

Klasifikasi Penyakit Alzheimer Menggunakan Metode *Hybrid Inception v3-RELM*

Annisa Widya Dwirani, Nurissaidah Ulinnuha

Matematika, Fakultas Sains dan Teknologi, UIN Sunan Ampel Surabaya

Email: annisawidya38@gmail.com, nuris.ulinnuha@uinsa.ac.id

Diajukan: 21-11-2025

Direvisi: 13-03-2026

Disetujui: 07-09-2026

Diterbitkan: 30-09-2026

Abstract. *Alzheimer's is the common cause of dementia and is characterized by damage to nerve cells (neurons) in areas of the brain that are important for cognitive function. The manual diagnostic process for small changes in brain biomarkers using Magnetic Resonance Imaging (MRI) often faces limitations in terms of accuracy and consistency. To address this, the Computer-Aided Diagnosis (CAD) approach was developed to assist in the automatic detection and classification process. This study uses data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and applies the hybrid Inception v3 method as a feature extractor and Regularized Extreme Learning Machine (RELM) as a classifier for the extracted features. The purpose of this research is to determine the best hyperparameter combination thru a tuning process to produce an optimal classification model for Alzheimer's patients MRI images. This model is designed to classify six levels of cognitive impairment severity: normal cognitive function, subjective memory complaints, early mild cognitive impairments, mild cognitive impairment, late mild cognitive impairment, and Alzheimer's disease. Based on the best results, the model with the best parameters was obtained at a value of $C=0.01$, the Swish activation function, 32768 neurons, and a batch size of 2048, with an accuracy of 81.67%, precision of 82.23%, recall of 81.78%, specificity of 96.23%, and an f1-score of 81.31%.*

Keywords: *Alzheimer's Disease, CNN, Inception v3, MRI, RELM*

Abstrak. *Alzheimer merupakan penyebab demensia yang paling umum dan ditandai oleh kerusakan sel saraf (neuron) pada area otak yang berperan penting dalam fungsi kognitif. Proses diagnosis manual terhadap perubahan kecil pada biomarker otak menggunakan citra Magnetic Resonance Imaging (MRI) sering menghadapi kendala dari segi akurasi dan konsistensi. Untuk mengatasi hal tersebut, pendekatan Computer Aided Diagnosis (CAD) dikembangkan untuk membantu proses deteksi dan klasifikasi secara otomatis. Penelitian ini menggunakan data dari Alzheimer's Disease Neuroimaging Initiative (ADNI) dan menerapkan metode hybrid Inception v3 sebagai ekstraktor fitur dan Regularized Extreme Learning Machine (RELM) sebagai pengklasifikasi hasil ekstraksi. Tujuan penelitian ini adalah menentukan kombinasi hyperparameter terbaik melalui proses tuning untuk menghasilkan model klasifikasi yang optimal terhadap citra MRI penderita Alzheimer. Model ini dirancang untuk mengklasifikasikan enam tingkat keparahan kondisi kognitif, yaitu cognitively normal, subjective memory complaints, early mild cognitive impairment, mild cognitive impairment, late mild cognitive impairment, dan Alzheimer's disease. Berdasarkan hasil pengujian, model dengan parameter terbaik diperoleh pada nilai $C=0.01$, fungsi aktivasi Swish, 32768 neuron, dan batch size 2048, dengan akurasi sebesar 81.67%, precision 82.23%, recall 81.78%, specificity 96.23%, dan f1-score 81.31%.*

Kata Kunci: *Alzheimer's Disease, CNN, Inception v3, MRI, RELM*

1. Pendahuluan

Alzheimer merupakan penyebab paling umum dari demensia dan ditandai oleh penurunan progresif pada fungsi memori, bahasa, pemecahan masalah, serta kemampuan kognitif akibat kerusakan

neuron pada area otak terkait (Breijyeh & Karaman, 2021). Penyakit ini memengaruhi sekitar 27 juta orang dan mencakup 60-70% dari seluruh kasus demensia, dengan satu kasus baru muncul setiap tiga detik. *World Health Organization* (WHO) mencatat lebih dari 55 juta penderita demensia pada 2019 dan diproyeksikan meningkat menjadi 139 juta pada 2050, dengan beban biaya global mencapai US\$1,3 triliun dan diprediksi melonjak menjadi US\$2,8 triliun pada 2030 (Long et al., 2024). Karena meningkatkan kerentanan terhadap penyakit lain, Alzheimer menjadi penyebab kematian kedua utama pada lansia dengan sekitar 2,4 juta kematian per tahun dan diperkirakan meningkat menjadi 5,8 juta pada tahun 2040 (Breijyeh & Karaman, 2021; Liang et al., 2021). Di Indonesia, jumlah penderita demensia diperkirakan mencapai 1,2 juta pada 2016 dan meningkat menjadi 4 juta pada 2050 (Alzheimer's Indonesia, 2019)., sementara kematian akibat Alzheimer pada 2020 tercatat sebanyak 27.054 kasus atau sekitar 1,6% dari total kematian (WHO, 2020).

Karena Alzheimer tidak dapat disembuhkan sepenuhnya, diagnosis dini menjadi penting untuk memperlambat kerusakan otak yang *irreversible* dan mempertahankan kemandirian pasien lebih lama (Shamrat et al., 2023). MRI merupakan teknik *neuroimaging* yang umum digunakan untuk mendeteksi Alzheimer pada tahap awal karena lebih aman dibandingkan PET, FDG-PET, maupun DTI, sebab tidak memerlukan penyuntikan isotop radioaktif (Zhao et al., 2023). Namun, interpretasi citra MRI secara manual masih memiliki keterbatasan dan berpotensi menimbulkan kesalahan diagnostik. Oleh karena itu, diperlukan dukungan sistem *Computer Aided Diagnosis* (CAD) untuk memberikan analisis lebih objektif, efisien, dan konsisten (Ismail et al., 2024) untuk membantu diagnosis dini penyakit Alzheimer. Berbagai penelitian sebelumnya telah menunjukkan efektivitas CAD berbasis citra MRI dalam diagnosis dini Alzheimer, dengan implementasi metode *machine learning* maupun *deep learning*. (Lin et al., 2023) melaporkan akurasi 89.7% untuk klasifikasi tiga kategori, sementara (Antor et al., 2021) memperoleh 92% menggunakan *Support Vector Machine* (SVM). Pendekatan CNN juga menunjukkan kinerja unggul, sebagaimana ditunjukkan oleh (Salehi et al., 2020) dengan akurasi 99% dan (El-Assy et al., 2024) yang meraih akurasi di atas 99% untuk klasifikasi tiga hingga lima tingkat keparahan Alzheimer.

