

## Komparasi Performa Model 3D CNN dalam Klasifikasi Demensia Alzheimer pada MRI Otak

Auriel Azril Ardin<sup>\*1</sup>, Dyah Aruming Tyas<sup>2</sup>

<sup>1,2</sup>Departemen Ilmu Komputer dan Elektronika, FMIPA UGM, Yogyakarta, Indonesia

e-mail: <sup>\*1</sup>auriel.azril.ardin@mail.ugm.ac.id, <sup>3</sup>dyah.aruming.t@ugm.ac.id

### Abstrak

Penyakit Alzheimer adalah jenis demensia akibat kerusakan pada neuron otak yang memengaruhi memori, bahasa, dan berpikir. Diagnosis manual sering rentan terhadap subjektivitas dan memakan waktu, sehingga diperlukan model otomatis seperti 3D CNN untuk klasifikasi tingkat keparahan Alzheimer. Namun, kompleksitas arsitektur 3D CNN menyebabkan waktu komputasi yang tinggi. Penelitian ini membandingkan tiga arsitektur model 3D CNN yaitu 3D ResNet, 3D ResNeXt + Bi-LSTM, dan 3D CNN + CLSTM untuk menentukan model yang optimal. Dataset yang digunakan diperoleh dari database ADNI. Performa model dievaluasi dengan menggunakan confusion matrix, akurasi, presisi, recall, F1-score dan waktu komputasi.

Hasil penelitian menunjukkan bahwa 3D ResNet memiliki akurasi pelatihan tertinggi mencapai 99,54% dan waktu komputasi pelatihan sebesar 57,61 detik/epoch. Model 3D ResNeXt + Bi-LSTM mencapai akurasi pengujian sebesar 99,33% dan waktu inferensi tercepat yaitu 0,0182 detik/sampel, namun waktu komputasi pelatihan terlama yaitu 117,68 detik/epoch. Sementara itu, 3D CNN + CLSTM mencapai akurasi uji sempurna 100% tetapi memiliki waktu inferensi terlama yaitu 0,0268 detik/sampel. Penelitian ini menunjukkan bahwa arsitektur yang sederhana tetap dapat memberikan performa optimal dengan waktu komputasi yang lebih efisien dibandingkan model yang lebih kompleks.

**Kata kunci**— Alzheimer, Machine Learning, 3D Convolutional Neural Network, Klasifikasi, SDGs.

### Abstract

Alzheimer's disease is a type of dementia caused by damage to brain neurons, affecting memory, language, and thinking. Manual diagnosis is often subjective and time-consuming, necessitating the development of automated models like 3D CNNs to classify Alzheimer's severity levels. However, the complexity of 3D CNN architectures leads to high computational time. This study compares three 3D CNN model architectures—3D ResNet, 3D ResNeXt + Bi-LSTM, and 3D CNN + CLSTM—to identify the optimal model. The dataset used was obtained from the ADNI database, and the models' performance was evaluated based on a confusion matrix, accuracy, precision, recall, F1-score, and computational time.

The results showed that 3D ResNet achieved the highest training accuracy of 99.54% with a computational time of 57.61 seconds per epoch during training. The 3D ResNeXt + Bi-LSTM model attained a testing accuracy of 99.33% and the fastest inference time of 0.0182 seconds per sample, but it required the longest training time of 117.68 seconds per epoch. Meanwhile, 3D CNN + CLSTM achieved a perfect testing accuracy of 100%, but with the longest inference time of 0.0268 seconds per sample. This study highlights that simpler architectures can still deliver optimal performance with more efficient computational times compared to more complex models.

**Keywords**— Alzheimer's Disease, Machine Learning, 3D Convolutional Neural Network, Classification, SDGs.

## 1. PENDAHULUAN

Alzheimer merupakan salah satu jenis demensia yang paling umum. Penyakit ini disebabkan oleh kerusakan sel saraf (neuron) di otak. Neuron yang pertama kali rusak yaitu neuron pada bagian otak yang bertanggung jawab pada fungsi memori, bahasa, dan berpikir [1]. Saat ini terdapat lebih dari 55 juta orang terkena demensia dan hampir 10 juta kasus baru didiagnosis setiap tahunnya. Hal ini mengindikasikan pentingnya upaya pencegahan sedini mungkin untuk mengatasi masalah ini [2].

Diagnosis alzheimer secara manual oleh tenaga medis pada citra *Magnetic Resonance Imaging* (MRI) melibatkan proses visualisasi dengan memeriksa pada area otak yang berpotensi alzheimer dan mencari tanda-tanda atau pola karakteristik yang mengindikasikan alzheimer. Diagnosis manual sangat bergantung pada pengalaman tenaga medis, sehingga rentan terhadap subjektivitas dan human error. Prosesnya juga memakan waktu lama dan sulit jika menghadapi data besar [3]. Oleh sebab itu, perlu adanya sistem otomatis yang dapat mengklasifikasikan alzheimer sesuai dengan tingkat keparahannya.

Pada penelitian sebelumnya mengenai klasifikasi alzheimer menunjukkan bahwa 2D CNN kurang efektif dalam menganalisis citra 3D MRI karena memecah volume 3D menjadi irisan 2D, sehingga berisiko kehilangan informasi spasial penting [4]. Hal ini membatasi deteksi perubahan struktural otak dan pengklasifikasian alzheimer berdasarkan tingkat keparahannya. Akibatnya, perubahan struktural di seluruh volume otak sulit untuk dideteksi. Oleh karena itu, pada penelitian ini digunakan 3D CNN untuk mengklasifikasikan alzheimer. 3D CNN mampu untuk menangkap informasi spasial dan temporal yang lebih baik daripada 2D CNN.

