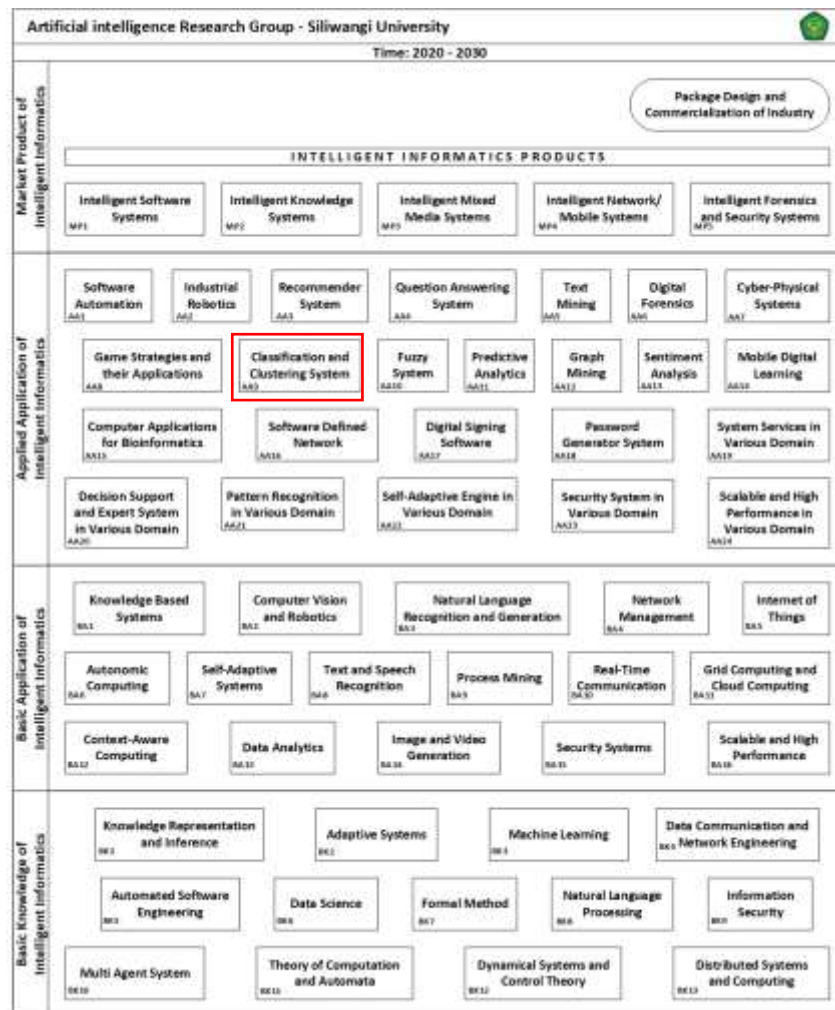


## BAB III

### METODOLOGI PENELITIAN

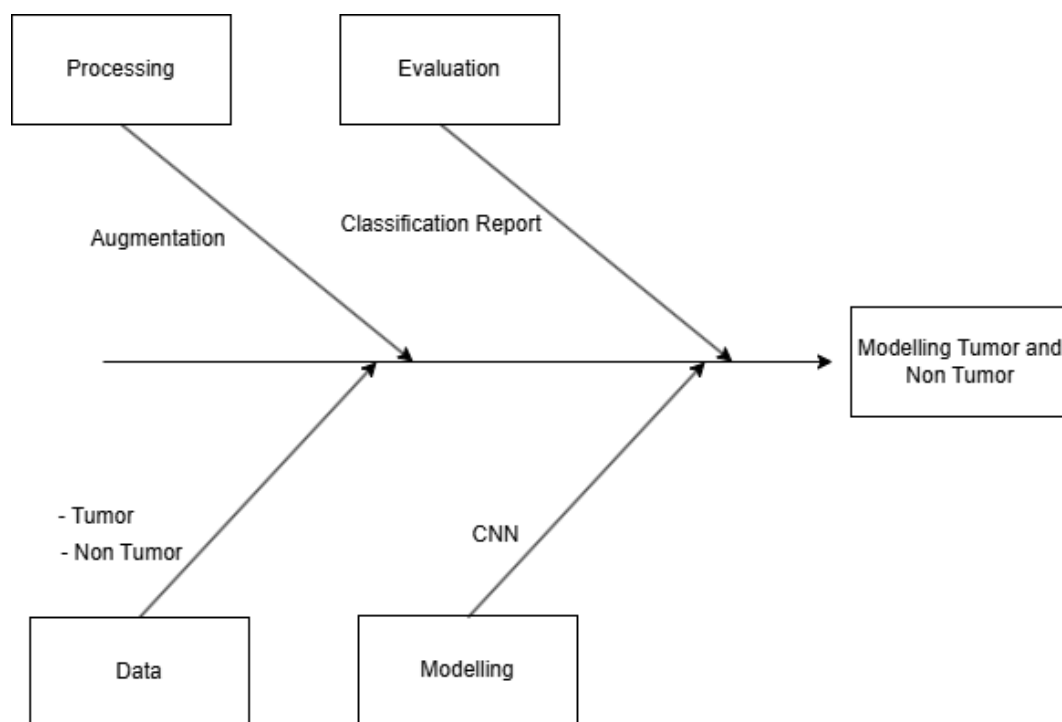
#### 3.1 Roadmap Penelitian

Rencana penelitian ini secara keseluruhan mendukung arah pengembangan riset Universitas Siliwangi di bidang *Artificial Intelligence*. Topik yang diangkat, yaitu *Classification*, sebagaimana ditunjukkan pada *roadmap* penelitian pada Gambar 3.1.



Gambar 3. 1 Roadmap Penelitian (AIS Universitas Siliwangi, 2020)

Gambar 3.1 menunjukkan tahapan pemilihan topik klasifikasi, yang menjadi dasar dari penelitian ini. Tujuan utama dari penelitian ini adalah pemodelan yang dapat digunakan sebagai proses klasifikasi. Dalam hal ini, penelitian difokuskan pada pengembangan model CNN menggunakan *augmentasi* untuk klasifikasi citra MRI untuk