



UIN SUSKA RIAU

© Hak cipta milik UIN Suska Riau



nic University of Sultan Syarif Kasim Riau

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

KLASIFIKASI CITRA STROKE MENGGUNAKAN AUGMENTASI DAN CONVOLUTIONAL NEURAL NETWORK EFFICIENTNET-B0

TUGAS AKHIR

Disusun Sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Teknik
Pada Jurusan Teknik Informatika

Oleh

NADILA HANDAYANI PUTRI

NIM. 11950120086



**FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU
PEKANBARU
2023**



Ha



1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERSETUJUAN

KLASIFIKASI CITRA STROKE MENGGUNAKAN AUGMENTASI DAN CONVOLUTIONAL NEURAL NETWORK EFFICIENTNET-B0

TUGAS AKHIR

Oleh

NADILA HANDAYANI PUTRI

NIM. 11950120086

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir
di Pekanbaru, pada tanggal 11 Mei 2023

Pembimbing I,

Jasril, S.Si., M.Sc

NIP. 19710215 200003 1 002

Pembimbing II,

Muhammad Irsyad, S.T., M.T

NIP. 19780508 200710 1 007

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PENGESAHAN

KLASIFIKASI CITRA STROKE MENGGUNAKAN AUGMENTASI DAN CONVOLUTIONAL NEURAL NETWORK EFFICIENTNET-B0

Oleh

NADILA HANDAYANI PUTRI

NIM. 11950120086

Telah dipertahankan di depan sidang dewan pengaji
sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik
pada Universitas Islam Negeri Sultan Syarif Kasim Riau

Pekanbaru, 11 Mei 2023

Mengesahkan,

Ketua Jurusan,

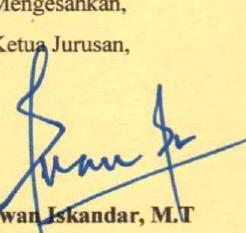
Iwan Iskandar, M.T

NIP. 19821216 201503 1 003



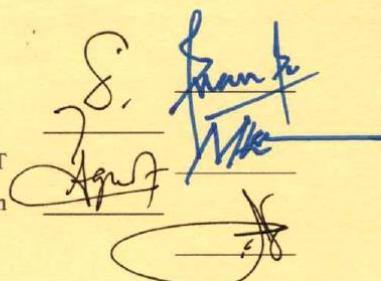
Dr. Hartono, M.Pd

NIP. 19640301 199203 1 003



DEWAN PENGUJI

Ketua	:	Iwan Iskandar, M.T
Pembimbing I	:	Jasril, S.Si., M.Sc
Pembimbing II	:	Muhammad Irsyad, S.T., M.T
Pengaji I	:	Surya Agustian, S.T., M.Kom
Pengaji II	:	Febi Yanto, M.Kom





LEMBAR HAK ATAS KEKAYAAN INTELEKTUAL

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

© Hak cipta milik UIN SUSKA Riau

State Islamic University of Sultan Syarif Kasim Riau

UIN SUSKA RIAU



LEMBAR PERNYATAAN

© Hak Cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan didalam daftar pustaka.

Pekanbaru, 17 Mei 2023

Yang membuat pernyataan,

NADILA HANDAYANI PUTRI

NIM 11950120086

UIN SUSKA RIAU



© Hak cipta milik UIN Suska Riau

State Islamic University of Sultan Syarif Kasim Riau

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

SURAT PERNYATAAN

Saya yang bertandatangan di bawah ini :

Nama : Nadila Handayani Putri
NIM : 11950120086
Tempat/Tgl.Lahir : Pekanbaru, 28 Februari 2002
Fakultas : Sains dan Teknologi
Prodi : Teknik Informatika
Judul Skripsi : Klasifikasi Citra Stroke Menggunakan Augmentasi dan Convolutional Neural Network EfficientNet-B0

Menyatakan dengan sebenar-benarnya bahwa :

1. Penulisan skripsi dengan judul sebagaimana tersebut di atas adalah hasil pemikiran dan penelitian saya sendiri
2. Semua kutipan pada karya tulis saya ini sudah disebutkan sumbernya
3. Oleh karena itu Skripsi saya ini, saya nyatakan bebas dari plagiat
4. Apabila dikemudian hari terbukti terdapat plagiat dalam penulisan skripsi saya tersebut, maka saya bersedia menerima sanksi sesuai peraturan perundang-undangan.

