

KLASIFIKASI PENYAKIT ALZHEIMER DARI SCAN MRI OTAK MENGUNAKAN CONVNEXT

Yehezkiel Stephanus Austin^{*1}, Haikal Irfano², Juan Young Christopher³, Lintang Cahyaning Sukma⁴,
Octo Perdana Putra⁵, Riyadh Ilham Ardhanto⁶, Novanto Yudistira⁷

^{1,2,3,4,5,6,7}Universitas Brawijaya, Malang

Email: ¹yehezkiel@student.ub.ac.id, ²haikalirfano@student.ub.ac.id, ³juanyoung1334@student.ub.ac.id,
⁴lintangcahya@student.ub.ac.id, ⁵octo_perdana@student.ub.ac.id, ⁶riyadhilham_@student.ub.ac.id,
⁷yudistira@ub.ac.id

^{*}Penulis Korespondensi

(Naskah masuk: 23 November 2023, diterima untuk diterbitkan: 19 November 2024)

Abstrak

Penyakit Alzheimer adalah gangguan *neurodegeneratif* yang menyebabkan penurunan fungsi kognitif yang signifikan. Penanganan penyakit ini dapat dilakukan melalui deteksi dini untuk meningkatkan kualitas kehidupan pasien melalui perawatan medis yang efisien dan tepat waktu. Teknologi machine learning dan neural network dapat mendukung deteksi dini melalui penggunaan model ConvNeXt yang telah dilatih dengan metode *transfer learning* menggunakan bobot awal dari ImageNet, dan di-fine-tune untuk mengklasifikasikan empat tingkat keparahan Alzheimer berdasarkan hasil pemindaian MRI otak, yaitu *Mild Demented*, *Moderate Demented*, *Non Demented*, dan *Very Mild Demented*. Penelitian ini akan menghasilkan model h5 dengan akurasi yang lebih baik daripada model lain sehingga dapat di-deploy pada aplikasi atau website untuk membantu deteksi dini klasifikasi tingkat keparahan Alzheimer.

Kata kunci: *ConvNeXt, Deteksi Dini, Fine-Tuning, ImageNet, Klasifikasi, Machine Learning, Model h5, Neural Network, Pemindaian MRI, Penyakit Alzheimer, Transfer Learning*

CLASSIFICATION OF ALZHEIMER'S DISEASE FROM BRAIN MRI SCANS USING CONVNEXT

Abstract

A Alzheimer's disease is a neurodegenerative disorder that causes significant cognitive decline. Early detection is crucial for managing this disease to improve patients' quality of life through efficient and timely medical care. Machine learning and neural network technology can support early detection through the use of the ConvNeXt model, which has been trained using transfer learning with initial weights from ImageNet and fine-tuned to classify four stages of Alzheimer's severity based on brain MRI scans: Non Demented, Very Mild Demented, Mild Demented, and Moderate Demented. This research will produce an h5 model with better accuracy than other models, enabling it to be deployed in applications or websites to assist in the early detection and classification of Alzheimer's severity.

Keywords: *ConvNeXt, Early Detection, Fine-Tuning, ImageNet, Machine Learning, Mild Demented, Moderate Demented, Neural Network, MRI Scans, Alzheimer's Disease, Transfer Learning*

1. PENDAHULUAN

Penyakit Alzheimer adalah penyakit gangguan otak yang umumnya mempengaruhi orang di atas 65 tahun. Penyakit ini pertama kali didefinisikan oleh Alois Alzheimer pada tahun 1906 di Kongres Psyatrist Jerman Selatan ke-37. Dalam kongres tersebut, Alzheimer mengungkapkan penemuannya tentang “*unusual disease of the cerebral cortex*” yang mengindikasikan adanya sebuah penyakit aneh pada seorang pasien bernama Auguste D. Penyakit ini

menyebabkan gejala kehilangan ingatan, disorientasi, dan halusinasi hingga kematian pasien (Hippius & Neundörfer, 2022)

Jumlah penderita penyakit Alzheimer di seluruh dunia meningkat dengan cepat. Hal ini dibuktikan dengan adanya 55 juta orang yang didiagnosis dengan demensia pada tahun 2020. Diprediksi, angka ini akan terus bertambah dari tahun ke tahun hingga menyentuh 78 juta pada tahun 2030 dan 139 juta pada tahun 2050. Di Indonesia, diperkirakan ada sekitar

1,2 juta orang dengan demensia pada tahun 2016, yang akan meningkat menjadi 2 juta orang pada tahun 2030 dan 4 juta orang pada tahun 2050 (Alzheimer's Indonesia, 2019). Estimasi total kerugian ekonomi yang disebabkan oleh demensia pada tahun 2015 menyentuh US\$818 juta dan diperkirakan akan meningkat menjadi US\$ 2,8 triliun pada tahun 2030 (Greenblat, 2023). Oleh karena banyaknya dampak yang disebabkan oleh penyakit ini, maka diperlukan pendekatan medis secara tradisional atau-pun modern.

Sejarah pendekatan medis secara tradisional terus berkembang seiring berjalannya waktu. Pada awalnya, melalui pengetahuan yang terbatas, fokus utama praktisi medis adalah meredakan gejala kehilangan ingatan dan disorientasi melalui obat-obatan dengan harapan untuk memperlambat perkembangan penyakit dan meningkatkan kualitas hidup pasien (Schachter & Davis, 2022). Perkembangan pendekatan medis lebih lanjut dapat dilihat pada penemuan dan penelitian mengenai hasil pemindaian otak melalui MRI yang dapat memberikan penglihatan mendalam terhadap perubahan struktural otak. Melalui hasil pemindaian MRI, praktisi medis dapat menganalisis, mengidentifikasi dan mendiagnosis pasien agar mendapatkan perawatan yang sesuai. Namun, seringkali terjadi ketidakakuratan dalam penilaian praktisi medis yang menyebabkan keterlambatan diagnosis pasien sehingga gejala penyakit Alzheimer semakin memburuk (Ferrari & Sorbi, 2021).

Pendekatan medis secara modern dapat diwujudkan dengan adanya perkembangan teknologi informasi dan kecerdasan buatan. Perkembangan teknologi ini memungkinkan *machine learning* untuk digunakan pada hasil pemindaian otak menggunakan MRI. Dengan *machine learning*, komputer dapat mempelajari pola-pola kompleks yang sebelumnya sulit untuk dipelajari oleh manusia (Oh et al., 2019). Melalui pendekatan ini, pendeteksian penyakit Alzheimer akan lebih akurat dan memberikan kesempatan bagi pasien untuk mendapatkan perawatan yang lebih baik.

