_shift_range = 0.1$, $horizontal_flip = true$.

3.5. EfficientNetV2B0

Convolutional Neural Networks (CNN) adalah model berbasis jaringan saraf tiruan yang terinspirasi dari ilmu biologi. CNN telah terbukti menjadi yang terbaik di bidangnya dengan mendapatkan nilai yang lebih baik daripada metode pembelajaran mesin tradisional dalam kompetisi untuk klasifikasi dan deteksi objek. Convolutional Neural Networks (CNN) sebagian besar terdiri dari lapisan konvolusi, lapisan aktivasi, lapisan penyatuan, dan lapisan yang terhubung penuh [17].

Model EfficientNet dikembangkan oleh tim riset Google pada tahun 2019. Model EfficientNet mencapai akurasi yang lebih tinggi dan efisiensi yang lebih baik dengan mengurangi ukuran parameter dan Floating point Operations Per Second (FLOPS) dibandingkan dengan model CNN yang ada di ImageNet secara keseluruhan. Khususnya, EfficientNetV2B0 8,4 kali lebih kecil dan 6,1 kali lebih

cepat daripada model CNN terbaik yang tersedia, sambil mencapai akurasi 84,4% dengan parameter 66 M dalam kompetisi ImageNet 53[18].

Penelitian ini memanfaatkan desain arsitektur model Convolutional Neural Network (CNN) untuk mengklasifikasikan citra jenis tumor otak. Arsitektur model ini terdiri dari lapisan input dengan resolusi (150 x 150), GlobalAveragePooling2D, Dropout, dan Dense layer. Fungsi dari GlobalAveragePooling2D adalah untuk mengurangi risiko overfitting pada model dengan mengurangi jumlah parameter [19]. Selain itu, GlobalAveragePooling2D membantu mengurangi kompleksitas model dan mempertahankan informasi yang diperlukan untuk melakukan klasifikasi. Penggunaan L2 Regularization juga membantu hasil pelatihan untuk menangani overfitting. Dengan menjaga bobot tetap kecil, L2 Regularization mengurangi risiko overfitting dengan mengurangi kompleksitas model tanpa mengorbankan kemampuannya untuk belajar pola yang signifikan dari data.. Lapisan Dense bertanggung jawab untuk menjalankan proses pembelajaran dan ekstraksi fitur yang lebih mendalam dan kompleks dari data gambar yang telah diproses sebelumnya oleh GlobalAveragePooling2D. Selanjutnya, Dropout merupakan teknik regulasi yang bertujuan untuk menghindari overfitting pada model. Detail arsitektur model yang digunakan pada penelitian ini dapat dilihat pada tabel 1.

Tabel 1. Arsitektur Model

Layer	Filter	Activation
Input (150, 150)	-	-
GlobalAveragePooling2D	-	-
Dropout	0.5	-
Dense	256	relu
Dense	4	softmax

3.6. Evaluasi Model

Dalam pengembangan model pada penelitian ini, evaluasi dilakukan menggunakan metrik seperti Akurasi, Presisi, Recall, dan Skor F1. Evaluasi ini bertujuan untuk menilai hasil dan kinerja model pada dataset klasifikasi tumor otak. Presisi, Recall, dan Skor F1 dihitung menggunakan algoritma tertentu untuk memberikan gambaran yang komprehensif tentang kualitas model. Algoritma perhitungan Presisi, Recall, dan Skor F1 digunakan sebagai berikut:

Accuracy

$$= \frac{\text{True Positive} + \text{True Negative}}{\text{True Positive} + \text{True Negative} + \text{False Positive} + \text{False Negative}}$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