Pada penelitian sebelumnya mengenai klasifikasi alzheimer menggunakan citra MRI menunjukkan bahwa model berbasis 3D CNN mampu untuk menangkap pola dinamis dari aktivitas neural dan mencapai akurasi sebesar 96,77% pada klasifikasi AD vs CN [5]. Sementara itu, Li *et al.* [6] mengintegrasikan *non-local blocks* dan *attention mechanisms* ke dalam 3D ResNet dan menghasilkan akurasi sebesar 93,65%. Model ResNeXt + Bi-LSTM yang dikembangkan Wang *et al.* [7] mencapai akurasi tertinggi 98,97% dengan memanfaatkan *residual connection* dan hubungan spasial antar fitur.

Sedangkan, penelitian yang dilakukan oleh Xia *et al.* [8] mengkombinasikan 3D CNN dan 3D CLSTM untuk menangkap informasi spasial dan interaksi temporal dengan akurasi sebesar 94,19%. Kumari *et al.* [9] menggunakan PB-CNN berbasis *patch* yang dioptimalkan dengan OBL-WOA dan mencapai akurasi sebesar 92,86%. Wu *et al.* [10] mengembangkan model AMSNet yang memanfaatkan mekanisme *multi-scale attention* untuk meningkatkan efisiensi ekstraksi fitur dengan hasil akurasi mencapai 91,3%. Kumari *et al.* [11] mengaplikasikan algoritma MOSA pada model 3D CNN untuk optimasi klasifikasi tiga kelas alzheimer (AD, MCI, CN) dan menghasilkan akurasi sebesar 92,86%. Selain itu, Dwivedi *et al.* [4] menggabungkan pendekatan 3D CNN dan SVM dan mampu mencapai akurasi 91,85%.




Akan tetapi, model 3D CNN memiliki arsitektur yang sangat kompleks yang berakibat pada waktu komputasi yang lama. Arsitektur yang rumit ini, disebabkan oleh banyaknya lapisan konvolusi dan pooling yang harus diproses, serta peningkatan jumlah parameter yang harus dioptimalkan selama proses pelatihan [12]. Hal ini berujung pada waktu komputasi yang signifikan, baik dalam pelatihan maupun inferensi. Oleh sebab itu, penelitian ini bertujuan untuk membandingkan performa model 3D CNN dalam menangani data 3D MRI untuk membuktikan bahwa arsitektur model 3D CNN yang sederhana tetap dapat mempertahankan performa model yang optimal. Penelitian ini juga mendukung Tujuan Pembangunan Berkelanjutan (*Sustainable Development Goals / SDGs*) yaitu kesehatan yang baik dan kesejahteraan (*Good Health and well-being*).

## 2. METODE PENELITIAN

### 2.1 Pengumpulan Data

Data yang digunakan berasal dari database ADNI (*Alzheimer's Disease Neuroimaging Initiative*). ADNI adalah lembaga yang berfokus untuk menangani alzheimer. Dataset dikategorikan menjadi tiga kelas meliputi *Cognitive Normal* (CN), *Mild Cognitive Impairment* (MCI), dan *Alzheimer's Disease* (AD) seperti yang ditunjukkan pada Tabel 1.. Dalam penelitian ini, dataset dibagi menjadi 2 yaitu set pelatihan dan set pengujian. Set pelatihan merupakan subset dataset yang digunakan untuk melatih model CNN. Data pada set pelatihan terbagi menjadi dua dataset yaitu data latih dan data validasi. Set pengujian merupakan subset dataset yang digunakan untuk menguji performa akhir model. Setelah model melalui proses pelatihan, set pengujian digunakan untuk mengukur performa model secara objektif.

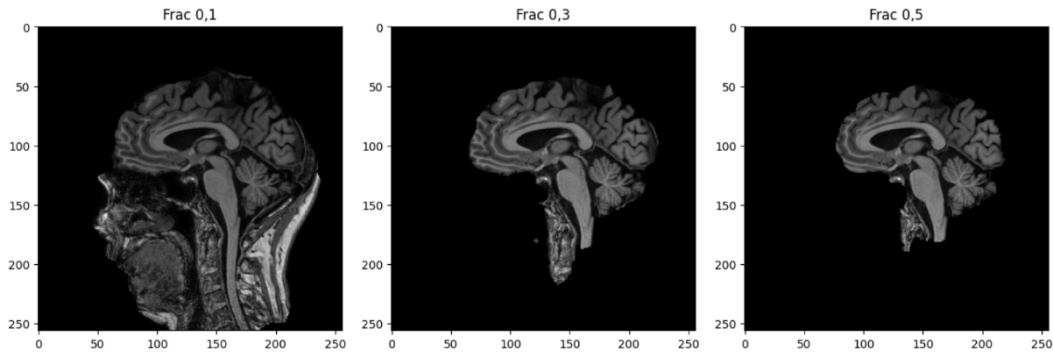
Tabel 1 Citra MRI pada setiap kelas

Kelas	Citra MRI
<i>Alzheimer's Disease</i>	
<i>Mild Cognitive Impairment</i>	
<i>Cognitive Normal</i>	

### 2.2 Preprocessing

#### 2.2.1 Skull Stripping

Citra MRI mentah mengandung noise dan informasi tidak relevan, seperti tulang tengkorak dan jaringan lunak yang dapat mengurangi akurasi analisis. Oleh sebab itu, diperlukan teknik untuk memisahkan jaringan otak dari struktur lain. Teknik tersebut yaitu *Brain Extraction Tools* (BET) pada *software* FSL (*FMRIB Software Library*) yang secara otomatis dapat mendeteksi batas otak menggunakan informasi spasial dan intensitas *voxel*. Dalam penelitian ini, digunakan parameter *-g* (*vertical gradient*) untuk menyesuaikan posisi citra dan *-f* (*fractional intensity threshold*) sebesar 0,5 untuk menentukan ambang batas segmentasi. Nilai ini memastikan bahwa area otak yang diekstraksi memiliki intensitas setidaknya 50% lebih tinggi dari rata-rata intensitas tengkorak. Hasil dari BET dapat dilihat pada Gambar 1.