Meskipun *machine learning* telah membantu praktisi medis mendeteksi penyakit Alzheimer, terdapat banyak metode *machine learning* yang tersedia. Setiap metode tersebut akan memiliki kemampuan yang berbeda-beda. Seperti yang disebutkan oleh teorema "*No Free lunch*", jika semua algoritma optimisasi dirata-ratakan pada semua masalah, maka semua algoritma akan memiliki kemampuan yang sama (Adam et al., 2019). Hal ini menunjukkan bahwa, seberapa efektif suatu algoritma bergantung dengan situasinya. Oleh karena itu, perlu penelitian lebih lanjut mengenai hal ini.

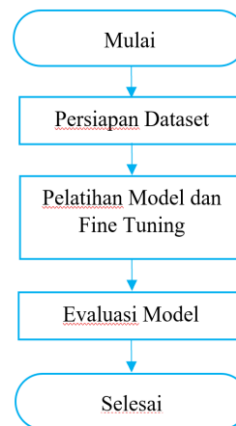
Salah satu metode machine learning yang dapat digunakan dalam sebuah klasifikasi adalah Convolutional Neural Network (CNN). CNN (Yann LeCun, 1980) adalah sebuah jenis jaringan saraf tiruan yang menggunakan beberapa lapisan konvolusi

untuk memproses informasi berdasarkan pola-pola inputnya. Algoritma CNN ini menjadi dasar bagi beberapa algoritma lainnya, seperti AlexNet (Krizhevsky, Sutskever, & Hinton, 2012), VGG (Karen Simonyan & Andrew Zisserman, 2014), dan ResNet (Kaiming He, Xiangyu Zhang, Shaoqing Ren, & Jian Sun, 2015). Seiring berjalannya waktu, algoritma ResNet yang menggunakan konsep CNN juga menjadi dasar dari algoritma ConvNext (Aditya Singh, 2022). Algoritma ConvNext memiliki akurasi yang tinggi dan layak untuk dijadikan sebagai sebuah metode dalam kasus penelitian.

Penelitian ini akan mengoptimalkan model ConvNeXt melalui transfer learning dari ImageNet dan fine-tuning pada data MRI otak, dengan tujuan untuk menghasilkan model klasifikasi yang lebih akurat untuk empat tingkatan keparahan penyakit Alzheimer. Model h5 sebagai hasil penelitian ini diharapkan dapat membuka jalan bagi pengembangan sistem pendukung keputusan yang dapat digunakan oleh praktisi medis dalam diagnosis lebih dini dan penanganan yang lebih efektif terhadap penyakit ini.

2. METODE PENELITIAN

Penelitian ini dilaksanakan dengan pendekatan metodologi yang sistematis untuk melatih dan mengevaluasi model arsitektur ConvNeXt. Penelitian akan diawali dengan persiapan *dataset*, pelatihan model, dan evaluasi model. Diagram alur penelitian dapat terlihat pada gambar 1.



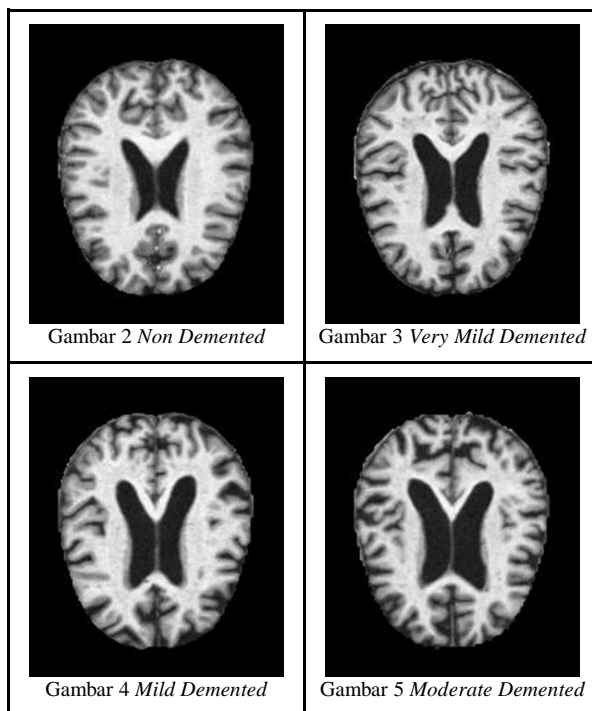
Gambar 1 Diagram Alur Penelitian

2.1. Persiapan Dataset

Dataset yang digunakan pada penelitian Klasifikasi Penyakit Alzheimer dari Scan MRI Otak Menggunakan ConvNeXt adalah "*Alzheimer's Dataset (4 class of images)*". *Dataset* ini diunduh dari <https://www.kaggle.com/datasets/tourist55/alzheimer-dataset-4-class-of-images> dengan lisensi dari Alzheimer's Disease Neuroimaging Initiative (ADNI). *Dataset* ini adalah *dataset* dengan gambar dari scan MRI otak manusia yang memiliki empat klasifikasi tingkat penyakit Alzheimer, yaitu *Mild Demented*, *Moderate Demented*, *Non Demented*, dan *Very Mild Demented*.

Dataset ini memiliki 6400 gambar yang dipisah menjadi empat folder klasifikasi, yaitu 3200 gambar *Non Demented*, 2240 gambar *Very Mild Demented*, 896 gambar *Mild Demented*, dan 64 gambar *Moderate Demented*. Sampel gambar dari keempat kelas alzheimer ini dapat dilihat pada gambar 2 sampai 5. Pembagian dataset ini dilakukan dengan rasio 60:20:20 untuk data latih (training), data evaluasi (validation), dan data uji (test). Berdasarkan jumlah rasio tersebut, maka jumlah gambar untuk masing-masing data adalah 3840 gambar data latih, 1289 gambar data evaluasi, dan 1280 gambar data uji.

Dataset juga akan dimasukkan ke dalam sebuah *image data generator* yang akan memproses gambar serta melakukan augmentasi agar model dapat belajar lebih efektif. Augmentasi yang akan dilakukan meliputi rotasi 90 derajat, *flip* kiri-kanan dan atas-bawah, penyesuaian kecerahan, serta kontras. Langkah-langkah ini diharapkan dapat membantu meningkatkan performa model dengan memperkaya variasi data yang tersedia.



2.2. Pelatihan Model dan Fine Tuning

Layer (type)	Output Shape	Param #
convnext_base (Functional)	(None, 1024)	87566464
dense_2 (Dense)	(None, 4)	4100
Total params: 87,570,564		
Trainable params: 87,570,564		
Non-trainable params: 0		

Gambar 6 Model ConvNeXt dengan Fine Tuning

Setelah persiapan *dataset* selesai, model ConvNeXt pada gambar 6 akan diinisiasi untuk

dilatih dengan *dataset* tersebut. Model akan diinisiasi dengan bantuan tensorflow keras yang menggunakan *base model* ConvNeXt dengan bobot dari imagenet, pengaturan gambar dengan ukuran 300x300 dan *max pooling*. *Layer base model* ini akan ditambahkan dengan *layer dense class_count* dengan fungsi aktivasi *softmax*. Model juga akan menggunakan Adam *optimizer* dengan *hyperparameter learning rate* yang ditentukan untuk mengurangi *error* saat proses pelatihan.