CNN, termasuk arsitektur *GoogleNet* dan pengembangannya dalam keluarga *Inception* telah banyak digunakan karena kemampuannya mempelajari fitur citra secara mendalam dengan efisiensi komputasi yang lebih baik (Zhong et al., 2015). *Inception* sebagai kelanjutan dari *GoogleNet* meningkatkan kedalaman jaringan tanpa penambahan parameter yang signifikan, sekaligus mengurangi risiko *overfitting* melalui optimalisasi struktur modulnya (Ahmed & Cho, 2020). *Inception v2* dikembangkan untuk mengurangi kompleksitas jaringan *convolution* dengan membuat arsitektur yang lebih lebar (Alamsyah & Fachrurrozi, 2019), sedangkan *Inception v3* menawarkan peningkatan akurasi, efisiensi komputasi, serta stabilitas pelatihan melalui perbaikan pada teknik *factorization*, regulasi, dan optimalisasi arsitektur (C. Lin et al., 2019). Berbagai penelitian menunjukkan efektivitas *Inception v3* dalam klasifikasi penyakit Alzheimer. Cui, dkk. mengimplementasikan arsitektur *Inception v3* untuk klasifikasi penyakit Alzheimer dari dataset ADNI dalam tiga tingkat keparahan dan menghasilkan akurasi sebesar 85.71% (Cui et al., 2019). Ghaffari, dkk. menerapkan tiga model pelatihan CNN (*ResNet101*, *Xception*, dan *Inception v3*) dan mendapatkan akurasi tertinggi 93.75% pada metode *Inception v3* untuk klasifikasi Alzheimer ke dalam empat tingkat keparahan dengan waktu *training* 15 jam (Ghaffari et al., 2022). Shamrat, dkk. menggunakan lima model CNN (VGG16, *MobileNet v2*, *AlexNet*, *ResNet50*, dan *Inception v3 modified*) untuk klasifikasi ke dalam enam tingkat keparahan, akurasi terbaik dihasilkan oleh *Inception v3* sebesar 98.68% dengan waktu *training* 6 hari (Shamrat et al., 2023). Meskipun demikian, kompleksitas struktur hirarki dan banyaknya parameter membuat proses pelatihan *Inception v3* cukup memakan waktu, sehingga pemisahan antara *feature learning* dan tahap klasifikasi menjadi diperlukan untuk mengatasi keterbatasan tersebut. Pendekatan ini membuka peluang penerapan metode *hybrid* yang mengombinasikan CNN sebagai ekstraktor fitur dengan algoritma klasifikasi yang lebih ringan dan cepat.

Salah satu algoritma yang banyak digunakan untuk meningkatkan efisiensi proses klasifikasi adalah *Extreme Learning Machine* (ELM). ELM merupakan algoritma *machine learning* berbasis *Single-Hidden Layer Feed Forward* (SLFN) yang dikenal sederhana, cepat (Li et al., 2022), dan memiliki kemampuan generalisasi yang baik karena bobot *output* dihitung menggunakan *Moore-Penrose inverse*, sementara bobot dan bias ditentukan secara acak (Aisah et al., 2023). Meskipun

demikian, ELM didasarkan pada prinsip *Empirical Risk Minimization* (ERM) yang cenderung rentan terhadap *overfitting*. Selain itu, ELM memiliki kontrol kapasitas model yang relatif lemah akibat penentuan solusi secara langsung melalui pendekatan *minimum norm least squares*, sehingga estimasi yang dihasilkan kurang *robust*, terutama ketika terdapat *outlier* atau heteroskedastisitas pada data (Deng et al., 2009). Untuk mengatasi masalah tersebut, dikembangkan *Regularized ELM* (RELM) dengan menambahkan parameter regularisasi pada proses penentuan bobot *output*, sehingga model lebih optimal. Beberapa penelitian menunjukkan hasil yang konsisten bahwa penerapan regularisasi pada ELM mampu meningkatkan stabilitas model dan performa klasifikasi dibandingkan ELM standar (Iosifidis et al., 2015; Martínez-Martínez et al., 2011; Yu et al., 2013). Oleh karena itu, penelitian ini mengusulkan pendekatan *hybrid* yang menggabungkan *Inception v3* sebagai ekstraktor fitur dengan RELM sebagai pengklasifikasi untuk membedakan enam tingkat keparahan penyakit Alzheimer menggunakan dataset *Alzheimer's Disease Neuroimaging Initiative* (ADNI). Pada penelitian ini, fitur tingkat tinggi yang dihasilkan dari *Inception v3* direduksi menggunakan *Principal Component Analysis* (PCA) untuk mengurangi dimensi fitur dengan mempertahankan sebagian besar variansi data sehingga proses klasifikasi menjadi lebih efisien. Selanjutnya, model dievaluasi menggunakan metode *Repeated Stratified K-Fold Cross Validation* untuk memastikan stabilitas performa klasifikasi. Selain itu, penelitian ini juga mengeksplorasi berbagai *hyperparameter* pada RELM, meliputi nilai parameter regularisasi (C), fungsi aktivasi, jumlah neuron pada *hidden layer*, serta *batch size*, guna memperoleh konfigurasi model yang optimal.