Gambar 1 Hasil BET pada berbagai nilai *fractional*

### 2.2.2 Normalisasi

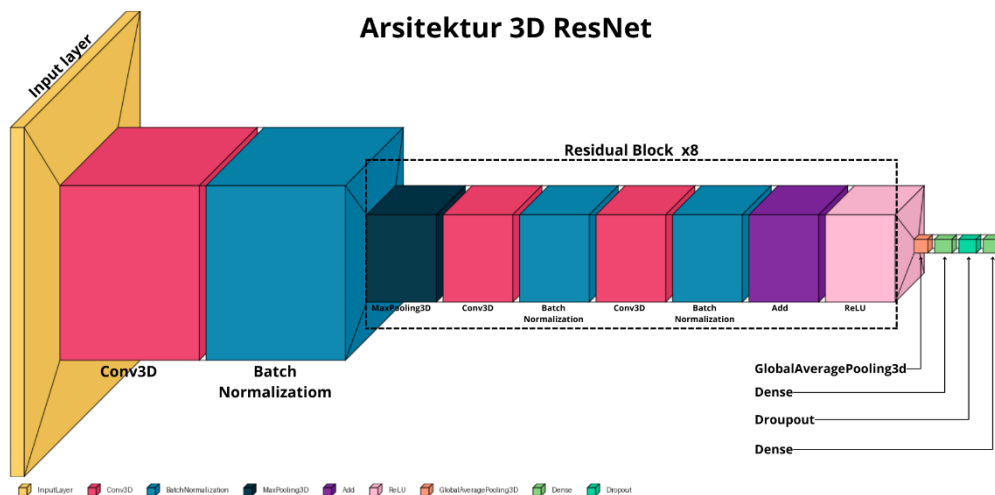
Setelah skull stripping, tahap selanjutnya adalah normalisasi citra MRI untuk memastikan intensitas voxel yang seragam dan dimensi yang konsisten. Rentang intensitas yang digunakan adalah antara -1000 dan 400, di mana nilai di luar rentang ini dianggap tidak relevan atau merupakan artefak. Nilai-nilai tersebut kemudian dinormalisasi ke rentang 0 hingga 1 untuk mempercepat konvergensi model dan membuat pelatihan lebih stabil dan efisien.

### 2.2.3 Augmentasi

Augmentasi data dilakukan untuk meningkatkan variasi data latih dan mencegah *overfitting*. Salah satu teknik yang digunakan adalah *resize volume*. Teknik ini mengubah dimensi citra MRI menjadi 64x128x128 menggunakan fungsi *ndimage.zoom*. Selanjutnya, citra dirotasi secara acak pada rentang  $-20^\circ$  hingga  $20^\circ$  dengan menggunakan fungsi *ndimage.rotate*. Proses rotasi hanya dilakukan pada data latih, sementara data validasi dan data uji hanya melalui *resize volume* agar tetap representatif.

## 2.3 Arsitektur CNN

### 2.3.1 3D ResNet

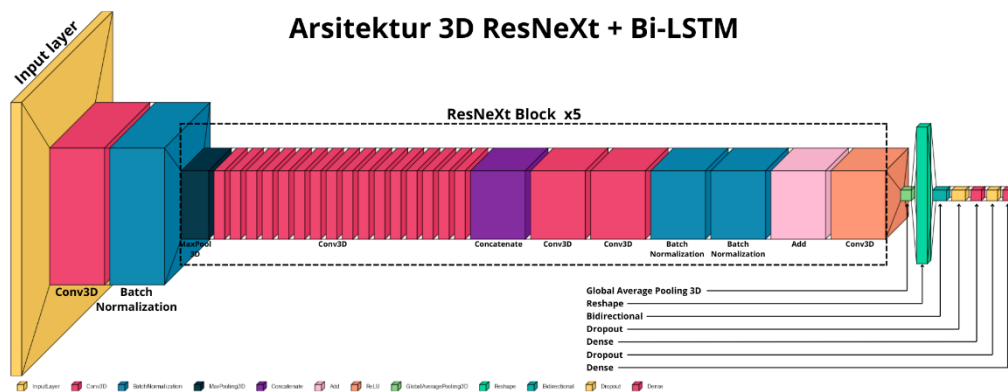


Gambar 2 Arsitektur 3D ResNet

Arsitektur 3D ResNet diawali dengan *input layer* yang menerima citra MRI berukuran 128x128x64x1. Proses awal melibatkan konvolusi 3D dengan 64 *filter* berukuran 7x7x7, diikuti oleh *batch normalization* dan *max pooling* untuk mengurangi dimensi spasial. Bagian utama

arsitektur ini terdiri dari *residual blocks* yang menggunakan *3D convolution* berukuran  $3 \times 3 \times 3$  dan *batch normalization*. Dengan adanya *residual* dan *shortcut connections*, arsitektur ini dapat mengurangi risiko *vanishing gradient* dan mempercepat konvergensi. Terdapat empat tahap *residual block* dengan *filter* yang bertambah secara bertahap (64, 128, 256, 512) dan *strides* sebesar 2. Setelah itu, *3D average pooling* digunakan untuk menghasilkan vektor fitur. Pada bagian akhir arsitektur, terdapat *fully connected layer* dengan 256 unit, diikuti *dropout* 0,5 dan *softmax activation* untuk klasifikasi. Dengan total 33.3 juta parameter, arsitektur ini memiliki kapasitas besar untuk mengekstraksi fitur volumetrik yang kompleks sehingga menjadikannya cocok untuk tugas klasifikasi citra 3D MRI yang membutuhkan analisis spasial mendalam.