Selain inisiasi model, fungsi *callback* juga akan dibuat dengan *hyperparameter batch size* sebagai jumlah ukuran kelompok data untuk proses pelatihan, *hyperparameter epoch* sebagai satuan masuknya *dataset* ke dalam model, *hyperparameter patience* sebagai ambang batas epoch untuk menentukan *learning rate*, *hyperparameter stop patience* sebagai ambang batas *epoch* untuk memberhentikan proses pelatihan, *hyperparameter threshold* sebagai ambang batas akurasi, *hyperparameter factor* sebagai angka yang digunakan untuk mengurangi *learning rate*, *hyperparameter ask epoch* sebagai ambang batas *epoch* penentu diberhentikan nya proses pelatihan, serta *hyperparameter batch size* sebagai jumlah *batch* yang dilakukan selama proses pelatihan untuk setiap *epoch*.

Model yang sudah diinisiasi akan dilatih dengan *hyperparameter x* sebagai data latih, parameter *epoch* sebagai satuan masuknya *dataset* kedalam model, *hyperparameter verbose* sebagai penanda dicetaknya proses pelatihan tiap iterasi, *hyperparameter callbacks* sebagai fungsi tambahan yang akan dijalankan selama proses pelatihan, *hyperparameter validation_data* sebagai data validasi, *hyperparameter validation_steps* sebagai jumlah langkah validasi yang dilakukan, dan *hyperparameter shuffle* sebagai penanda diacaknya data pelatihan untuk setiap *epoch*-nya.

Tabel 1. Konfigurasi Hyperparameter

Hyperparameter	Nilai
<i>Learning rate</i>	0.0001
<i>Batch Size</i>	10
<i>Batches</i>	513
<i>Epoch</i>	30
<i>Ask Epoch</i>	10
<i>Patience</i>	1
<i>Stop Patience</i>	5
<i>Threshold</i>	0.9
<i>Factor</i>	0.5

2.3. Evaluasi Model

Proses evaluasi yang digunakan dalam penelitian ini akan menggunakan nilai *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) yang didapatkan dari *Confusion Matrix*. *Confusion Matrix* adalah sebuah matriks performa model yang menunjukkan jumlah data aktual yang diprediksi benar dan jumlah data aktual yang diprediksi salah untuk setiap kelasnya. TP adalah jumlah data aktual positif yang diprediksi positif oleh model, sedangkan TN adalah jumlah data

aktual negatif yang diprediksi negatif oleh model. FP adalah jumlah data aktual negatif yang diprediksi positif oleh model, sedangkan FN adalah jumlah data aktual positif yang diprediksi negatif oleh model. Keempat nilai ini akan digunakan untuk menghitung nilai *Precision*, *Recall*, *F1-Score* dan *Accuracy* dari model melalui rumus pada persamaan 1 sampai 4.

$$Precision = \frac{TP}{TP + FP}$$

Persamaan 1 Rumus *Precision*

$$Recall = \frac{TP}{TP + FN}$$

Persamaan 2 Rumus *Recall*

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

Persamaan 3 Rumus *F1 Score*

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

Persamaan 4 Rumus *Accuracy*

Nilai *Precision* adalah nilai yang mengukur proporsi TP dari semua kasus yang diprediksi positif, sedangkan nilai *Recall* adalah nilai yang mengukur proporsi TP dari semua kasus aktual positif. Selain itu, model akan memiliki nilai *F1 Score* sebagai nilai yang mengukur harmoni diantara *Precision* dan *Recall*. Model juga akan memiliki nilai *accuracy* sebagai nilai jumlah prediksi benar model dari seluruh prediksi model.

2.4. Perangkat Lunak dan Perangkat Keras

Bahasa pemrograman yang digunakan dalam penelitian ini adalah bahasa pemrograman python. Penelitian ini juga akan menggunakan *library* os, time, itertools, numpy, pandas, matplotlib, sklearn, dan tensorflow. *Library* os digunakan untuk mengakses direktori fail *dataset*. Selanjutnya, *library* time digunakan untuk menghitung waktu proses pelatihan model. Selain itu, *library* itertools digunakan untuk membantu proses iterasi data. *Library* numpy, pandas, matplotlib, dan sklearn juga akan digunakan dalam proses perhitungan evaluasi dari model yang sudah dilatih. Terakhir, *library* tensorflow akan digunakan untuk proses pembuatan, pelatihan, dan evaluasi model.

Proses penelitian akan dijalankan dengan menggunakan layanan komputasi awan dari Kaggle. Layanan ini memiliki kapasitas memori sebesar 73.1GB. Layanan ini juga akan menggunakan akselerasi GPU P100 yang memiliki RAM CPU sebesar 13GB dan memori GPU sebesar 15.9GB. Dengan menggunakan layanan dan perangkat keras tersebut, proses pelatihan dan validasi model terjadi selama 2 jam, 7 menit, dan 41.48 detik.

3. TINJAUAN PUSTAKA

3.1. Klasifikasi Alzheimer

Pemilihan metode klasifikasi dengan machine learning memiliki peranan penting dalam menangani permasalahan keterlambatan deteksi tingkat keparahan Alzheimer. Data dan metode klasifikasi yang tepat akan membantu praktisi medis mengidentifikasi pola-pola khas dari penyakit Alzheimer. Metode klasifikasi ini dapat digunakan terhadap data MRI, data Neuropsikologis, dan data Genetik.

3.1.1. Klasifikasi berdasarkan Data MRI

Data MRI (Magnetic Resonance Imaging) merupakan data yang paling umum yang digunakan dalam klasifikasi Alzheimer. Convolutional Neural Networks (CNN) adalah salah satu varian jaringan saraf tiruan yang sesuai untuk analisis citra medis seperti MRI otak. Hal ini berkaitan dengan cara CNN beroperasi, yakni cara yang menyerupai kemampuan manusia dalam mengenali pola dalam citra, melalui konvolusi pada citra masukan dengan menggunakan filter kecil yang bergeser secara bertahap melintasi citra. Proses ini memungkinkan CNN untuk mengekstrak fitur-fitur yang penting dalam konteks klasifikasi. Langkah awal dalam proses klasifikasi ini melibatkan penggunaan gambar MRI otak sebagai data masukan. Gambar ini kemudian mengalami tahap pra-pemrosesan. Tahapan pra-pemrosesan mencakup normalisasi intensitas, pemangkasan, serta pengurangan *noise* (Ali Nawaz et al., 2020). Gambar-gambar yang telah melalui tahap pra-pemrosesan berikutnya digunakan untuk melatih model CNN. Model ini terdiri dari beberapa lapisan konvolusi, lapisan penggabungan (pooling) untuk mengurangi dimensi gambar, dan lapisan-lapisan sepenuhnya terhubung untuk membuat keputusan.