Demikianlah surat pernyataan ini saya buat dengan penuh kesadaran dan tanpa paksaan dari pihak manapun juga

Pekanbaru, 16 Mei 2023
Yang membuat pernyataan

Nadila Handayani Putri
NIM. 11950120086





LEMBAR PERSEMPAHAN

Terima kasih kepada Allah SWT yang melimpahkan segala rahmatnya sehingga skripsi ini bisa terselesaikan dengan baik. Skripsi ini saya persembahkan untuk yang tersayang papa, mama, adik, dan teman-teman terdekat.

Semoga Tugas Akhir ini bisa bermanfaat bagi pembaca. Amiin.

Hak Cipta Dilindungi Undang-Undang

© Hak cipta milik Universitas Islam Sultan Syarif Kasim Riau

Klasifikasi Citra Stroke Menggunakan Augmentasi dan Convolutional Neural Network EfficientNet-B0

Nadila Handayani Putri, Jasril*, Muhammad Irsyad, Surya Agustian, Febi Yanto

Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru,
Indonesia

Email: ¹nadilahandyni@gmail.com, ^{2,*}jasril@uin-suska.ac.id, ³irsyadtech@uin-suska.ac.id,

⁴surya.agustian@uin-suska.ac.id, ⁵febiyanto@uin-suska.ac.id

Email Penulis Korespondensi: jasril@uin-suska.ac.id

Abstrak—Stroke adalah gangguan fungsional otak secara tiba-tiba yang disebabkan oleh disfungsi otak klinis fokal dan global selama 24 jam bahkan lebih. Sebanyak 15 juta jiwa meninggal diakibatkan oleh stroke setiap tahunnya. Pasien yang terkena stroke harus segera ditangani sehingga bisa meminimalisir resiko kerusakan otak. Salah satu pendukung diagnosis stroke yaitu melalui analisa citra hasil pemindaian Computed Tomography (CT-Scan). Seiring perkembangan zaman, teknologi pengolahan citra memungkinkan untuk mendeteksi pola stroke pada citra otak, dapat membantu dokter dan ahli radiologi dalam proses diagnosis dan pengobatan. Penelitian ini bertujuan untuk melakukan perbandingan tingkat akurasi menggunakan augmentasi dan tanpa augmentasi serta hyperparameter menggunakan Convolutional Neural Network dalam arsitektur EfficientNet-B0 untuk melakukan klasifikasi citra stroke iskemik, hemoragik, dan otak normal. Augmentasi data yang digunakan adalah dengan melakukan rotation, horizontal flip, dan pengaturan contrast pada data asli. Data uji disediakan sebanyak 20% dari porsi data asli dan augmentasi, dan 80% lainnya digunakan untuk proses training pencarian model optimal. Pencarian model berdasarkan komposisi data latih dan validasi dengan perbandingan 70:30, 80:20, dan 90:10. Hasil eksperimen menunjukkan performa yang terbaik diperoleh pada gabungan citra asli dan augmentasi, dengan akurasi dan F1-score berurut-turut sebesar 97%, 93%, dan 94% untuk data uji citra asli, citra augmentasi, dan citra gabungan. Penggabungan citra asli dan augmentasi untuk data training telah menunjukkan bahwa model cukup robust untuk dapat menghasilkan akurasi yang tinggi.

Kata Kunci: Augmentasi; CNN; Deep Learning; EfficientNet-B0; Citra Stroke

Abstract—A stroke is a sudden onset of brain dysfunction, lasting for 24 hours or longer, resulting from clinically focal and global brain dysfunction. As many as 15 million people die from stroke each year. The stroke patients need an immediate treatment to minimize the risk of brain damage. One of the proponents for the stroke diagnosis is through a computed tomography (CT) image. In recent years, the image processing techniques capable to detect stroke patterns in a brain image, it can be useful for doctors and radiologists in doing diagnosis and treatment. This study aims to compare the level of accuracy using augmentation and without augmentation and hyperparameters using the Convolutional Neural Network in the EfficientNet-B0 architecture to classify ischemic, hemorrhagic, and normal brain stroke images. The data augmentation is produced by rotating, horizontal flipping, and contrast tuning of the original data. Testing data is provided as much as 20% of the portion of the original and augmented data, and the other 80% is used for the training process to find the optimal model. The model search is based on the composition of the training and validation data with a ratio of 70:30, 80:20 and 90:10. The experimental results show that the best performance is obtained for the combined original and augmented images, with accuracies of 97%, 93%, and 94%, respectively, for the three types of data-test: original, augmented, and combined. The merging of original and augmented images for training data has shown that the model is robust enough in producing high accuracy results.