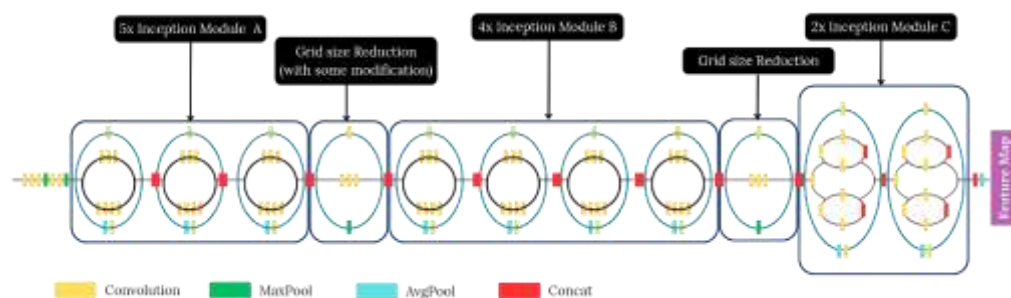
2. Tinjauan Pustaka

2.1. Alzheimer

Alzheimer merupakan penyakit yang terjadi pada bagian tengah otak yang menyebabkan hilangnya memori secara bertahap, gangguan kognitif, dan tekanan emosional (ud din dar et al., 2023). Cita MRI pada kelas *Cognitively Normal* (CN) umumnya menunjukkan struktur otak yang masih utuh tanpa tanda atrofi atau penumpukan plak patologis (Xie et al., 2023), sedangkan pada *Subjective Memory Complaints* (SMC) mulai tampak perubahan halus berupa penurunan volume *hippocampus* atau korteks *entorhinal* (Dhana et al., 2022). Pada *Early Mild Cognitive Impairment* (EMCI), perubahan struktural menjadi lebih jelas dan disertai indikasi kerusakan mikrostruktur *white matter* (Kang et al., 2020), sementara pada *Mild Cognitive Impairment* (MCI) penipisan korteks di area *temporal medial* dan *entorhinal* serta perubahan *white matter* semakin nyata (Song et al., 2021). Individu dengan *Late MCI* (LMCI) menunjukkan penurunan volume otak yang lebih besar, khususnya di *hippocampus* serta kerusakan *white matter* dan penipisan korteks yang lebih menonjol (Prado & Rojas, 2021). Pada tahap *Alzheimer's Disease* (AD), citra MRI memperlihatkan perubahan struktural paling signifikan, dengan atrofi luas pada *hippocampus*, korteks *entorhinal*, dan *lobus temporal medial* serta penurunan integritas *white matter* yang mengganggu konektivitas antarruang otak (Yamanakkanavar et al., 2020).

2.2. Inception v3

Inception v3 merupakan arsitektur CNN berlapis 42 yang dirancang untuk ekstraksi fitur secara efisien melalui teknik seperti *factorized convolutions*, yang mengganti *large convolution* dengan beberapa *small convolution* untuk mengurangi parameter, serta *asymmetric convolution* guna meningkatkan efisiensi pemrosesan (Ramaneswaran et al., 2021).

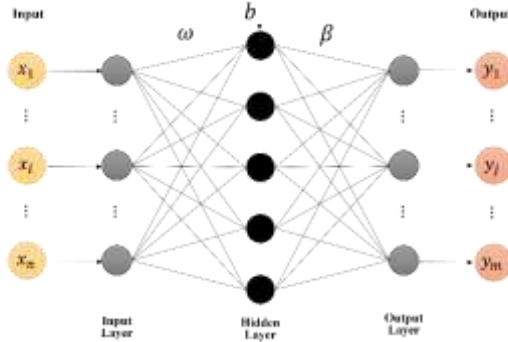


Gambar 1. Arsitektur *Inception v3 Feature Learning*

Pada tahap ekstraksi fitur, citra yang telah diproses dilewatkan melalui sejumlah *convolution layer* dalam *inception module* yang bekerja paralel, menghasilkan *feature maps* dari berbagai skala dan karakteristik sebelum digabungkan menjadi satu representasi fitur yang komprehensif. Output akhir dari model kemudian diperoleh dalam bentuk vektor fitur berdimensi 1×2048 (Shen & Li, 2020). Arsitektur *Inception v3* sebagai *feature learning* ditunjukkan pada Gambar 1.

2.3. Regularized Extreme Learning Machine (RELM)

Extreme Learning Machine (ELM) dikembangkan untuk melatih jaringan *single hidden layer feedforward* (SLFN) dengan keunggulan berupa penentuan bobot input dan bias secara acak, sehingga pemetaan neuron tersembunyi dapat terdiferensiasi secara kontinu (Wang et al., 2022). Dengan karakteristik tersebut, ELM membentuk sistem linier di mana bobot keluaran dapat dihitung secara langsung menggunakan *generalized inverse* (Cheng et al., 2020). Arsitektur ELM ditunjukkan pada Gambar 2. Meskipun ELM menawarkan pelatihan yang cepat, metode ini memiliki kelemahan karena rentan menghasilkan *overfitting* dan memberikan kontrol kapasitas yang lemah. *Regularized Extreme Learning Machine* (RELM) dikembangkan untuk mengatasi keterbatasan tersebut dengan menerapkan prinsip *Structural Risk Minimization* (SRM), namun tetap mempertahankan efisiensi komputasi ELM karena bobot keluarannya tetap dihitung menggunakan *invers Moore–Penrose* (Várkonyi & Buza, 2019). Secara umum, ELM dapat ditunjukkan pada Persamaan (1).



Gambar 2. Arsitektur ELM

$$H\beta = Y \quad (1)$$

Di mana H adalah *output* matriks *hidden layer*, β adalah matriks bobot *output*, dan Y adalah matriks label yang benar.

$$H = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_N) \end{bmatrix} = \begin{bmatrix} g(\omega_1 x_1 + b_1) & \cdots & g(\omega_L x_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(\omega_1 x_N + b_1) & \cdots & g(\omega_L x_N + b_L) \end{bmatrix}_{N \times L} \quad (2)$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix} = \begin{bmatrix} \beta_{1,1} & \cdots & \beta_{1,m} \\ \vdots & \ddots & \vdots \\ \beta_{L,1} & \cdots & \beta_{L,m} \end{bmatrix}_{L \times m} \quad (3)$$

$$Y = \begin{bmatrix} y_1^T \\ \vdots \\ y_L^T \end{bmatrix} = \begin{bmatrix} y_{1,1} & \cdots & y_{1,m} \\ \vdots & \ddots & \vdots \\ y_{N,1} & \cdots & y_{N,m} \end{bmatrix}_{N \times m} \quad (4)$$

Variabel x_i menyatakan vektor fitur sampel ke- i , w_j adalah bobot yang menghubungkan *input* dengan neuron tersembunyi ke- j , b_j merupakan bias pada neuron tersebut, $g(\cdot)$ adalah fungsi aktivasi, dan m jumlah kelas *output*. (Bartlett, 1998) menyatakan bahwa semakin kecil *norm* dari vektor bobot, semakin baik kemampuan generalisasi model. Oleh karena itu, regulasi pada ELM bertujuan meminimalkan kesalahan pelatihan sekaligus memperkecil *norm* bobot keluaran, yang dapat dihitung menggunakan Persamaan (5).

$$B = \begin{cases} \left(H^T H + \frac{I}{C} \right)^{-1} H^T Y & \text{if } N > L \\ H^T \left(H H^T + \frac{I}{C} \right)^{-1} Y & \text{if } N \leq L \end{cases} \quad (5)$$

Di mana H^T menyatakan transpose dari matriks H , I adalah matriks identitas, C adalah parameter regularisasi, notasi N menunjukkan jumlah sampel, dan L jumlah neuron pada *hidden layer*. Maka, *output* dari ELM yang diregularisasi dapat dihitung menggunakan Persamaan (6).

$$Y' = \begin{cases} h(x) \left(H^T H + \frac{I}{C} \right)^{-1} H^T Y & \text{if } N < L \\ h(x) H^T \left(H^T H + \frac{I}{C} \right)^{-1} & \text{if } N > L \end{cases} \quad (6)$$

Di mana Y' merupakan matriks prediksi *output* model, sedangkan $h(x)$ menyatakan vektor *output* dari *hidden later* untuk data *input* x .