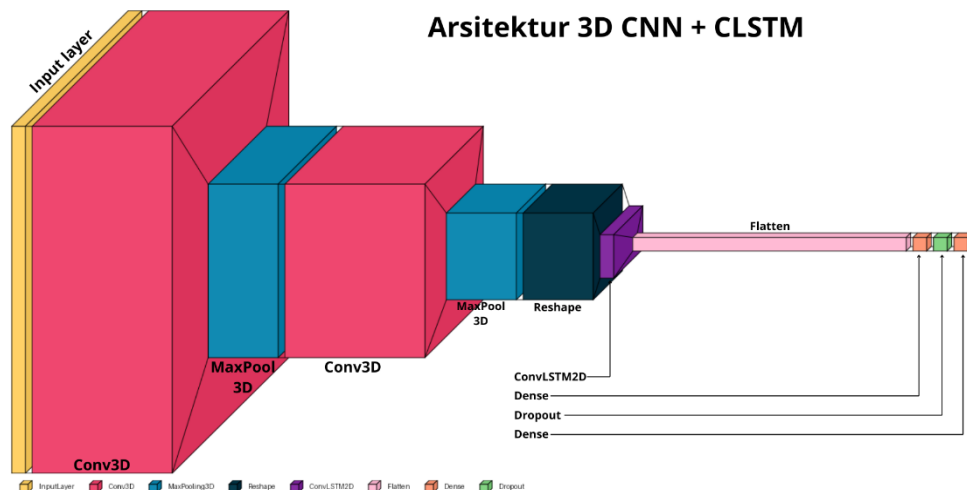
### 2.3.2 3D ResNeXt + Bi-LSTM



Gambar 3 Arsitektur 3D ResNeXt +Bi-LSTM

Arsitektur 3D ResNeXt + Bi-LSTM diawali dengan *input layer* berukuran  $128 \times 128 \times 64 \times 1$  yang dapat menerima data volumetrik. Pada tahap awal, terdapat *3D convolution* layer dengan 32 *filter* ( $7 \times 7 \times 7$ ) dan *strides* 2. Untuk ekstraksi fitur awal dan pengurangan resolusi spasial, digunakan *batch normalization* dan *3D maxpooling*. Bagian utama arsitektur ini menggunakan lima *ResNeXt blocks* dengan *filter* bertambah dari 64 hingga 512, beberapa di antaranya menggunakan *strides* sebesar 2 untuk *downsampling*. Fitur yang diekstraksi diratakan dengan *3D global average pooling* dan diproses melalui Bi-LSTM untuk menangkap pola temporal. Di akhir arsitektur, terdapat *dense layer* dengan 128 unit dan *softmax activation* untuk klasifikasi multikelas. Dengan total 2,6 juta parameter, arsitektur ini mampu menangani masalah klasifikasi yang membutuhkan analisis mendalam pada dimensi spasial dan temporal.

### 2.3.3 3D CNN + CLSTM



Gambar 4 Arsitektur 3D CNN + CLSTM

Arsitektur 3D CNN + CLSTM diawali dengan *input layer* berukuran 128x128x64x1. Tahap awal melibatkan *3D convolutional layer* dengan 32 *filter* (3x3x3), *ReLU activation*, dan *padding "same"*, serta *3D MaxPooling* untuk mengurangi dimensi spasial. Proses ini diulang pada lapisan kedua dengan 64 *filter* untuk menangkap fitur yang lebih kompleks. Setelah proses konvolusi dan *pooling*, *output* dari lapisan konvolusi diubah bentuknya untuk disesuaikan dengan masukan pada CLSTM. Lapisan CLSTM menggunakan 64 *filter* (3x3) dan *padding "same"* untuk mempertahankan dimensi output. Data diratakan menggunakan *flatten* dan diproses pada *fully connected layer* untuk klasifikasi. Pada akhir arsitektur, terdapat *dense layer* dengan 128 *neuron* dan *ReLU activation*, diikuti *dropout 0,5*, dan *softmax activation* untuk *output* klasifikasi. Dengan total 4.5 juta parameter, arsitektur ini dapat memberikan representasi fitur yang mendalam dan dapat menangkap hubungan temporal dengan baik.

#### 2.4 Pelatihan Model

Selama proses pelatihan model, data latih diacak dengan menggunakan fungsi *.shuffle()* untuk mencegah *overfitting*. Kemudian, data dibagi ke dalam *batch* kecil dengan variasi 2, 4, dan 8. Teknik *prefetching* digunakan untuk memuat batch berikutnya secara paralel. Pelatihan dilakukan dengan metode *fit()* hingga 150 *epoch*, disertai dengan *callback* seperti *EarlyStopping*, *ModelCheckpoint*, dan *StopOnPerfectAccuracy* untuk menghemat waktu dan sumber daya komputasi. *Adam optimizer* diterapkan dengan *learning rate* dinamis menggunakan *ExponentialDecay* dengan variasi *learning rate* awal sebesar 0,001, 0,0001, 0,00001 dan 0,000001. Fungsi *loss* yang digunakan adalah *sparse\_categorical\_crossentropy*. Mode eksekusi *eager* diaktifkan untuk memudahkan *debugging*. Durasi pelatihan dicatat untuk analisis waktu komputasi model selama pelatihan.