Keywords: Augmentation; CNN; Deep Learning; EfficientNet-B0; Stroke Images

1. PENDAHULUAN

Berdasarkan keterangan dari World Health Organization (WHO), Stroke adalah disfungsi otak secara tiba-tiba yang disebabkan oleh disfungsi neurologic fokal dan fungsional selama 24 jam bahkan lebih [1]. Berdasarkan data Badan Penyelenggara Jaminan Sosial (BPJS), terjadi peningkatan biaya pelayanan stroke tiap tahunnya. Pada tahun 2016 mencapai angka 1,43 triliun, kemudian naik menjadi 2,18 triliun pada tahun 2017, dan mencapai 2,56 triliun pada tahun berikutnya [2]. Tingginya angka biaya pelayanan tersebut karena penyakit stroke ini memerlukan terapi, alat kesehatan yang mumpuni, dan membutuhkan waktu yang lama untuk mencapai kesembuhannya. Bahkan, orang-orang yang kehilangan nyawa setiap tahun di dunia disebabkan oleh stroke, mencapai 15 juta jiwa berdasarkan statistik dari WHO [3].

Secara umum stroke diklasifikasikan menjadi dua, stroke iskemik dan hemoragik. Stroke iskemik terjadi pada saat pembuluh darah pembawa oksigen menuju otak tersumbat. Sedangkan stroke hemoragik terjadi pada saat pecahnya pembuluh darah di salah satu area otak yang mengakibatkan pendarahan [4]. Kedua jenis stroke tersebut merupakan stroke yang sangat berbahaya karena dapat menyebabkan kematian atau kecacatan dalam kurun waktu yang lama. Upaya yang dilakukan untuk mengetahui stroke salah satunya dengan melakukan pemeriksaan menggunakan CT. [5]

Computed Tomography (CT) merupakan alat yang sering kali digunakan untuk menghasilkan citraan pada otak. Alat tersebut akan menunjukkan representasi bagian otak yang terdeteksi stroke. Seorang radiolog memiliki tugas yang cukup besar untuk menangani pasien stroke sehingga sangat penting memiliki pengetahuan mengenai gambaran radiologis dan menetapkan pengobatan yang tepat terhadap pasien [6]. Seiring

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tesis Pengembangan Model Klasifikasi Citra Stroke Menggunakan Deep Learning

Perkembangan zaman, teknologi pengolahan citra memungkinkan memberi informasi mengenai stroke guna membantu dokter dan ahli radiologi dalam proses diagnosis dan pengobatan.

Penelitian terkait pemanfaatan pengolahan citra untuk klasifikasi citra stroke iskemik, stroke hemoragik yang sudah dilakukan diantaranya yaitu menggunakan Random Forest. Penelitian ini memperoleh hasil terbaik dengan akurasi precision, recall, dan f1-score mencapai 96% [7]. Dalam penelitian lain oleh Priya Govindarajan dkk menunjukkan bahwa klasifikasi menggunakan ANN mengungguli algoritma lain dengan akurasi mencapai 95% [8].

Pada bidang Artificial Intelligence saat ini terdapat metode Deep Learning, yang merupakan sebuah metode baru. Ada beberapa perbandingan antara Deep Learning dengan Machine Learning, menurut beberapa studi yang telah dikerjakan ole Kevin Leonard Kohsasih dkk memperlihatkan bahwa metode Deep Learning mendapatkan performa yang lebih tinggi dengan akurasi mencapai 98% dibandingkan menggunakan algoritma Machine Learning [9]. Penelitian selanjutnya menggunakan Convolutional Neural Network (CNN) [10] dan mendapatkan akurasi sebesar 94,2% yang menunjukkan bahwa CNN memiliki performa yang terbaik.