Selama proses pelatihan, CNN mempelajari untuk mengenali ciri-ciri patologis atau perubahan struktural yang terkait dengan Alzheimer pada gambar MRI tersebut. Proses pembelajaran ini melibatkan penyesuaian bobot-bobot jaringan untuk meminimalkan kesalahan dalam klasifikasi terhadap data pelatihan. CNN mengenali pola-pola seperti plak amyloid, atrofi otak, atau perubahan volume pada wilayah tertentu yang berhubungan dengan Alzheimer. Setelah model CNN dilatih dengan cukup baik, model tersebut dapat digunakan untuk menganalisis gambar MRI otak yang belum pernah dilihat sebelumnya. Proses ini melibatkan pengambilan gambar masukan, dan model CNN akan menghasilkan output yang mengindikasikan apakah gambar tersebut menampilkan tanda-tanda penyakit Alzheimer. Hasil ini dapat berguna untuk diagnosis dini atau peningkatan evaluasi kesehatan pada pasien.

3.1.2. Klasifikasi berdasarkan Data Neuropsikologis

Klasifikasi Alzheimer berdasarkan data neuropsikologi dengan memanfaatkan algoritma

Machine Learning seperti Random Forest adalah pendekatan yang efektif dalam mengidentifikasi potensi risiko Alzheimer. Random Forest adalah algoritma Machine Learning yang terbukti efektif dalam mengklasifikasikan data yang kompleks, termasuk data neuropsikologi. Cara kerja algoritma ini melibatkan pembagian data menjadi sejumlah pohon keputusan atau *decision tree* (Tyler C. Hammond et al., 2020). Setiap pohon dibuat dengan metode bootstrapping, yang mengambil sampel acak dari data pelatihan dengan penggantian. Selanjutnya, algoritma ini memilih subset acak dari fitur-fitur yang tersedia di setiap node keputusan dalam masing-masing pohon. Ketika data tes neuropsikologi baru diperkenalkan sebagai input ke model Random Forest, model ini akan memproses data melalui setiap pohon keputusan secara terpisah. Setiap pohon memberikan prediksi kelas yang berbeda. Hasil dari seluruh pohon keputusan digabungkan, dan kelas dengan dukungan mayoritas akan menjadi hasil prediksi akhir dari model.

Langkah awal dalam prosedur ini melibatkan pengumpulan data dari tes neuropsikologi, yang melibatkan berbagai jenis tes kognitif, seperti tes kecerdasan, tes memori, serta evaluasi kemampuan bahasa dan fungsi kognitif lainnya. Data ini umumnya mencakup hasil skor tes dari individu yang telah menjalani tes tersebut. Setelah data neuropsikologi terkumpul, langkah berikutnya adalah tahap pra-pemrosesan data. Pra-pemrosesan melibatkan normalisasi data untuk mengatasi potensi bias atau ketidaksetaraan dalam rentang hasil tes. Data yang telah melalui pra-pemrosesan kemudian digunakan sebagai *input* dalam model *machine learning* berbasis Random Forest. Selama proses pelatihan, Random Forest belajar untuk mengenali pola dan korelasi yang terdapat dalam data neuropsikologi yang berkaitan dengan risiko Alzheimer. Dengan demikian, algoritma ini dapat memberikan prediksi yang akurat terkait dengan apakah individu tersebut memiliki risiko terkena Alzheimer atau tidak.

3.1.2. Klasifikasi berdasarkan Data Genetik

Klasifikasi berdasarkan data genetik dengan memanfaatkan algoritma Machine Learning Support Vector Machines (SVM) adalah pendekatan yang juga umum digunakan. Langkah awal dalam proses ini adalah pengumpulan data genetik, yang mungkin termasuk *Single Nucleotide Polymorphism* (SNP) dan variasi genetik lainnya yang telah diidentifikasi sebagai potensi penyebab Alzheimer (Xianglian Meng et al., 2022). Setelah data genetik terkumpul, langkah selanjutnya adalah pra-pemrosesan data genetik. Pra-pemrosesan ini mencakup normalisasi data untuk memastikan bahwa berbagai variabel genetik memiliki tingkat kontribusi yang setara dalam analisis klasifikasi. Data yang telah di-preproses ini kemudian digunakan sebagai input dalam model Machine Learning berbasis SVM.

SVM adalah algoritma Machine Learning yang unggul dalam memisahkan kelas pada data yang kompleks seperti data genetik. Cara kerja algoritma SVM melibatkan penciptaan *hyperplane* yang memisahkan dua kelas dengan jarak maksimal antara mereka (Zhuo Sun et al., 2018). *Hyperplane* ini dipilih untuk memaksimalkan margin antara dua kelas, yang menghasilkan prediksi yang sangat akurat. Selama proses pelatihan, SVM belajar untuk mengidentifikasi pola genetik yang berkorelasi dengan risiko Alzheimer. Ini melibatkan penyesuaian parameter SVM agar *hyperplane* yang digunakan dapat secara efektif memisahkan kelas yang mewakili individu yang berisiko tinggi dan individu yang berisiko rendah terhadap Alzheimer berdasarkan data genetik mereka. Ketika data genetik individu baru diperkenalkan sebagai *input* ke model SVM yang telah dilatih, algoritma ini akan menjalankan data melalui *hyperplane* yang telah dihasilkan selama pelatihan. SVM kemudian memberikan prediksi tentang apakah individu tersebut memiliki faktor risiko genetik yang tinggi atau rendah terhadap Alzheimer. Proses ini memungkinkan peneliti untuk mengidentifikasi faktor risiko genetik secara lebih akurat dan memahami hubungan antara polimorfisme genetik tertentu dengan perkembangan Alzheimer.

3.2. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) adalah suatu metode algoritma machine learning yang terinspirasi oleh struktur dan fungsi otak manusia. CNN terdiri atas tiga bagian utama, yaitu *convolution layer*, *pooling layer*, dan, *fully connected layer*. Lapisan konvolusi merupakan bagian yang mengidentifikasi pola pada input dengan menggunakan filter. Lapisan *pooling* berfungsi untuk mengurangi dimensi spasial dari fitur dengan *down-sampling*. Terakhir, yaitu lapisan *fully connected* adalah lapisan yang digunakan untuk menyambungkan setiap saraf dari lapisan sebelumnya ke lapisan saat ini (Yamashita et al., 2018).

Contoh penelitian dengan implementasi metode CNN yaitu "Implementasi Metode CNN untuk Klasifikasi Gambar Jamur Pada Analisis Pengolahan Gambar", penggunaan CNN untuk mengklasifikasikan gambar jamur *Agaricus* dan *Amanita* dilakukan dalam dua tahap utama, yaitu pembelajaran karakteristik yang terdiri dari 2 tahap konvolusi dan 2 tahap *pooling*. Kemudian, dilanjutkan dengan tahap klasifikasi dimana arsitektur terbaik didapatkan dengan membandingkan beberapa parameter yakni *epoch* dan *optimizer*. Diperoleh perbandingan akurasi dari 3 jenis *optimizer* yaitu 62% untuk *optimizer* Adam dengan *epoch* 100, 54% untuk *optimizer* RMSProp dengan *epoch* 50, dan 55% untuk *optimizer* SGD dengan *epoch* 50. Hasil akurasi yang diperoleh untuk mengklasifikasikan gambar jamur menggunakan CNN yakni sebesar 62% dengan menggunakan skenario perbandingan data

train validation 80% : 20%, ukuran kernel 3x3, *optimizer* Adam, 100 *epoch*, dan *learning rate* sebesar 0,001 (Oktavia Putri, O. N., 2020).