#### 2.5 Evaluasi Model

Setelah model melalui proses pelatihan, evaluasi terhadap model dilakukan untuk menilai kinerja model dalam memprediksi data yang belum pernah digunakan sebelumnya. Dalam evaluasi ini, beberapa metrik seperti akurasi, presisi, *recall*, *F1 score* dan *confusion matrix* digunakan untuk memberikan gambaran yang komprehensif mengenai kinerja model.

### 3. HASIL DAN PEMBAHASAN

Pengujian model 3D ResNet, 3D ResNeXt + Bi-LSTM, dan 3D CNN + CLSTM dilakukan dengan menggunakan konfigurasi *learning rate*, *batch size*, dan *split ratio* yang telah ditentukan. Setiap model memiliki kelebihan dan kekurangan, baik dari segi pelatihan, pengujian, metrik evaluasi yang didapatkan, hingga efisiensi komputasi model.

Tabel 2 Perbandingan hasil penelitian 3D ResNet, 3D ResNeXt + Bi-LSTM, dan 3D CNN + CLSTM

Model	Train						Test					Ket
	Acc	Loss	Val. Acc	Val. Loss	Epoch	Waktu Komp. (s/epoch)	Acc	Prec	Rec	F1	Waktu Inf. (s/sampel)	
3D ResNet	99,54%	0,034	95,83%	0,206	142 (ES)	57,61	98,67%	98,67%	98,67%	98,67%	0,0196	Learning rate = 0,00001 Batch size = 2 Split ratio = 0,90
3D ResNeXt + Bi-LSTM	95,83%	0,158	91,67%	0,565	126 (ES)	117,68	99,33%	99,33%	99,33%	99,33%	0,0182	Learning rate = 0,0001 Batch size = 2 Split ratio = 0,90
3D CNN + CLSTM	90,28%	0,263	91,67%	0,380	125 (ES)	32,72	100%	100%	100%	100%	0,0268	Learning rate = 0,0001 Batch size = 8 Split ratio = 0,90

Pengujian model 3D ResNet menunjukkan performa yang sangat baik selama proses pelatihan. Dengan menggunakan *learning rate* sebesar 0,00001, *batch size* 2, dan *split ratio* 90:10.

Berdasarkan grafik akurasi pelatihan pada Gambar 5 (a) menunjukkan bahwa akurasi pelatihan mengalami peningkatan yang stabil hingga mendekati 100% setelah sekitar 60 *epoch* dengan akurasi akhir mencapai 99,54%. Sedangkan, pada akurasi validasi mengalami peningkatan tetapi terdapat fluktuasi yang cukup besar hingga mencapai akurasi validasi sebesar 95,83%.

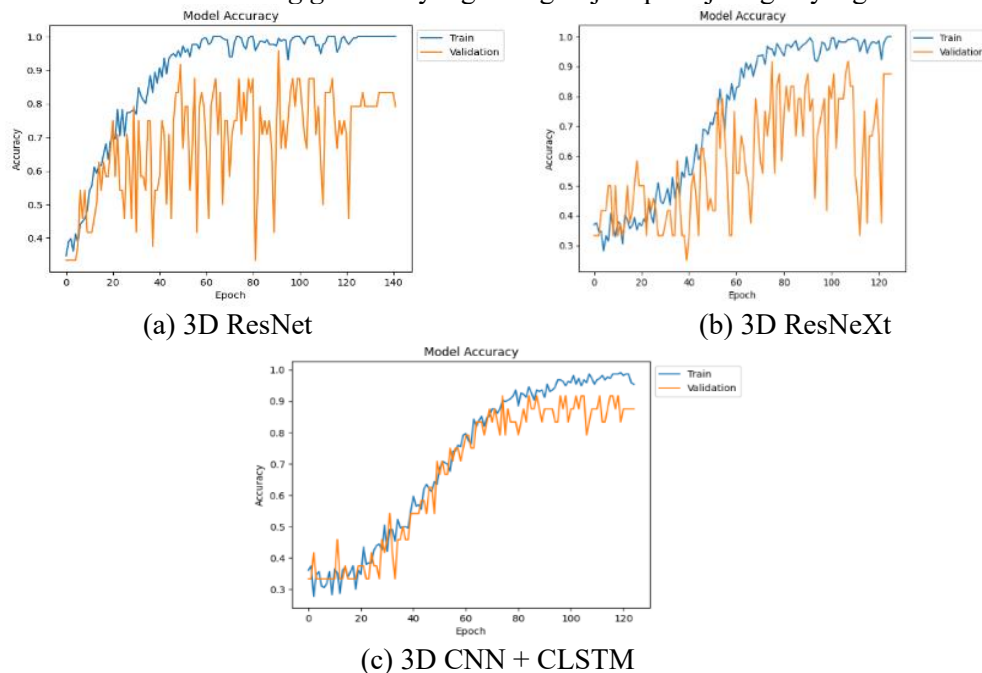
Sementara itu, Grafik *loss* pada Gambar 6 (a) menunjukkan bahwa *loss* pada data latih mengalami penurunan yang konsisten hingga mencapai nilai minimum sebesar 0,034 pada epoch terakhir. Sedangkan, *loss* pada data validasi mengalami fluktuasi yang signifikan dan relatif lebih tinggi dibandingkan *loss* pada data latih. Tren tersebut mengindikasikan bahwa model mengalami *overfitting* selama proses pelatihan.