2.4. Confusion Matrix

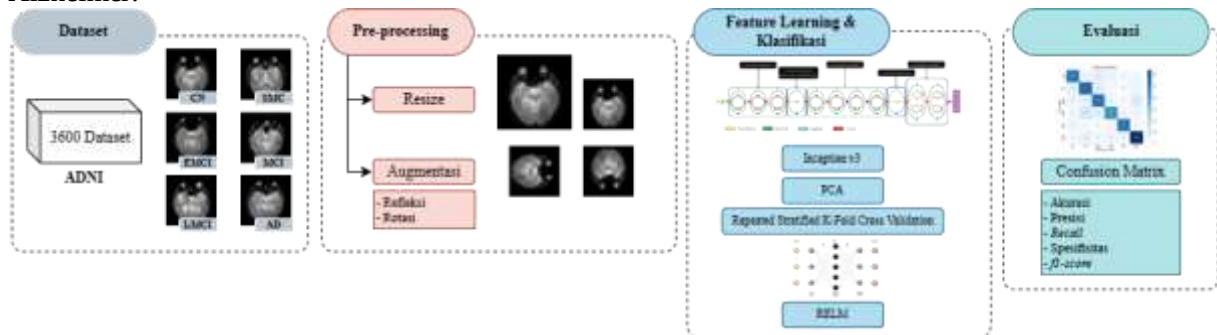
Confusion matrix didefinisikan sebagai matriks dengan elemen berupa bilangan bulat positif yang merepresentasikan jumlah instances dari pasangan kombinasi label kelas yang diprediksi dan kelas aktual (Markoulidakis & Markoulidakis, 2024). *Confusion matrix* umumnya digunakan untuk mengevaluasi kinerja dari algoritma klasifikasi yang dapat dihitung menggunakan rumus pada Tabel 1. Di mana, TP (True Positive), FP (False Positive), TN (True Negative), dan FN (False Negative).

Tabel 1. Rumus Metrik Evaluasi

Metrik	Deskripsi
$Accuracy = \sum_{i=0}^n \frac{TP_i}{TP_i + FP_i + TN_i + FN_i}$	Proporsi prediksi yang benar dari seluruh sampel yang diuji.
$Precision = \frac{\sum_{i=1}^n \frac{TP_i}{TP_i + FP_i}}{n}$	Seberapa tepat model dalam mengidentifikasi kelas positif.
$Sensitivity = \frac{\sum_{i=1}^n \frac{TP_i}{TP_i + FN_i}}{n}$	Kemampuan model menemukan seluruh sampel positif yang sebenarnya positif.
$Specificity = \frac{\sum_{i=1}^n \frac{TN_i}{TN_i + FP_i}}{n}$	Kemampuan model mengenali sampel negatif dengan benar.
$f1 - score = \frac{\sum_{i=1}^n \frac{2 \times Precision_i \times Recall_i}{Precision_i + Recall_i}}{n}$	Rata-rata harmonik metrik tunggal antara presisi dan <i>recall</i> .
Di mana n adalah banyaknya kelas dan i menunjukkan kelas ke- i .	

3. Metode Penelitian

Gambar 3 menunjukkan alur penelitian yang dimulai dari pengumpulan dataset citra MRI dari *Alzheimer's Disease Neuroimaging Initiative* (ADNI) yang terdiri dari enam tingkat keparahan penyakit Alzheimer.



Gambar 3. Flowchart dalam Penelitian Ini

Selanjutnya, citra diproses melalui tahap *pre-processing* yang mencakup proses *resize* untuk menyesuaikan ukuran citra menjadi 299×299 piksel sesuai dengan kebutuhan arsitektur Inception v3 tanpa memodifikasi gambar asli, dan pada penelitian ini digunakan interpolasi bilinear karena

menghasilkan kualitas citra yang lebih baik dibandingkan *nearest-neighbor* (Ahn et al., 2018). Selain itu, dilakukan augmentasi citra untuk mencegah *overfitting* pada pelatihan *neural network* dengan melakukan berbagai transformasi seperti rotasi dan refleksi terhadap citra asli (Awaluddin et al., 2023). Selanjutnya, model Inception v3 digunakan sebagai *feature learning* untuk memperoleh representasi fitur tingkat tinggi, yang kemudian direduksi menggunakan *Principal Component Analysis* (PCA) untuk mengurangi dimensi fitur serta meningkatkan efisiensi komputasi. Proses klasifikasi dilakukan menggunakan *Regularized ELM* (RELM) dengan beberapa parameter yang diuji meliputi nilai parameter regularisasi C, jumlah neuron pada *hidden layer*, dan fungsi aktivasi. Pada tahap evaluasi, model diuji menggunakan metode *Repeated Stratified K-Fold Cross Validation* untuk memastikan distribusi kelas yang seimbang pada setiap fold. kinerja model kemudian di evaluasi menggunakan metrik klasifikasi seperti akurasi, presisi, recall, spesifisitas, dan f1-score serta dianalisis menggunakan *confusion matrix* untuk menilai kemampuan model dalam membedakan setiap tingkat keparahan penyakit Alzheimer.

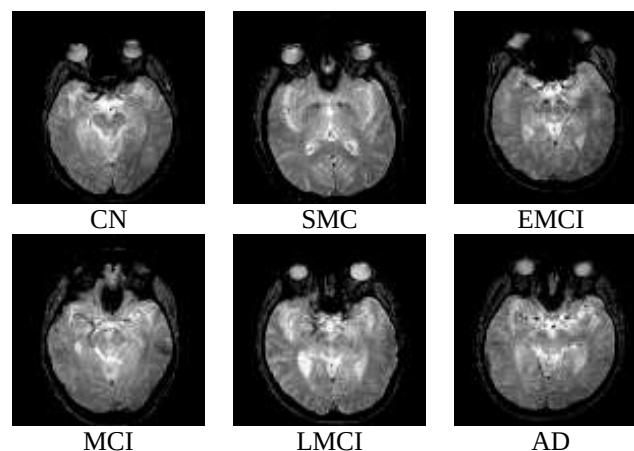
4. Hasil dan Pembahasan

4.1. Dataset

Penelitian ini memanfaatkan data yang diambil dari ADNI, basis data tersebut menyediakan pemindaian citra MRI berbobot *T2 Star* dengan irisan *Axial* dari 883 subjek yang berbeda dengan 6 kelas. Sampel data dapat dilihat pada Gambar 4.