mendeteksi adanya indikasi tumor maupun non-tumor. Selanjutnya, Gambar 3.2 menyajikan *diagram fishbone* yang menggambarkan kerangka penelitian sebagai bagian dari roadmap pengembangan.



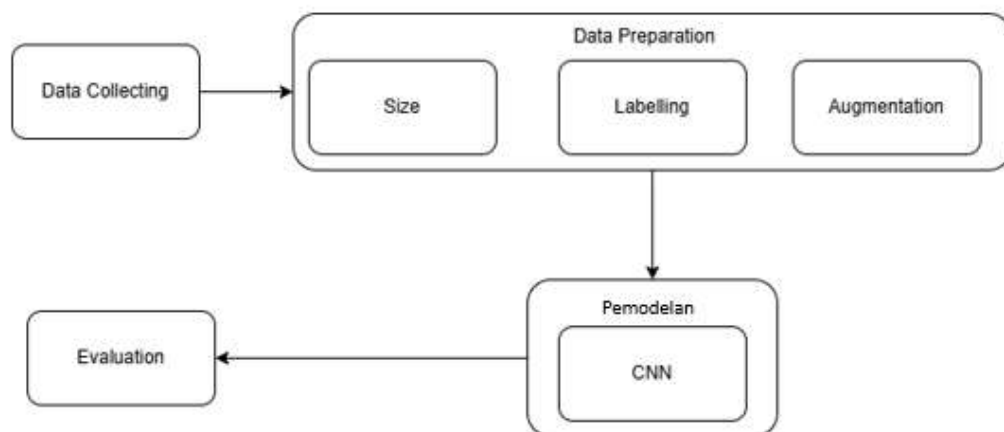
Gambar 3. 2 *Fishbone* Penelitian

Gambar 3.2 merupakan diagram *fishbone* yang menggambarkan model klasifikasi dalam penelitian ini. Penelitian ini, dirancang untuk membangun sebuah model yang mampu mengklasifikasikan kondisi tumor dan non-tumor. Data yang digunakan berasal dari citra MRI, baik yang menunjukkan adanya tumor maupun yang tidak. Data tersebut akan melalui proses augmentasi untuk meningkatkan

variasinya. Selanjutnya, model klasifikasi dibangun menggunakan algoritma *Convolutional Neural Network* (CNN) untuk mengklasifikasikan citra MRI .