Berdasarkan *confusion matrix* pada Gambar 7 (a) menunjukkan bahwa model berhasil memprediksi seluruh sampel pada kelas AD dan CN dengan benar. Sedangkan, pada kelas MCI, model berhasil mengklasifikasikan dengan benar sebanyak 48 dari 50 sampel data uji.

Secara umum, pengujian model dengan menggunakan data uji yang belum pernah digunakan sebelumnya menunjukkan hasil akurasi, presisi, *recall*, dan *F1-score* yang sangat baik yaitu masing-masing sebesar 98,67%. Hasil tersebut menandakan bahwa model ini mampu mempelajari pola data dengan baik sehingga dapat menggeneralisasi data baru dengan akurasi yang sangat baik.

Model 3D ResNet memiliki waktu komputasi sebesar 57,61 detik/*epoch*, yang tergolong sedang dibandingkan model lainnya. Waktu inferensi model ini tercatat sebesar 0,01962 detik/sampel. Meskipun waktu komputasi tidak secepat 3D CNN + CLSTM, waktu inferensi yang relatif rendah menunjukkan efisiensi dalam memproses data setelah pelatihan selesai.

Keunggulan utama dari model 3D ResNet yaitu kemampuannya dalam mempelajari fitur-fitur yang kompleks dalam citra MRI. Dengan adanya *residual connections*, model ini dapat menghindari masalah *vanishing gradient* yang sering terjadi pada jaringan yang dalam.



Gambar 5 Grafik akurasi pelatihan

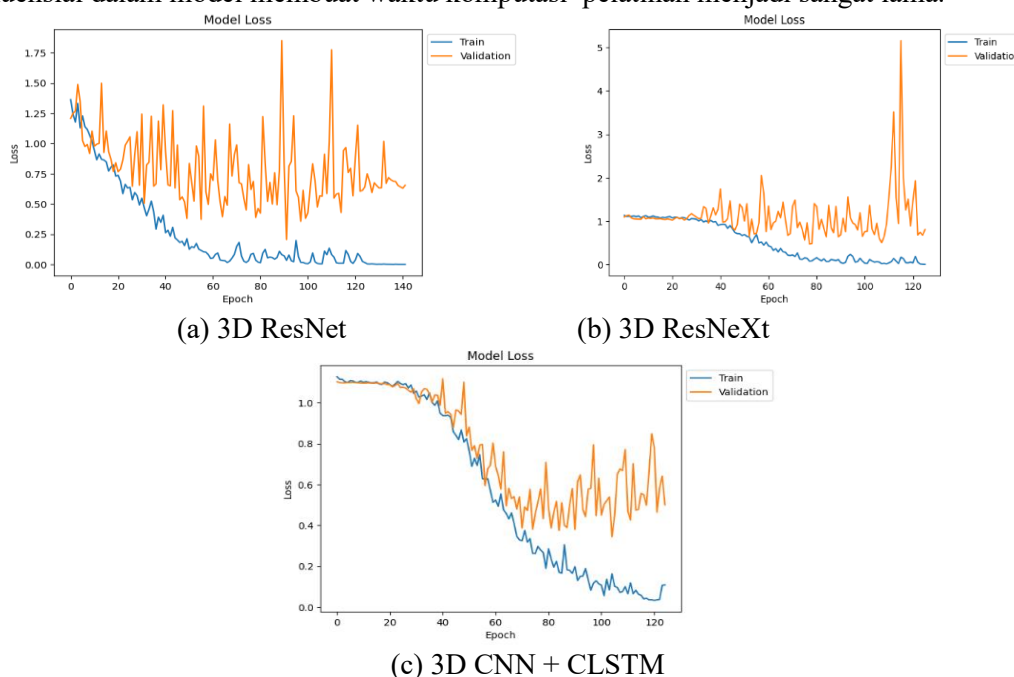
Pengujian model 3D ResNeXt + Bi-LSTM menggunakan *learning rate* sebesar 0,0001, *batch size* 2, dan *split ratio* 90:10, Sebagaimana yang terdapat pada Gambar 5 (b) akurasi pelatihan meningkat secara stabil hingga mendekati nilai yang sempurna yaitu 95,83%. Sedangkan, akurasi validasi mengalami fluktuasi yang signifikan dan diakhiri dengan akurasi. Sementara itu, grafik *loss* pada Gambar 6 (b) menunjukkan terlihat pola penurunan *loss* yang konsisten hingga mencapai 0,158. Sedangkan, *loss* pada data validasi mengalami fluktuasi yang



cukup tajam. Tren tersebut mengindikasikan bahwa model mengalami *overfitting* selama proses pelatihan.

Kemudian, model 3D ResNeXt + Bi-LSTM menunjukkan waktu komputasi paling tinggi sebesar 117,68 detik/*epoch*. Hal ini kemungkinan disebabkan oleh kompleksitas tambahan dari arsitektur Bi-LSTM yang memerlukan komputasi lebih intensif. Namun, model ini memiliki waktu inferensi yang paling cepat, yaitu 0,01822 detik per sampel. Hal ini menandakan bahwa meskipun proses pelatihan memakan waktu lama, proses inferensi setelah model dilatih dapat dilakukan dengan cepat.