4.2. Pre-Processing

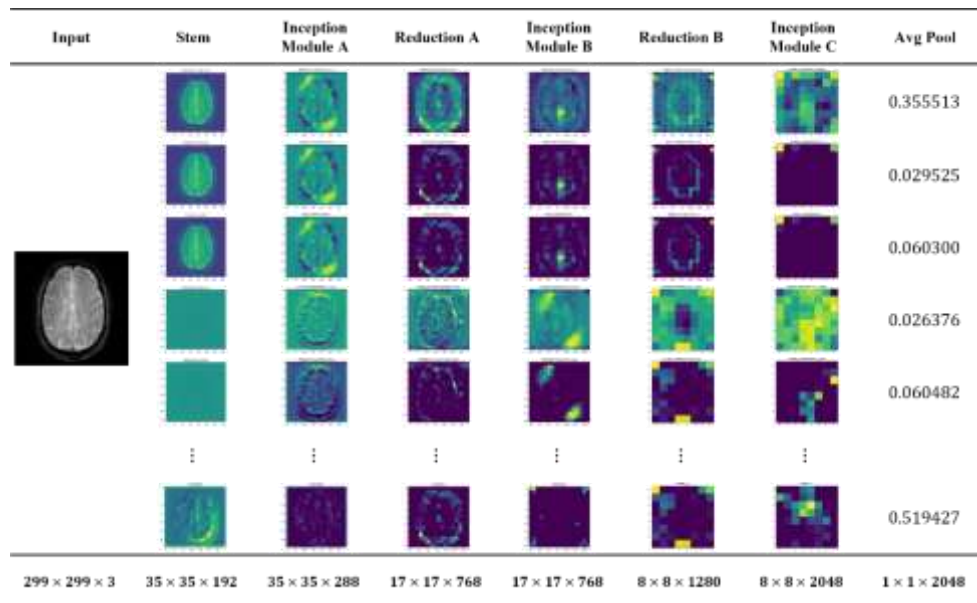
Proses *resize* dilakukan untuk memastikan bahwa dimensi citra sesuai dengan arsitektur Inception v3, ukuran citra diubah dari 256×256 menjadi 299×299 menggunakan metode *interpolasi bilinear*. Selanjutnya, augmentasi citra dalam penelitian ini dilakukan menggunakan teknik translasi, refleksi, dan rotasi. Hasil augmentasi setiap kelas menghasilkan 600 citra sehingga dataset akhir terdiri dari 3600 citra.



Gambar 4. Dataset Citra MRI ADNI

4.3. Feature Learning

Dalam arsitektur berbasis *deep learning*, setiap lapisan *convolutional* menghasilkan transformasi tertentu terhadap *input* citra sehingga informasi visual yang awalnya bersifat sederhana secara bertahap direpresentasikan dalam bentuk yang semakin abstrak dan selektif. Gambar 5 menunjukkan bahwa lapisan awal (*stem*) masih menyoroti struktur dasar citra, sementara lapisan-lapisan berikutnya mulai memunculkan pola yang lebih kompleks sebagai hasil dari kombinasi operasi *convolution* di dalam *Inception Module*. Reduksi resolusi pada blok *Reduction* menyebabkan fitur yang bersifat global menjadi lebih dominan, sedangkan blok *Inception* selanjutnya menghasilkan representasi yang semakin abstrak dan kaya secara semantik. Pada akhirnya, proses *Global Average Pooling* mengonsolidasikan seluruh peta fitur tersebut menjadi vektor berukuran 1×2048 yang kemudian digunakan sebagai *input* untuk tahap klasifikasi menggunakan RELM.



Gambar 5. Ilustrasi Perubahan Feature Map pada Inception v3

4.4. Klasifikasi

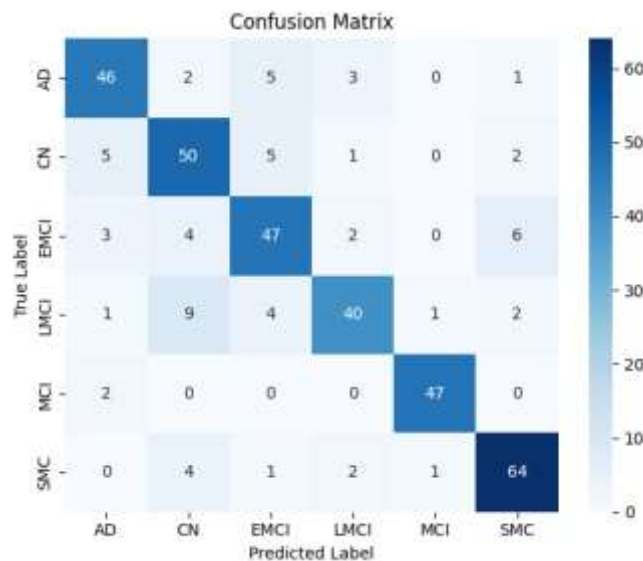
Tahap klasifikasi pada penelitian ini menggunakan percobaan *hyperparameter*, yaitu fungsi aktivasi (*Sigmoid*, *ReLU*, *Softplus*, *Swish*, dan *Mish*), koefisien regulasi (0.01, 0.1, 1, 10, 100, 1000, 10000, dan 100000), neuron (256, 512, 1024, 2048, 4096, 8192, dan 32768), serta *batch size* (256, 512, 1024, dan 2048). Setiap kombinasi diuji menggunakan skema *Repeated Stratified K-Fold Cross Validation* untuk memastikan proporsi kelas tetap seimbang dan evaluasi lebih stabil. Pada setiap *fold*, fitur hasil ekstraksi *Inception v3* terlebih dahulu dinormalisasi dan direduksi menggunakan PCA sebelum digunakan sebagai *input* RELM. Secara keseluruhan, kombinasi terbaik dalam penelitian ini diperoleh pada fungsi aktivasi *Swish* dengan nilai koefisien regulasi $C = 0.01$, 32768 neuron, dan *batch size* 2048 yang menghasilkan akurasi tertinggi 0.810.