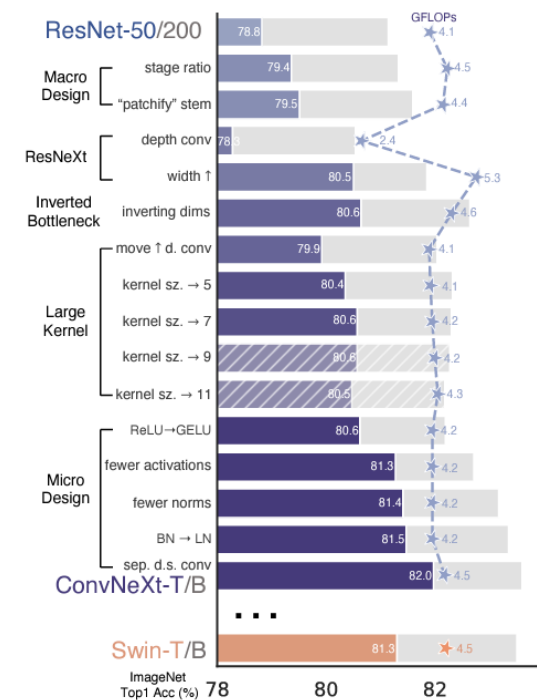
“Implementasi Convolutional Neural Network (CNN) Untuk Klasifikasi Batik Tanah Liat Sumatera Barat”, Metode *neural network convolutional* berhasil digunakan untuk mengklasifikasikan gambar batik tanah liat Sumatera Barat. Metode ini menggunakan library keras dan tensorflow dengan bahasa pemrograman python. Pada hasil pelatihan data *train* didapat akurasi sebesar 98.75%, sedangkan tahap pelatihan data uji didapat akurasi sebesar 62.5% (Azmi, K., Defit, S., Sumijan., 2023). Tingkat akurasi ini cukup baik, dan sudah layak digunakan sebagai referensi untuk membangun aplikasi yang nyata untuk mengidentifikasi motif batik atau objek lainnya secara umum. Kemiripan warna, kemiripan motif, ukuran jarak pengambilan gambar, kualitas cahaya, dan posisi gambar menyebabkan beberapa gambar diklasifikasikan kurang tepat karena hampir setiap daerah di Indonesia memiliki motif batik daerahnya sendiri.

“Klasifikasi Tutupan Lahan Melalui Citra Satelit SPOT-6 dengan Metode Convolutional Neural Network (CNN)”, Dalam penelitian ini, sebuah sistem telah dirancang untuk mengkategorikan lima jenis tutupan lahan: hutan, bukit gundul, sawah, pemukiman, dan sungai. Data primer, yang dikumpulkan melalui satelit SPOT 6, digunakan sebagai data uji. Sebanyak 350 gambar tutupan lahan, atau data latih dan data uji, digunakan sebagai data uji dalam penelitian ini. Model CNN yang diusulkan untuk penelitian ini terdiri dari 3 hidden layer, 1 *fully connected layer*, sebuah filter ukuran 3x3 dengan output channel secara berurutan 8,16 dan 32, serta aktivasi *softmax*. Berdasarkan hasil pengujian sistem yang telah dilakukan, model CNN yang diusung oleh peneliti mampu untuk mengklasifikasikan 5 kelas tutupan lahan dengan performansi terbaik yang dihasilkan yaitu akurasi 95.45%, *loss* 0.2457, *precision* 0.92, *recall* 0.92 dan *f-1 score* 0.92 (Magdalena, R., Saidah, S., Pratiwi, N, K, C, P., Putra, A, T., 2021).

3.3. ConvNeXt

Salah satu model CNN untuk melakukan klasifikasi adalah ConvNeXt. ConvNeXt adalah model ConvNet dengan dasar arsitektur ResNet yang mengalami modernisasi dengan konsep hierarchical vision transformer (Swin) (Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, & Zheng Zhang, 2021). ResNet akan dilatih dengan teknik yang sama seperti *vision transformer* dengan AdamW *Optimizer*, *Epochs*, *Heavy Data Augmentation*, dan *Regulation*. Pada gambar 7 dapat terlihat bahwa *macro design* model dimodifikasi dengan *stage ratio* yang menyesuaikan jumlah blok pada setiap stage untuk meningkatkan akurasi dan *patchify stem* yang membuat *sliding windows* bersifat seperti *vision transformer*. Konsep *depth conv* dan *width* milik

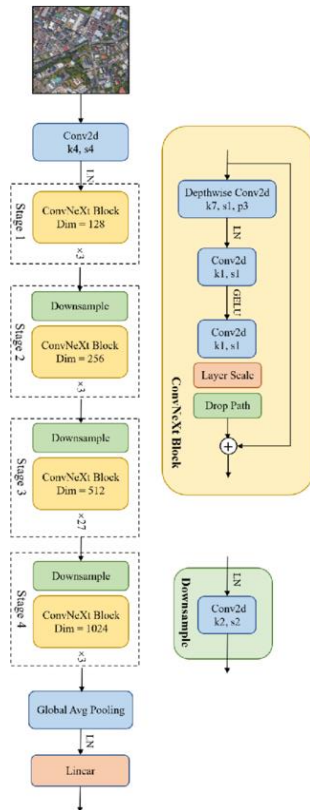
ResNeXt akan digunakan untuk meningkatkan akurasi dari model. Selain itu, akurasi model akan ditingkatkan dengan *inverted bottleneck* serta mengubah ukuran kernel yang digunakan. ConvNeXt juga akan menggunakan *micro design* model yang dimodifikasi melalui *Rectified Linear Unit* (ReLU) yang diganti dengan *Gaussian Linear Unit* (GeLU), penggunaan fungsi aktivasi yang lebih jarang, penggunaan lapisan normalisasi yang lebih sedikit dan *batch normalization* yang diganti dengan *layer normalization* (Liu et al., 2022).



Gambar 7 Desain ConvNeXt (Liu et al., 2022)

Arsitektur ConvNeXt yang terlihat pada gambar 8 terdiri dari Conv2d, 4 *stage*, *global average pooling*, dan *linear*. Stage 1 terdiri dari ConvNeXt Block sedangkan stage 2 sampai 4 terdiri dari Downsample dan ConvNeXt Block. ConvNeXt Block terdiri dari Depthwise Conv2d, 2 Conv2d, Layer Scale dan DropPath. Di sisi lain, Downsample terdiri dari Conv2d.