### 3.2 Tahapan Penelitian

Metode penelitian untuk memprediksi penyakit tumor otak dengan mengembangkan model CNN. pada penelitian ini melibatkan beberapa langkah seperti yang ditunjukkan dalam gambar 3.3, menunjukan model CNN dikembangkan dengan menggunakan *augmentasi* dalam meningkatkan akurasi prediksi penyakit tumor otak.



Gambar 3. 3 Tahapan Penelitian (Saeedi et al., 2023)

Gambar 3.3 merupakan beberapa tahapan utama dalam proses penelitian. Pengembangan model memerlukan data gambar hasil pemeriksaan MRI terhadap penyakit tumor otak, sehingga pada penelitian ini membutuhkan proses *data collecting* (Nabila et al., 2024). Hasil dari *data collection* akan diteruskan ke dalam proses *data preparation* untuk mempersiapkan data sebelum dilatih oleh model. Data akan disamakan ukurannya supaya ukuran inputnya konsisten.

Kemudian, data diberi label menjadi dua *class*, yaitu tumor dan non tumor (Shanjida et al., 2022). Selanjutnya, pada proses *training* menggunakan model yang dikombinasikan dengan teknik *augmentasi*. Tahapan terakhir yaitu evaluasi model dengan mengukur akurasi dari performa model (Maulana et al., 2024). Penjelasan detail untuk setiap prosesnya yaitu sebagai berikut:

### 1. *Data Collecting*

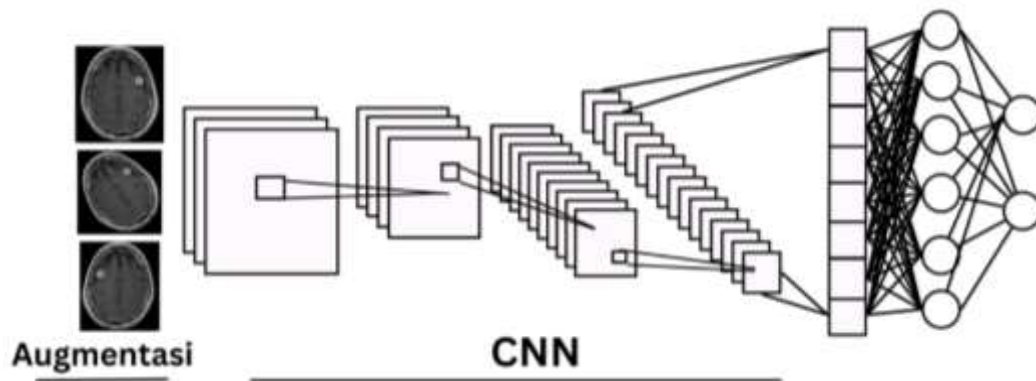
Data membutuhkan data gambar hasil MRI dalam penyakit tumor otak. Dua kategori data gambar yang dibutuhkan, yaitu hasil MRI terdapat tumor otak dan tidak terdapat tumor otak (Nabila et al., 2024). Data yang dibutuhkan untuk model tersedia di *website* kaggle. Pada tahap ini data diambil dari *website* tersebut. Jumlah data tumor 982 dan non tumor sebanyak 493.

### 2. *Data Preparation*

Pada proses *data preparation* yaitu proses menyiapkan data sebelum *training* model. Supaya data lebih mendapatkan input yang konsisten maka semua gambar ukurannya disamakan (Agustina et al., 2024). Kemudian, data dikasih label untuk membedakan gambar *class* tumor dan *class* non tumor.