4.5. Evaluasi

Berdasarkan hasil dari *confusion matrix* pada Gambar 6, model menunjukkan kinerja *True Positive* (TP) tertinggi pada kelas SMC dengan 64 kasus terklasifikasi benar. Selain itu, kelas CN mendapatkan 50 kasus teridentifikasi dengan tepat. Hal ini mengindikasikan bahwa ciri-ciri fitur pencitraan untuk status keluhan subjektif telah terekstraksi dan dikenali dengan baik oleh model. Namun, kesalahan klasifikasi yang berpotensi berdampak secara klinis yaitu 5 kasus CN yang salah diprediksi sebagai AD serta 2 kasus AD diprediksi sebagai CN. Hal ini menunjukkan perlunya pengembangan model lebih lanjut untuk meminimalkan kesalahan klasifikasi yang berpotensi berdampak pada proses diagnosis. Tanda (-) tidak dilaporkan pada penelitian terkait.

Tabel 2. Perbandingan Hasil Penelitian dengan Penelitian Terdahulu Menggunakan dataset ADNI

Sumber	Metode	Data	Kelas	Akurasi	Presisi	Recall	Spes	F1-Score	Waktu	
									Train	Inference
(Shamrat et al., 2023)	<i>Inception v3</i>	60000	6	98.68%	98.68%	98.68	99.74%	98.68%	6d	-
(Ghaffari et al., 2022)	<i>Inception v3</i>	336	4	93.75%	95.00%	93.75%	97.92%	93.65%	5h	-
(Alayba et al., 2024)	CNN-ANN	1296	5	98.80%	97.08%	98.90%	99.50%	-	-	8m 45s
Penelitian ini	<i>Inception v3-RELM</i>	3600	6	81.67%	82.23%	81.78%	96.23%	81.31%	6h 35m	10.255s



Gambar 6. Confusion Matrix

Perbandingan hasil penelitian sebelumnya pada Tabel 2 menunjukkan bahwa penelitian oleh (Shamrat et al., 2023) dan (Ghaffari et al., 2022) yang menggunakan *Inception v3* murni menghasilkan metrik evaluasi sangat tinggi, namun membutuhkan waktu pelatihan yang lama, masing-masing 6 hari dan 15 jam. Selain itu, (Ghaffari et al., 2022) menggunakan dataset yang jauh lebih kecil serta jumlah kelas yang lebih sedikit dibandingkan penelitian ini, tetapi tetap membutuhkan waktu pelatihan yang lebih lama. Sebagai perbandingan, penelitian ini memanfaatkan *feature learning* dari *Inception v3* yang memerlukan waktu pelatihan sekitar 6 jam 35 menit pada dataset berjumlah 3600 citra. Sementara itu, penelitian (Alayba et al., 2024) yang menggunakan model *hybrid* menunjukkan waktu *inference* sekitar 8 menit 45 detik pada dataset 1296 citra dengan lima kelas klasifikasi. Pada penelitian ini, dengan menggunakan dataset yang lebih besar dan kelas yang lebih banyak, model *Inception v3-RELM* menghasilkan waktu *inference*/klasifikasi sebesar 10,255 detik pada lingkungan GPU setara 2× NVIDIA T4. Perbedaan waktu komputasi tersebut dapat dipengaruhi oleh ukuran dataset, jumlah kelas klasifikasi, konfigurasi model, serta spesifikasi perangkat keras yang digunakan pada masing-masing penelitian. Namun demikian, perbandingan secara langsung terhadap waktu penelitian masih memiliki keterbatasan karena tidak semua penelitian sebelumnya melaporkan waktu pelatihan dan *inference* secara rinci. Secara keseluruhan, hasil penelitian ini menunjukkan bahwa pendekatan *hybrid Inception v3-RELM* berpotensi memberikan keseimbangan antara akurasi klasifikasi dan efisiensi komputasi dalam proses klasifikasi citra MRI Alzheimer.

5. Simpulan

Hasil penelitian menunjukkan bahwa kombinasi *Inception v3* sebagai *deep feature extractor* dan *Regularized Extreme Learning Machine* (RELM) sebagai pengklasifikasi mampu memberikan kinerja yang efektif dalam mendeteksi tingkat perkembangan Alzheimer. *Inception v3* berhasil mengekstraksi 2048 fitur yang representatif terhadap karakteristik citra MRI, sementara RELM dengan konfigurasi optimal berupa nilai $C = 0,01$, fungsi aktivasi *Swish*, 32768 neuron, dan *batch size* 2048 menghasilkan performa klasifikasi dengan akurasi 81,67%, presisi 82,23%, *recall* 81,78%, spesifisitas 96,23%, dan *f1-score* 81,31%. Jika dibandingkan dengan beberapa penelitian sebelumnya yang menggunakan arsitektur *Inception v3* secara *end-to-end*, pendekatan *hybrid* pada penelitian ini menunjukkan efisiensi komputasi yang lebih baik pada proses pelatihan serta waktu klasifikasi yang relatif lebih cepat dibandingkan pendekatan *hybrid* pada penelitian sebelumnya. Perbedaan performa akurasi dengan penelitian sebelumnya dapat dipengaruhi oleh beberapa faktor, seperti perbedaan ukuran

dataset, jumlah kelas klasifikasi, konfigurasi model, serta strategi pelatihan yang digunakan pada masing-masing penelitian.

Secara keseluruhan, pendekatan *hybrid Inception v3-RELM* menunjukkan potensi dalam memberikan keseimbangan antara efisiensi komputasi dan kemampuan ekstraksi fitur dalam klasifikasi citra MRI Alzheimer. Pengembangan pada penelitian selanjutnya dapat difokuskan pada optimasi parameter RELM secara lebih adaptif, eksplorasi arsitektur *deep learning* lain melalui pendekatan *transfer learning*, serta modifikasi pada bagian ekstraksi fitur untuk meningkatkan kualitas representasi fitur yang dihasilkan. Selain itu, eksplorasi teknik augmentasi data yang lebih beragam juga dapat dipertimbangkan untuk meningkatkan *robustness* model terhadap variasi data.