Contoh penelitian studi kasus dengan ConvNeXt yaitu penelitian untuk klasifikasi *Fine-Grained* dalam jurnal “ConvNeXt-Based Fine-Grained Image Classification and Bilinear Attention Mechanism Model”. Penelitian ini dilakukan menggunakan tiga dataset klasifikasi *Fine-Grained* yang diakui secara internasional, yaitu CUB200-2011, FGVC-Aircraft, dan Stanford Cars. Hasil dari studi ini menunjukkan bahwa model ConvNeXt dapat berhasil digunakan dalam tugas klasifikasi *Fine-Grained* dan memberikan hasil yang lebih baik dibandingkan dengan jaringan klasifikasi tradisional (Liu et al., 2022).



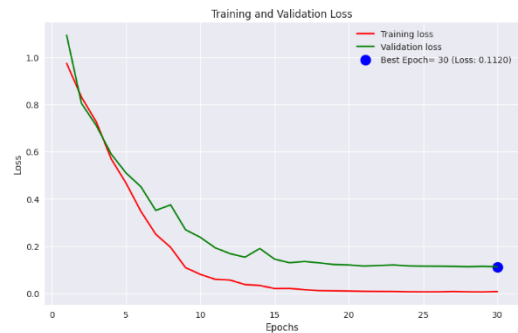
Gambar 8 Arsitektur ConvNeXt (Chen et al., 2023)

Contoh lain penelitian studi kasus lainnya dilakukan oleh Zhang et al. Dalam penelitian ini, mereka mengusulkan sebuah model deteksi objek berbasis ConvNeXt yang disebut AFCNet. Tujuan dari penelitian ini adalah untuk meningkatkan kinerja deteksi objek dengan mengatasi keterbatasan detektor berbasis *anchor*. Dari penelitian tersebut dapat disimpulkan bahwa AFCNet, yang merupakan model deteksi objek berbasis ConvNeXt, berhasil meningkatkan kinerja deteksi objek pada *dataset* IIOPE dan PASCAL VOC 2007. AFCNet mengungguli detektor-detektor lain seperti Faster R-CNN, RetinaNet, FCOS, ATSS, dan TOOD dengan perbedaan masing-masing sebesar 10,6%, 15,8%, 22,3%, 15,6%, dan 16,4%. Dengan demikian, studi ini menunjukkan bahwa model deteksi objek berbasis ConvNeXt seperti AFCNet memiliki potensi dalam meningkatkan kinerja deteksi objek. Dengan memanfaatkan keunggulan ConvNeXt dan menggabungkannya dengan peningkatan-peningkatan lainnya, AFCNet berhasil mencapai hasil yang lebih baik dalam mendeteksi objek, baik pada *dataset* gambar inframerah maupun *dataset* umum (Zhang et al., 2023).

4. HASIL DAN DISKUSI

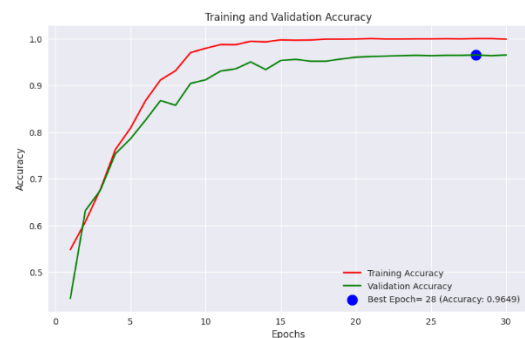
Hasil pelatihan dan validasi model akan memiliki grafik *training validation loss*, *training validation accuracy*, *training validation recall*, *training validation precision*, dan *training validation f1-score*. Grafik *training validation loss* adalah grafik yang menunjukkan tingkat *error* pada model saat

melakukan proses pelatihan dan validasi pada setiap *epoch*-nya, sedangkan grafik *training validation accuracy* adalah grafik yang menunjukkan tingkat akurasi model saat melakukan proses pelatihan dan validasi pada setiap *epoch*-nya. Grafik *training validation recall* adalah grafik yang menunjukkan nilai *recall* model saat melakukan proses pelatihan dan validasi pada setiap *epoch*-nya, sedangkan grafik *training validation precision* adalah grafik yang menunjukkan nilai *presisi* model saat melakukan proses pelatihan dan validasi pada setiap *epoch*-nya. Grafik *training validation f1-score* adalah grafik yang menunjukkan nilai *f1* model saat melakukan proses pelatihan dan validasi pada setiap *epoch*-nya.



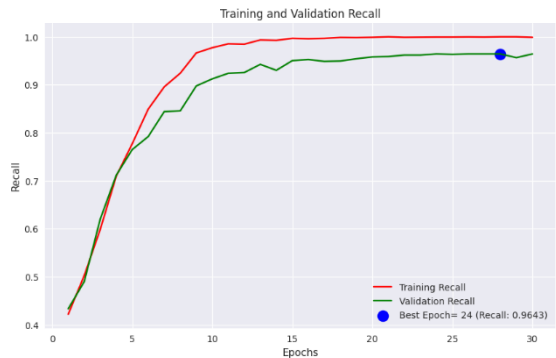
Gambar 9 Hasil loss pelatihan dan validasi model

Berdasarkan grafik *training and validation loss* pada gambar 9, dapat dilihat bahwa hasil *error* proses pelatihan dan validasi model mengalami penurunan. Hal ini menunjukkan bahwa model berhasil mengalami peningkatan proses pembelajaran saat proses pelatihan dan validasi. *Epoch* terbaik model saat melakukan proses validasi adalah *epoch* 30 dengan *error* sebesar 0.1120.

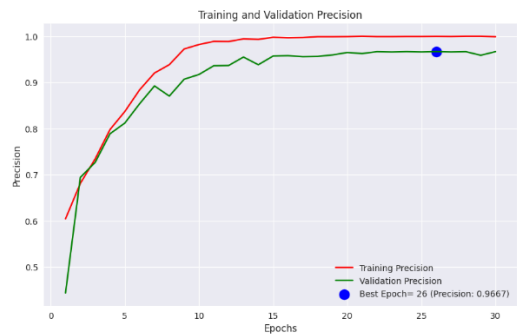


Gambar 10 Hasil akurasi pelatihan dan validasi model

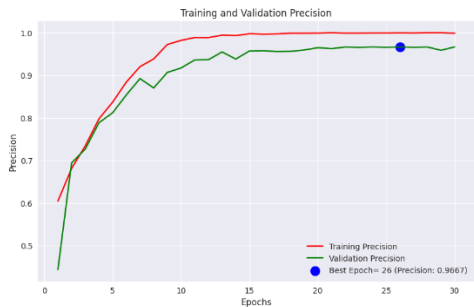
Selanjutnya berdasarkan grafik *training and validation accuracy* pada gambar 10, dapat dilihat bahwa hasil akurasi proses pelatihan dan validasi mengalami peningkatan yang signifikan. Hal ini menunjukkan bahwa model memiliki akurasi yang tinggi saat proses pelatihan dan validasi. *Epoch* terbaik model saat melakukan proses validasi adalah *epoch* 24 dengan akurasi sebesar 0.9649.



