

BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan hasil analisis, pengujian sistem, dan evaluasi performa model yang telah dilaksanakan untuk menjawab seluruh rumusan masalah dalam penelitian ini, maka diperoleh beberapa kesimpulan sebagai berikut:

1. Model DenseNet121 yang terintegrasi dengan *Convolutional Block Attention Module* (CBAM) berhasil digunakan untuk mengidentifikasi slice citra MRI yang mengandung dan tidak mengandung lesi stroke. Dari empat skema yang dibandingkan, yaitu M1 (*DWI All Slices*), M2 (*RGB Fusion All Slices*), M3 (*DWI Important Slices*), dan M4 (*RGB Fusion Important Slices*). M3 dengan input DWI-only dan strategi *Important Slices* dipilih sebagai model terbaik. Model M3 memperoleh Accuracy sebesar 84,65%, *Precision* sebesar 83,53%, *Recall* sebesar 88,57%, *F1-Score* sebesar 0,8598, dan AUC-ROC sebesar 0,9195. Pemilihan M3 didasarkan terutama pada *F1-Score* dan AUC-ROC yang menunjukkan kemampuan model dalam membedakan *slice* yang mengandung dan tidak mengandung lesi stroke.
2. Model NVAutoNet dari *framework* DeepISLES berhasil digunakan untuk melakukan segmentasi lesi stroke pada citra MRI. Pada 50 data pengujian ISLES 2022, diperoleh *Mean Dice Score* 0,8347, *Mean Precision* 0,8699, dan *Mean Recall* 0,8221. Sebanyak 46 dari 50 data memperoleh *Dice Score* $\geq 0,7$ dan 47 dari 50 data memperoleh *Dice Score* $\geq 0,5$. Hasil segmentasi selanjutnya digunakan untuk menghitung volume lesi berdasarkan jumlah voxel dan ukuran voxel, yang kemudian menjadi dasar dalam menentukan tingkat keparahan stroke.
3. Validasi eksternal dilakukan pada data MRI rumah sakit untuk menilai kemampuan model pada sumber data yang berbeda dari dataset

pengembangan. Validasi mencakup klasifikasi pada tingkat pasien dan segmentasi lesi.

a. Validasi Klasifikasi

Model M3 menghasilkan keputusan pasien melalui agregasi prediksi *slice*. Pengujian beberapa *threshold* probabilitas menunjukkan bahwa *threshold* 0,80 menghasilkan selisih proporsi *slice* positif terbesar, yaitu 0,3495, dengan proporsi *slice* positif sebesar 0,8333 pada pasien stroke dan 0,4839 pada pasien normal. Keputusan pasien ditentukan menggunakan *majority rule*, yaitu minimal 50% *slice* positif dikategorikan sebagai stroke. Pasien stroke kode 914320 memiliki 83,33% *slice* positif dan diklasifikasikan sebagai stroke, sedangkan pasien normal kode 678531 memiliki 48,39% *slice* positif dan diklasifikasikan sebagai normal. Kedua hasil tersebut sesuai dengan diagnosis radiologi.

b. Validasi Segmentasi

Hasil segmentasi eksternal diverifikasi secara kualitatif melalui *probability map*, citra DWI, dan lokasi anatomis pada laporan radiologi. Pada pasien 914320, probabilitas maksimum *probability map* sebesar 0,481887, sehingga *threshold* 0,4 digunakan untuk memperoleh 342 voxel kandidat lesi. Setelah *connected component analysis* dan pemeriksaan visual, *lesion mask* yang sesuai dengan area abnormal pada DWI dan deskripsi radiologi late subacute infark pada capsula eksterna kiri menghasilkan estimasi volume lesi sebesar 6,67 mL, yang termasuk kategori ringan. Validasi ini bersifat kualitatif karena tidak tersedia *ground truth mask* untuk evaluasi menggunakan Dice Score.

4. Hasil klasifikasi, segmentasi, estimasi volume lesi, dan penentuan tingkat keparahan berhasil diintegrasikan ke dalam aplikasi berbasis Streamlit. Aplikasi menghubungkan proses input citra MRI, *preprocessing*, klasifikasi

DenseNet121-CBAM, segmentasi NVAutoNet, hingga penyajian hasil estimasi volume dan tingkat keparahan dalam satu antarmuka.

5.2.Saran Pengembangan

Berdasarkan keterbatasan operasional dan ruang lingkup pelaksanaan eksperimen dalam penelitian ini, beberapa saran rekonstruktif yang dirasa paling perlu untuk diajukan bagi pengembangan studi lanjutan adalah sebagai berikut:

1. Pengayaan Ragam Dataset Klinis:

Penelitian berikutnya disarankan untuk memperluas cakupan data uji menggunakan variasi dataset multi-senter dari rumah sakit lokal untuk menguji ketahanan (*robustness*) model klasifikasi DenseNet121-CBAM dan segmentasi NVAutoNet terhadap perbedaan variasi spesifikasi mesin *scanner* MRI di lapangan.

2. Penerapan Optimasi Hiperparameter Otomatis:

Perlu dilakukan eksplorasi metode *hyperparameter tuning* otomatis tingkat lanjut, seperti *Bayesian Optimization* atau *Genetic Algorithm*, khusus pada penentuan penambatan nilai ambang batas (*threshold probability*) klasifikasi guna meminimalkan penumpukan eror *False Positive* pada potongan gambar tepi lesi.

3. Kompresi Bobot Jaringan untuk Produksi:

Untuk skenario penempatan *dashboard* Streamlit pada peladen komputasi awan komersial (*cloud production server*), disarankan untuk mengimplementasikan teknik kompresi model seperti *quantization* atau *model pruning* pada berkas bobot *ensemble network* (model0.ts hingga model14.ts) agar proses pemuatan (*loading checkpoint*) dan inferensi real-time dapat berjalan lebih efisien tanpa membutuhkan spesifikasi VRAM GPU yang terlalu besar.