Referensi

- Ahmed, S., & Cho, S. H. (2020). Hand gesture recognition using an IR-UWB radar with an inception module-based classifier. *Sensors (Switzerland)*, 20(2). <https://doi.org/10.3390/s20020564>
- Ahn, J., Park, J., Park, D., Paek, J., & Ko, J. (2018). Convolutional neural network-based classification system design with compressed wireless sensor network images. *PLoS ONE*, 13(5), 1–25. <https://doi.org/10.1371/journal.pone.0196251>
- Aisah, S. N., Candra, D., Novitasari, R., & Farida, Y. (2023). Perbandingan Metode Extreme Learning Machine (ELM) dan Kernel Extreme Learning Machine (KELM) Pada Klasifikasi Penyakit Cedera Panggul. *Journal Fourier*, 12(2), 69–78. <https://doi.org/10.14421/fourier.2023.122.69-78>
- Alamsyah, D., & Fachrurrozi, M. (2019). Faster R-CNN with inception v2 for fingertip detection in homogenous background image. *Journal of Physics: Conference Series*, 1196(1). <https://doi.org/10.1088/1742-6596/1196/1/012017>
- Alayba, A. M., Senan, E. M., & Alshudukhi, J. S. (2024). Enhancing early detection of Alzheimer's disease through hybrid models based on feature fusion of multi-CNN and handcrafted features. *Scientific Reports*, 14(1). <https://doi.org/10.1038/s41598-024-82544-y>
- Alzheimer's Indonesia. (2019). *Statistik tentang Demensia*. <https://alzi.or.id/statistik-tentang-demensia/>
- Antor, M. B., Jamil, A. H. M. S., Mamtaz, M., Khan, M. M., Aljahdali, S., Kaur, M., Singh, P., & Masud, M. (2021). A Comparative Analysis of Machine Learning Algorithms to Predict Alzheimer's Disease. *Journal of Healthcare Engineering*, 2021. <https://doi.org/10.1155/2021/9917919>
- Awaluddin, B., Chao, C., & Chiou, J.-S. (2023). Investigating Effective Geometric Transformation for Image Augmentation to Improve Static Hand Gestures with a Pre-Trained Convolutional Neural Network. *Mathematics*, 11(23), 4783. <https://doi.org/10.3390/math11234783>
- Bartlett, P. L. (1998). The sample complexity of pattern classification with neural networks: the size of the weights is more important than the size of the network. *IEEE Transactions on Information Theory*, 44(2), 525–536. <https://doi.org/10.1109/18.661502>
- Breijyeh, Z., & Karaman, R. (2021). a Comprehensive Review on Alzheimer's Disease. *World Journal of Pharmacy and Pharmaceutical Sciences*, 10(7), 1170. <https://doi.org/10.20959/wjpps20217-19427>
- Chen, Q., Wei, H., Rashid, M., & Cai, Z. (2021). Kernel extreme learning machine based hierarchical machine learning for multi-type and concurrent fault diagnosis. *Measurement: Journal of the International Measurement Confederation*, 184(July), 109923. <https://doi.org/10.1016/j.measurement.2021.109923>
- Cheng, X., Feng, Z. K., & Niu, W. J. (2020). Forecasting Monthly Runoff Time Series by Single-Layer Feedforward Artificial Neural Network and Grey Wolf Optimizer. *IEEE Access*, 8, 157346–157355. <https://doi.org/10.1109/ACCESS.2020.3019574>
- Cui, Z., Gao, Z., Leng, J., Zhang, T., Quan, P., & Zhao, W. (2019). Alzheimer's Disease Diagnosis Using Enhanced Inception Network Based on Brain Magnetic Resonance Image. *Proceedings - 2019 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2019*, 2324–2330. <https://doi.org/10.1109/BIBM47256.2019.8983046>
- Dementia, U. K. (2021). *What is Dementia*. <https://keystonemedical.com.sg/dementia/>
- Deng, W., Zheng, Q., & Chen, L. (2009). Regularized Extreme Learning Machine. *IE*.