Gambar 11 Hasil recall pelatihan dan validasi model



Gambar 12 Hasil precision pelatihan dan validasi model

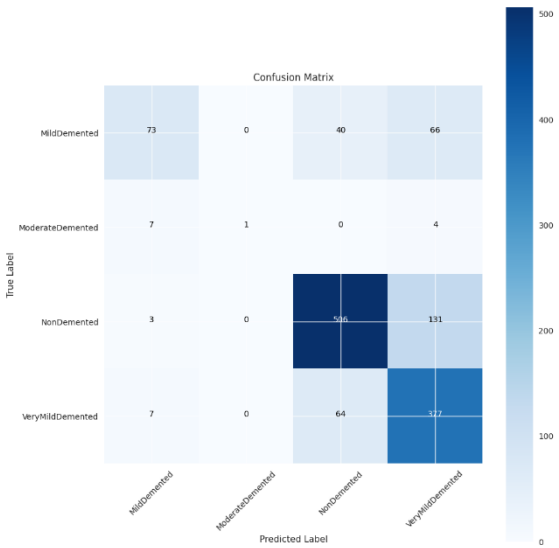


Gambar 13 Hasil f1-score pelatihan dan validasi model

Selain itu berdasarkan grafik training and validation recall, precision, dan f1 pada gambar 11 sampai 13, dapat dilihat bahwa ketiga nilai tersebut mengalami peningkatan pada proses pelatihan dan validasi. Hal ini menunjukkan bahwa model memiliki nilai recall, precision, dan f1 yang tinggi saat proses pelatihan dan validasi. Recall terbaik model saat proses validasi didapatkan saat *epoch* ke 24 dengan nilai 0.9643, sedangkan precision terbaik model saat proses validasi didapatkan saat *epoch* ke 26 dengan nilai 0.9667. Selanjutnya dapat dilihat juga bahwa nilai F1 terbaik model saat proses validasi didapatkan saat *epoch* ke 26 dengan nilai 0.9654.

Selain grafik pelatihan dan validasi, model juga akan melakukan proses uji coba yang menghasilkan confusion matrix. Berdasarkan confusion matrix pada gambar 14 dapat dilihat jumlah kebenaran prediksi untuk setiap kelasnya. Kelas *MildDemented* mendapatkan 73 prediksi benar dan 106 prediksi salah. Kelas *ModerateDemented* mendapatkan 1 prediksi benar dan 11 prediksi salah. Kelas *NonDemented* mendapatkan 506 prediksi benar dan 134 prediksi salah. Kelas *VeryMildDemented*

mendapatkan 377 prediksi benar dan 71 prediksi salah.



Gambar 14 Hasil uji coba model

Melalui visualisasi confusion matrix, akan didapatkan nilai *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Keempat nilai ini akan digunakan untuk menghitung nilai *Precision*, *Recall*, dan *F1-Score* dari model untuk setiap kelas yang ada. Nilai evaluasi tersebut dapat dilihat pada tabel 2.

Tabel 2. Hasil Precision, Recall, dan F1

Class	P	R	F1
Mild	0.81	0.41	0.54
Demented	1.00	0.08	0.15
Moderate	0.83	0.79	0.81
Demented	0.65	0.84	0.73
Non			
Demented			
Very Mild			
Demented			

Proses *testing* juga akan menghasilkan nilai akurasi dari model yang digunakan. Nilai akurasi yang didapatkan oleh model ConvNeXt adalah 0.75. Perbandingan akurasi model ConvNeXt dengan model lain dapat dilihat pada tabel 3. Pada tabel tersebut dapat dilihat bahwa model ConvNeXt memiliki akurasi yang lebih baik dibandingkan dengan model lainnya.

Tabel 3. Perbandingan Akurasi Model

Model	Accuracy
ConvNeXt	0.75
EfficientNet b3	0.73
ResNet50	0.68
MobileNet	0.68

Penelitian dilanjutkan dengan membandingkan kembali model ConvNeXt yang digunakan dengan optimizer berbeda untuk mengurangi *error* yang terjadi saat proses pelatihan dan validasi. Optimizer yang dapat digunakan antara lain adalah Adam Optimizer, Adamax Optimizer, dan AdamW Optimizer. Tabel perbandingan model ConvNeXt

dengan penggunaan optimizer berbeda dapat dilihat pada tabel 4.

Tabel 4. Perbandingan Akurasi Model

Model	Accuracy
ConvNeXt (Adam)	0.79
ConvNeXt (Adamax)	0.75
ConvNeXt (AdamW)	0.79
ConvNeXt (SGD)	0.56

5. KESIMPULAN

Tujuan dari penelitian ini adalah untuk melihat seberapa baik ConvNeXt untuk mengklasifikasikan penyakit Alzheimer. ConvNeXt merupakan model dari ConvNet dengan dasar arsitektur ResNet yang mengalami modernisasi dengan konsep hierarchical vision transformer (Swin). Meskipun menurut paper “*A ConvNet for 2020*” hasil akurasi dari ConvNeXt cukup baik dalam hal klasifikasi gambar, namun hasil yang diperoleh dalam penelitian ini belum memenuhi ekspektasi yang diharapkan.

Penelitian ini menghadapi tantangan karena model ConvNeXt yang menggunakan memori yang besar sehingga proses pelatihan model sangat lama. Kurangnya data dalam beberapa label juga menyebabkan terjadinya *imbalance* dataset, sehingga model kurang belajar mengenai label-label tersebut. Selain itu, dugaan adanya kesalahan selama proses pelatihan dan pengujian model juga memunculkan tantangan yang signifikan dalam pencapaian akurasi yang diinginkan. Akurasi proses uji coba model yang sudah dilatih dalam penelitian ini adalah sebesar 0.75. Akurasi yang dicapai ini tidaklah cukup tinggi untuk menjadikan ConvNeXt sebagai algoritma yang cocok untuk diagnosis penyakit Alzheimer, namun sudah cukup tinggi dibandingkan dengan model lainnya seperti EfficientNet b3, ResNet 50, dan MobileNet.

Namun penting untuk dicatat bahwa akurasi yang dicapai dapat bervariasi tergantung pada kumpulan data yang digunakan dan modifikasi arsitektur ConvNeXt yang digunakan. Meskipun demikian, hasil yang diperoleh dari penelitian ini dapat dinilai kurang memuaskan sehingga ConvNeXt tidak dapat disarankan untuk menjadi algoritma yang berguna untuk klasifikasi diagnosis penyakit Alzheimer menggunakan data scan MRI otak. Oleh karena itu, penelitian lebih lanjut diperlukan, baik melalui penambahan dataset untuk mencegah distribusi yang tidak seimbang, penambahan augmentasi gambar agar model dapat belajar lebih efektif, maupun eksplorasi modifikasi arsitektur model ConvNeXt agar lebih sesuai untuk tugas klasifikasi ini.