Performa model saat diuji dengan menggunakan data yang belum pernah digunakan sebelumnya menunjukkan hasil yang sangat baik dengan akurasi, presisi, *recall*, dan *F1-score* sebesar 99,33%. Hasil tersebut membuktikan keunggulan kombinasi 3D ResNeXt dan Bi-LSTM dalam menangkap pola spasial dan temporal dalam citra MRI. 3D ResNeXt memiliki struktur *grouped convolution* yang efisien untuk ekstraksi fitur dan Bi-LSTM, sehingga mampu mempelajari hubungan temporal dengan baik. Akan tetapi, kompleksitas arsitektur dan proses sekuensial dalam model membuat waktu komputasi pelatihan menjadi sangat lama.



Gambar 6 Grafik *loss* pelatihan model

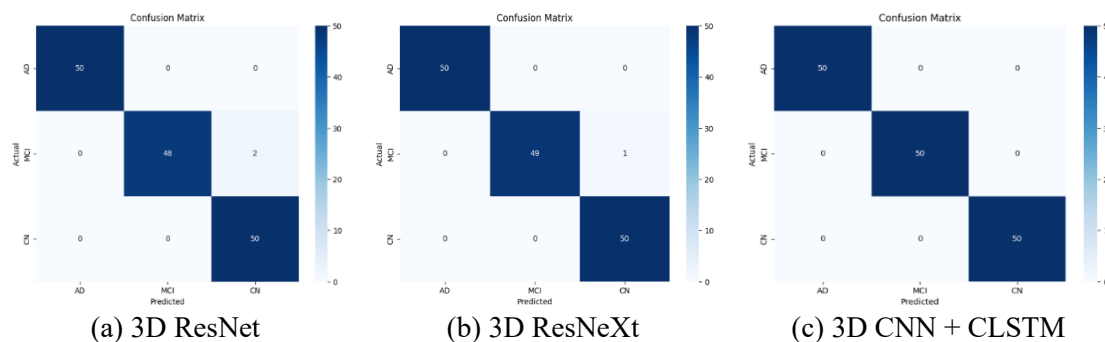
Pengujian model 3D CNN + CLSTM menggunakan *learning rate* sebesar 0,0001, batch size 8, dan split ratio 90:10. Berdasarkan grafik akurasi pelatihan model 3D CNN + CLSTM seperti yang terdapat pada Gambar 5 (c), performa model pada tahap pelatihan mengalami peningkatan yang stabil seiring bertambahnya jumlah *epoch* dengan akurasi pelatihan sebesar 90,28. Begitu pula dengan akurasi validasi yang mengalami peningkatan dan mulai stagnan sekitar 75 *epoch* dengan mencapai akurasi sebesar 91,67%. Tren ini menunjukkan bahwa model dapat mempelajari pola data dengan baik.

Kemudian, pada grafik *loss* model 3D CN + CLSTM sebagaimana yang ditunjukkan pada Gambar 6 (c), *loss* pada data latih mengalami penurunan yang konsisten hingga mencapai 0,263. Sedangkan, *loss* pada data validasi mengalami penurunan dengan fluktuasi yang lebih signifikan dibandingkan *loss* pada data latih hingga mencapai 0,3804. Tren grafik *loss* tersebut menunjukkan bahwa model berhasil meminimalkan kesalahan prediksi pada data latih.

Di sisi lain, model 3D CNN + CLSTM menunjukkan efisiensi komputasi tertinggi dengan waktu komputasi hanya 32,72 detik/*epoch*. Waktu tersebut adalah yang tercepat dibandingkan dua model lainnya. Akan tetapi, waktu inferensi model ini lebih lama, yakni 0,02684 detik/sampel. Meskipun model ini unggul dalam kecepatan pelatihan, waktu inferensi yang



sedikit lebih lama dapat menjadi pertimbangan dalam aplikasi *real-time* yang membutuhkan kecepatan tinggi saat melakukan prediksi.



Gambar 7 Confusion matrix

Hasil pengujian model 3D CNN + CLSTM menunjukkan bahwa model memiliki tingkat akurasi yang sempurna untuk mengklasifikasikan alzheimer sesuai dengan tingkat keparahannya sebagaimana yang terdapat pada Gambar 7 (c). Masing-masing kelas memiliki 50 sampel dan model mampu mengklasifikasikan seluruh sampel dengan benar dan tidak ada kesalahan klasifikasi yang ditunjukkan oleh nilai *True Positive* (TP) sebanyak 50 dan *False Negative* (FN) sebanyak 0. Dengan demikian, model mampu mencapai akurasi, presisi, *recall*, dan *F1-score* yang sempurna masing-masing sebesar 100% untuk setiap kelasnya.

Hasil tersebut mencerminkan efektivitas kombinasi 3D CNN dan CLSTM dalam mempelajari pola data citra MRI. 3D CNN berperan dalam mempelajari pola spasial yang kompleks dalam citra 3D MRI. Sedangkan, CLSTM berperan dalam menangkap informasi temporal maupun sekuensial dalam citra 3D MRI. Selain itu, pemilihan parameter yang optimal juga berperan penting dalam pelatihan model sehingga memberikan hasil yang sangat baik.