- Dhana, A., Decarli, C., Dhana, K., Desai, P., Krueger, K., Evans, D. A., & Rajan, K. B. (2022). Association of Subjective Memory Complaints with White Matter Hyperintensities and Cognitive Decline among Older Adults in Chicago, Illinois. *JAMA Network Open*, 5(4), E227512. <https://doi.org/10.1001/jamanetworkopen.2022.7512>
- El-Assy, A. M., Amer, H. M., Ibrahim, H. M., & Mohamed, M. A. (2024). A novel CNN architecture for accurate early detection and classification of Alzheimer's disease using MRI data. *Scientific Reports*, 14(1), 1–19. <https://doi.org/10.1038/s41598-024-53733-6>
- Ghaffari, H., Tavakoli, H., & Jahromi, G. P. (2022). Deep transfer learning-based fully automated detection and classification of Alzheimer's disease on brain MRI. *British Journal of Radiology*, 95(1136). <https://doi.org/10.1259/bjr.20211253>
- Iosifidis, A., Tefas, A., & Pitas, I. (2015). Regularized Extreme Learning Machine for Large-scale Media Content Analysis. *Procedia Computer Science*, 53, 420–427. <https://doi.org/10.1016/j.procs.2015.07.319>
- Ismail, M. Y., Sunyoto, A., & Purwanto, A. (2024). Klasifikasi Penyakit Alzheimer pada Citra Medis Magnetic Resonance Images dengan Arsitektur DenseNet121. *Jurnal Ilmiah Komputasi*, 23(2), 275–282. <https://doi.org/10.32409/jikstik.23.2.3600>
- Kang, L., Jiang, J., Huang, J., & Zhang, T. (2020). Identifying Early Mild Cognitive Impairment by Multi-Modality MRI-Based Deep Learning. *Frontiers in Aging Neuroscience*, 12(September), 1–10. <https://doi.org/10.3389/fnagi.2020.00206>
- Li, C., Zhou, J., Dias, D., & Gui, Y. (2022). A Kernel Extreme Learning Machine-Grey Wolf Optimizer (KELM-GWO) Model to Predict Uniaxial Compressive Strength of Rock. *Applied Sciences (Switzerland)*, 12(17). <https://doi.org/10.3390/app12178468>
- Liang, C.-S., Li, D.-J., Yang, F.-C., Tseng, P.-T., Carvalho, A. F., Stubbs, B., Thompson, T., Mueller, C., Shin, J. Il, Radua, J., Stewart, R., Rajji, T. K., Tu, Y.-K., Chen, T.-Y., Yeh, T.-C., Tsai, C.-K., Yu, C.-L., Pan, C.-C., & Chu, C.-S. (2021). Mortality rates in Alzheimer's disease and non-Alzheimer's dementias: a systematic review and meta-analysis. *The Lancet. Healthy Longevity*, 2(8), e479–e488. [https://doi.org/10.1016/S2666-7568\(21\)00140-9](https://doi.org/10.1016/S2666-7568(21)00140-9)
- Lin, C., Li, L., Luo, W., Wang, K. C. P., & Guo, J. (2019). Transfer learning based traffic sign recognition using inception-v3 model. *Periodica Polytechnica Transportation Engineering*, 47(3), 242–250. <https://doi.org/10.3311/PPtr.11480>
- Lin, Q., Che, C., Hu, H., Zhao, X., & Li, S. (2023). A Comprehensive Study on Early Alzheimer's Disease Detection through Advanced Machine Learning Techniques on MRI Data. *Academic Journal of Science and Technology*, 8(1), 281–285. <https://doi.org/10.54097/ajst.v8i1.14334>
- Long, S., Benoist, C., & Weidner, W. (2024). World Alzheimer Report 2024. *Alzheimer's Disease International*, 94.
- Markoulidakis, I., & Markoulidakis, G. (2024). Probabilistic Confusion Matrix: A Novel Method for Machine Learning Algorithm Generalized Performance Analysis. *Technologies*, 12(7). <https://doi.org/10.3390/technologies12070113>
- Martínez-Martínez, J., Escandell-Montero, P., Olivas, E., Martín-Guerrero, J., Magdalena, R., & Gómez-Sanchís, J. (2011). Regularized extreme learning machine for regression problems. *Neurocomputing - IJON*, 74, 3716–3721. <https://doi.org/10.1016/j.neucom.2011.06.013>
- Prado, J. J., & Rojas, I. (2021). Machine Learning for Diagnosis of Alzheimer's Disease and Early Stages. *BioMedInformatics*, 1(3), 182–200. <https://doi.org/10.3390/biomedinformatics1030012>
- Ramaneswaran, S., Srinivasan, K., Vincent, P. M. D. R., & Chang, C. Y. (2021). Hybrid Inception v3 XGBoost Model for Acute Lymphoblastic Leukemia Classification. *Computational and Mathematical Methods in Medicine*, 2021, 1–10. <https://doi.org/10.1155/2021/2577375>
- Salehi, A. W., Baglat, P., Sharma, B. B., Gupta, G., & Upadhyay, A. (2020). A CNN Model: Earlier Diagnosis and Classification of Alzheimer Disease using MRI. *Proceedings - International Conference on Smart Electronics and Communication, ICOSSEC 2020, November*, 156–161. <https://doi.org/10.1109/ICOSSEC49089.2020.9215402>
- Shamrat, F. M. J. M., Akter, S., Azam, S., Karim, A., Ghosh, P., Tasnim, Z., Hasib, K. M., De Boer, F., & Ahmed, K. (2023). AlzheimerNet: An Effective Deep Learning Based Proposition for

- Alzheimer's Disease Stages Classification From Functional Brain Changes in Magnetic Resonance Images. *IEEE Access*, 11, 16376–16395. <https://doi.org/10.1109/ACCESS.2023.3244952>
- Shen, W., & Li, X. (2020). Facial expression recognition based on bidirectional gated recurrent units within deep residual network. *International Journal of Intelligent Computing and Cybernetics*, 13(4), 527–543. <https://doi.org/10.1108/IJICC-07-2020-0088>
- Song, Y., Xu, W., Chen, S., Hu, G., Ge, H., Xue, C., Qi, W., Lin, X., & Chen, J. (2021). Functional MRI-Specific Alterations in Salience Network in Mild Cognitive Impairment: An ALE Meta-Analysis. *Frontiers in Aging Neuroscience*, 13(July), 1–14. <https://doi.org/10.3389/fnagi.2021.695210>
- ud din dar, G., Bhagat, A., Ansarullah, S. I., Othman, M. T. Ben, Hamid, Y., Alkahtani, H. K., Ullah, I., & Hamam, H. (2023). A Novel Framework for Classification of Different Alzheimer's Disease Stages Using CNN Model. *Electronics*, 12(2). <https://doi.org/10.3390/electronics12020469>
- Várkonyi, D., & Buza, K. (2019). *Extreme Learning Machines with Regularization for the Classification of Gene Expression Data*.
- Wang, J., Lu, S., Wang, S. H., & Zhang, Y. D. (2022). A review on extreme learning machine. *Multimedia Tools and Applications*, 81(29), 41611–41660. <https://doi.org/10.1007/s11042-021-11007-7>
- WHO. (2020). *Indonesia: Alzheimer/Dementia*. World Health Rankings. <https://www.worldlifeexpectancy.com/id/indonesia-alzheimers-dementia>
- Xie, L., Das, S. R., Wisse, L. E. M., Ittyerah, R., Flores, R. De, Shaw, L. M., Yushkevich, P. A., Wolk, D. A., & Initiative, A. D. N. (2023). Baseline structural MRI and plasma biomarkers predict longitudinal structural atrophy and cognitive decline in early Alzheimer's disease. *Alzheimer's Research & Therapy*, 15(1), 79. <https://doi.org/10.1186/s13195-023-01210-z>
- Yamanakkanavar, N., Choi, J. Y., & Lee, B. (2020). MRI segmentation and classification of human brain using deep learning for diagnosis of alzheimer's disease: A survey. *Sensors (Switzerland)*, 20(11), 1–31. <https://doi.org/10.3390/s20113243>
- Yu, Q., Miche, Y., Eirola, E., van Heeswijk, M., Séverin, E., & Lendasse, A. (2013). Regularized extreme learning machine for regression with missing data. *Neurocomputing*, 102, 45–51. <https://doi.org/https://doi.org/10.1016/j.neucom.2012.02.040>
- Zhao, Z., Chuah, J. H., Lai, K. W., Chow, C. O., Gochoo, M., Dhanalakshmi, S., Wang, N., Bao, W., & Wu, X. (2023). Conventional machine learning and deep learning in Alzheimer's disease diagnosis using neuroimaging: A review. *Frontiers in Computational Neuroscience*, 17. <https://doi.org/10.3389/fncom.2023.1038636>
- Zhong, Z., Jin, L., & Xie, Z. (2015). High Performance Offline Handwritten Chinese Character Recognition Using GoogLeNet and Directional Feature Maps. *International Conference on Document Analysis and Recognition (ICDAR)*, 846–850. <https://doi.org/10.1109/ICDAR.2015.73338881>

[Halaman ini sengaja dikosongkan]