DAFTAR PUSTAKA

ADAM, S. P., ALEXANDROPOULOS, S. A. N., PARDALOS, P. M., & VRAHATIS, M. N. 2019. No free lunch theorem: A review.

Springer Optimization and Its Applications, 145, 57–82. https://doi.org/10.1007/978-3-030-12767-1_5/COVER

ADITYA SINGH. 2022, February. *ConvNext: The Return Of Convolution Networks*. <https://medium.com/augmented-startups/convnext-the-return-of-convolution-networks-e70cbe8dabcc>

Alzheimer's Indonesia. 2019. *Statistik tentang Demensia - Alzheimer Indonesia*. Last Accessed: 14 September 2023. <https://alzi.or.id/statistik-tentang-demensia/>

FERRARI, C., & SORBI, S. 2021. The Complexity Of Alzheimer's Disease: An Evolving Puzzle. *Physiological Reviews*, 101(3), 1047–1081. https://doi.org/10.1152/PHYSREV.00015.2020/ASSET/IMAGES/LARGE/PHYSREV.00015.2020_F009.JPEG

GREENBLAT, C. 2023. *Dementia*. Last Accessed: 14 September, 2023. <https://www.who.int/news-room/fact-sheets/detail/dementia>

HIPPIUS, H., & NEUNDÖRFER, G. 2022. The discovery of Alzheimer's disease. *https://Doi.Org/10.31887/DCNS.2003.5.1/Hhippius*, 5(1), 101–108. <https://doi.org/10.31887/DCNS.2003.5.1/HHIPPIUS>

LI, Z., GU, T., LI, B., XU, W., HE, X., & HUI, X. 2022. ConvNeXt-Based Fine-Grained Image Classification and Bilinear Attention Mechanism Model. *Applied Sciences (Switzerland)*, 12(18). <https://doi.org/10.3390/app12189016>

LIU, Z., MAO, H., WU, C. Y., FEICHTENHOFER, C., DARRELL, T., & XIE, S. 2022. A ConvNet for the 2020s. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2022-June*, 11966–11976. <https://doi.org/10.1109/CVPR52688.2022.01167>

OH, K., CHUNG, Y. C., KIM, K. W., KIM, W. S., & OH, I. S. 2019. Classification and Visualization of Alzheimer's Disease using Volumetric Convolutional Neural Network and Transfer Learning. *Scientific Reports 2019 9:1*, 9(1), 1–16. <https://doi.org/10.1038/s41598-019-54548-6>

SCHACHTER, A. S., & DAVIS, K. L. 2022. Alzheimer's disease. *https://Doi.Org/10.31887/DCNS.2000.2.2/Asschachter*, 2(2), 91–100. <https://doi.org/10.31887/DCNS.2000.2.2/ASSCHACHTER>

YAMASHITA, R., NISHIO, M., DO, R. K. G., & TOGASHI, K. 2018. Convolutional neural networks: an overview and application in radiology. *Insights into Imaging*, 9(4), 611–629. <https://doi.org/10.1007/S13244-018-0639-9/FIGURES/15>

- OKTAVIA PUTRI, O, N. 2020. Implementasi Metode CNN Dalam Klasifikasi Gambar Jamur Pada Analisis Image Processing. <https://dspace.uui.ac.id/bitstream/handle/123456789/23677/16611103%20Oktavia%20Nurima%20Putri.pdf?sequence=1&isAllowed=y>
- AZMI, K., DEFIT, S., SUMIJAN. 2023. implementasi Convolutional Neural Network (CNN) Untuk Klasifikasi Batik Tanah Liat Sumatera Barat. <https://ejurnal.sttdumai.ac.id/index.php/unitek/article/download/504/350>
- MAGDALENA, R., SAIDAH, S., PRATIWI, N, K, C, P., PUTRA, A, T. 2021. Klasifikasi Tutupan Lahan Melalui Citra Satelit SPOT-6 dengan Metode Convolutional Neural Network (CNN). <https://jurnal.untan.ac.id/index.php/jepin/article/download/48195/75676591380>
- ZHANG, Y., XU, A., LAN, D., ZHANG, X., YIN, J., & GOH, H. H. 2023. ConvNeXt-based anchor-free object detection model for infrared image of power equipment. *Energy Reports*, 9, 1121–1132. <https://doi.org/10.1016/j.egyr.2023.04.145>
- KUNDARAM, S., & PATHAK, K., 2020. Deep Learning-Based Alzheimer Disease Detection. *Lecture Notes in Electrical Engineering*. https://doi.org/10.1007/978-981-15-5546-6_50
- HAMMOND, T., XING, X., WANG, C., MA, D., NHO, K., CRANE, P., ELAHI, F., ZIEGLER, D., LIANG, G., CHENG, Q., YANCKELLO, L., JACOBS, N., & LIN, A., 2020. β -amyloid and tau drive early Alzheimer's disease decline while glucose hypometabolism drives late decline. *Communications Biology*, 3. <https://doi.org/10.1038/s42003-020-1079-x>.
- XU, L., LIANG, G., LIAO, C., CHEN, G., & CHANG, C., 2019. k-Skip-n-Gram-RF: A Random Forest Based Method for Alzheimer's Disease Protein Identification. *Frontiers in Genetics*, 10. <https://doi.org/10.3389/fgene.2019.00033>.
- MENG, X., WEI, Q., MENG, L., LIU, J., WU, Y., & LIU, W., 2022. Feature Fusion and Detection in Alzheimer's Disease Using a Novel Genetic Multi-Kernel SVM Based on MRI Imaging and Gene Data. *Genes*, 13. <https://doi.org/10.3390/genes13050837>.
- SUN, Z., QIAO, Y., LELIEVELDT, B., STARING, M., & INITIATIVE, A., 2018. Integrating spatial-anatomical regularization and structure sparsity into SVM: Improving interpretation of Alzheimer's disease classification. *NeuroImage*, 178, pp. 445-460. <https://doi.org/10.1016/j.neuroimage.2018.05.015>.
- NAWAZ, A., ANWAR, S., LIAQAT, R., IQBAL, J., BAGCI, U., & MAJID, M., 2020. Deep Convolutional Neural Network based Classification of Alzheimer's Disease using MRI Data. 2020 IEEE 23rd International Multitopic Conference (INMIC), pp. 1-6. <https://doi.org/10.1109/INMIC50486.2020.9318172>.
- CHEN, S., OGAWA, Y., & SEKIMOTO, Y. 2023. Large-scale individual building extraction from open-source satellite imagery via super-resolution-based instance segmentation approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 195, 129–152. <https://doi.org/10.1016/j.isprsjprs.2022.11.006>