#### 4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa model 3D ResNet memiliki akurasi pelatihan tertinggi sebesar 99,54% dan performa uji sebesar 98,67% untuk semua *performance metrics*. Model ini juga memiliki waktu komputasi pelatihan yang lebih efisien dibandingkan dengan model 3D ResNeXt + Bi-LSTM yaitu sebesar 57,61 detik/*epoch*, meskipun sedikit lebih lambat pada waktu inferensi yaitu sebesar 0,0196 detik/sampel. Sementara itu, model 3D ResNeXt + Bi-LSTM menunjukkan akurasi uji sebesar 99,33% dan waktu inferensi tercepat yaitu 0,0182 detik/sampel, meskipun membutuhkan waktu komputasi tertinggi mencapai 117,68 detik/*epoch* akibat kompleksitas arsitekturnya. Selanjutnya, model 3D CNN + CLSTM memiliki performa terbaik dibanding model lainnya dengan nilai akurasi, presisi, *recall*, dan *F1-score* yang sempurna yaitu 100%. Akan tetapi, model ini membutuhkan waktu inferensi paling lama yaitu 0,0268 detik/sampel. Setiap model memiliki parameter yang berbeda untuk mencapai performa optimalnya dan semakin kompleks arsitektur model maka waktu komputasi pelatihan cenderung menjadi semakin lama, meskipun hal ini tidak selalu berlaku untuk waktu inferensi.

#### 5. SARAN

Untuk penelitian selanjutnya dapat dilakukan pengembangan dengan menggunakan teknik seperti *mixed precision training*, *model pruning*, dan *quantization* untuk menyederhanakan model dengan tujuan mengoptimalkan waktu komputasi. Teknik-teknik ini dapat mempercepat waktu pelatihan dan inferensi serta mengurangi biaya komputasi.

## DAFTAR PUSTAKA

- [1] Alzheimer's Association., "2023 Alzheimer's disease facts and figures," *Alzheimer's & Dementia*, vol. 19, no. 4, pp. 1598–1695, Apr. 2023, doi: 10.1002/alz.13016.
- [2] World Health Organization, "Dementia," 2023. Accessed: Oct. 18, 2023. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/dementia>
- [3] K. A. N. N. P. Gunawardena, R. N. Rajapakse, and N. D. Kodikara, "Applying convolutional neural networks for pre-detection of alzheimer's disease from structural MRI data," in *2017 24th International Conference on Mechatronics and Machine Vision in Practice (M2VIP)*, Auckland: IEEE, Nov. 2017, pp. 1–7. doi: 10.1109/M2VIP.2017.8211486.
- [4] S. Dwivedi, T. Goel, R. Sharma, and R. Murugan, "Structural MRI based Alzheimer's Disease prognosis using 3D Convolutional Neural Network and Support Vector Machine," in *2021 Advanced Communication Technologies and Signal Processing (ACTS)*, Rourkela, India: IEEE, Dec. 2021, pp. 1–4. doi: 10.1109/ACTS53447.2021.9708107.
- [5] R. S. Swarun, V. Navya, B. M C, and V. Thomas, "A Deep Learning Based Classification of Alzheimer's Disease Using 3D CNN," in *2024 1st International Conference on Trends in Engineering Systems and Technologies (ICTEST)*, Kochi, India: IEEE, Apr. 2024, pp. 1–5. doi: 10.1109/ICTEST60614.2024.10576122.
- [6] Y. Li, Y. Zhang, J. Wu, X. Zhang, L. Han, and X. Cui, "Multi-Attention-based Global 3D ResNet for Alzheimer's Disease Diagnosis," in *2024 International Joint Conference on Neural Networks (IJCNN)*, Yokohama, Japan: IEEE, Jun. 2024, pp. 1–8. doi: 10.1109/IJCNN60899.2024.10650984.
- [7] X. Wang, J. Yi, and Y. Li, "Application of Fusion Model of 3D-ResNeXt and Bi-LSTM Network in Alzheimer's Disease Classification," in *2022 6th International Conference on Communication and Information Systems (ICCIS)*, Chongqing, China: IEEE, Oct. 2022, pp. 136–140. doi: 10.1109/ICCIS56375.2022.9998141.
- [8] Z. Xia *et al.*, "A Novel End-to-End Hybrid Network for Alzheimer's Disease Detection Using 3D CNN and 3D CLSTM," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, Iowa City, IA, USA: IEEE, Apr. 2020, pp. 1–4. doi: 10.1109/ISBI45749.2020.9098621.
- [9] R. Kumari, S. Goel, and S. Das, "A Patch Based 3D CNN Approach for Diagnosing Early Stages of Alzheimer's Disease by Applying OBL-WOA Algorithm," in *2023 8th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA)*, Chengdu, China: IEEE, Apr. 2023, pp. 455–459. doi: 10.1109/ICCCBDA56900.2023.10154706.
- [10] Y. Wu, Y. Zhou, W. Zeng, Q. Qian, and M. Song, "An Attention-Based 3D CNN With Multi-Scale Integration Block for Alzheimer's Disease Classification," *IEEE J. Biomed. Health Inform.*, vol. 26, no. 11, pp. 5665–5673, Nov. 2022, doi: 10.1109/JBHI.2022.3197331.
- [11] R. Kumari, S. Goel, and S. Das, "A 3D Convolutional Neural Network Approach for Diagnosing Alzheimer's Disease using Modified Owl Search Optimization Technique," in *TENCON 2022 - 2022 IEEE Region 10 Conference (TENCON)*, Hong Kong, Hong Kong: IEEE, Nov. 2022, pp. 1–7. doi: 10.1109/TENCON55691.2022.9977604.
- [12] N. Noshiri, M. A. Beck, C. P. Bidinosti, and C. J. Henry, "A comprehensive review of 3D convolutional neural network-based classification techniques of diseased and defective crops using non-UAV-based hyperspectral images," *Smart Agricultural Technology*, vol. 5, p. 100316, Oct. 2023, doi: 10.1016/j.atech.2023.100316.