### 3. Pemodelan

Pemodelan menggunakan data yang sudah disiapkan pada tahapan sebelumnya. Berikut gambar 3.2 merupakan arsitektur dari pengembangan model CNN menggunakan *augmentasi* untuk memprediksi penyakit tumor otak.



Gambar 3. 4 Arsitektur Model (Sejuti & Islam, 2021)

Gambar 3.2 merupakan arsitektur dari pengembangan model CNN. *Augmentasi* membuat model memilih random pada gambar untuk melakukan rotasi maupun *flip horizontal*, sehingga model dapat menerima gambar yang lebih bervariasi (Saeedi et al., 2023). Rotasi dilakukan dengan derajat 10% dari 360°, yang artinya gambar akan diputar secara acak dengan sudut -36° hingga +36°. Kemudian, model mempelajari pola dari data dan mengklasifikasikan menjadi dua yaitu tumor dan non tumor (Citra R et al., 2024). Pembagian data *training* 80% dan data *test* 20 % memastikan model untuk dilatih pada sebagian data tertentu dan diuji pada data yang berbeda (Maulana et al., 2024). Model dengan menerima variasi data tersebut dapat membuat model mengurangi resiko *overfitting*.

#### 4. *Evaluation*

Hasil pemodelan di tahapan ini diukur performanya. Uji performa model menggunakan tabel *confusion matrix* untuk mengevaluasi kinerja model. tabel *confusion matrix* menunjukkan perbandingan antara prediksi yang dihasilkan oleh model dengan label sebenarnya pada *dataset* (Maulana et al., 2024). Kemudian,

model dilihat performa model dengan melihat akurasi yang dihasilkan. Berikut persamaan (3.1) untuk menghitung akurasi dari model (Rasheed et al., 2023).

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3.1)$$

Persamaan (3.1) merupakan perhitungan untuk akurasi, komponen nilainya diambil dari tabel *confusion matrix* (Shanjida et al., 2022). *True positive* (TP) merupakan jumlah memprediksi positif dan prediksinya benar. *True negative* (TN) adalah jumlah memprediksi negative dan prediksinya benar negative. *False Positive* (FP) merupakan jumlah error dari model ketika memprediksi positif tetapi sebenarnya negatif. *False negative* (FN) merupakan error dari model ketika memprediksi kelas negatif, tetapi sebenarnya adalah positif.

Performa model dievaluasi berdasarkan nilai recall dari pengujian. Persamaan (3.2) digunakan untuk menghitung recall dari model (Citra R et al., 2024).

$$recall = \frac{TP}{TP+FN} \quad (3.2)$$

Persamaan (3.2) merupakan rasio antara jumlah prediksi benar yang dilakukan oleh sistem terhadap total jumlah sampel aktual dalam kategori tersebut. Recall menunjukkan seberapa baik model dalam menangkap semua kasus positif yang ada. Nilai recall yang tinggi menunjukkan bahwa model berhasil mendeteksi sebagian besar sampel positif.

Kinerja model dinilai berdasarkan nilai presisi dari pengujian. Persamaan (3.3) digunakan untuk menghitung tingkat presisi model (Fakhri et al., 2024).

$$Presisi = \frac{TP}{TP+FP} \quad (3.3)$$

Persamaan (3.3) menggambarkan rasio prediksi positif yang benar terhadap total prediksi positif yang dihasilkan oleh model. Tingginya nilai presisi menunjukkan bahwa model jarang melakukan kesalahan dalam mengklasifikasikan sampel positif.

Performa model dievaluasi berdasarkan nilai F1-score dari pengujian. Persamaan (3.4) digunakan untuk menghitung F1-score model (Hossain, Islam, Rahman, et al., 2023).

$$F1score = 2x \frac{precision \times recall}{precision + recall} \quad (3.4)$$

Persamaan (3.4) merupakan rata-rata dari presisi dan recall, yang berfungsi menyeimbangkan kedua metrik tersebut. Metrik ini sangat berguna dalam situasi dengan ketidakseimbangan kelas, karena membantu model mempertimbangkan baik presisi maupun recall secara seimbang.