Penelitian stroke Deep Learning menggunakan CNN telah dilakukan pada tahun 2018 yang menghasilkan akurasi hampir 99% [11]. Penelitian klasifikasi citra stroke menggunakan CNN juga telah dikerjakan oleh J T Marbun dkk pada tahun 2018 dengan akurasi mencapai 90%. Menurut penelitian tersebut, hasil klasifikasi tergantung seberapa banyak data yang dipakai pada proses training. Semakin banyak data yang dipakai, maka semakin tinggi nilai akurasinya [12]. Berdasarkan penelitian tentang studi kasus stroke oleh M. S. Atshan menggunakan dataset MRI, bahwa arsitektur EfficientNet-B0 dapat meningkatkan akurasi dibandingkan dengan menggunakan CNN tanpa arsitektur [13]. Namun, penelitian tersebut masih belum menggunakan augmentasi. Maka dari itu, pada penelitian ini penulis melakukan eksperimen klasifikasi citra stroke dengan CNN arsitektur EfficientNet-B0 menggunakan augmentasi data. Menurut penelitian [14] dengan menggunakan augmentasi data dapat meningkatkan akurasi algoritma. Penggunaan augmentasi ini juga dimaksudkan untuk memperkaya citra sekaligus pengoptimalan data. Berdasarkan studi [15] augmentasi data dapat mengurangi overfitting atau dengan kata lain skor yang diperoleh pada data uji rendah meskipun skor akurasi yang diperoleh pada data uji tinggi.

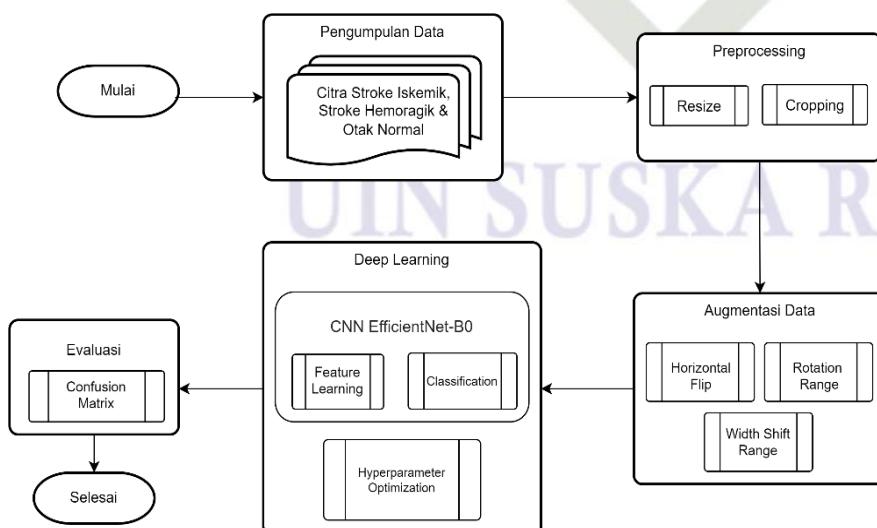
Penelitian ini bertujuan untuk melakukan perbandingan tingkat akurasi CNN arsitektur EfficientNet-B0 menggunakan augmentasi data dan tanpa augmentasi data serta hyperparameter untuk klasifikasi citra stroke iskemik, stroke hemoragik, dan otak normal.

Teorema No Free Lunch (NFL) secara singkat menyatakan bahwa semua algoritma pengoptimalan berkinerja sama baiknya [16]. Dengan memilih parameter yang tepat, pengoptimalan kinerja algoritma dapat dilakukan dengan baik. Dalam penelitian ini penulis menggunakan learning rate, batch size, dense, dan optimizer untuk kombinasi hyperparameter optimization. Hasil penelitian [17] menunjukkan bahwa dengan menggunakan hyperparameter dapat memperoleh akurasi yang tinggi pada metode yang ditawarkan.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Langkah-langkah penelitian untuk klasifikasi citra stroke iskemik, stroke hemoragik, dan otak normal menggunakan augmentasi, digambarkan dalam diagram pada Gambar 1. Tahapan-tahapan yang dikerjakan adalah sebagai berikut:



Langkah-langkah penelitian untuk klasifikasi citra stroke iskemik, stroke hemoragik, dan otak normal menggunakan augmentasi, digambarkan dalam diagram pada Gambar 1. Tahapan-tahapan yang dikerjakan adalah sebagai berikut:

2.2 Dilarang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah,
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Gambar 1. Langkah – langkah Penelitian Dari gambar 1 dapat dijelaskan:

Pengumpulan Data

Data citra CTScan diperoleh dari Rumah Sakit Syafira Pekanbaru, yang dikumpulkan antara Oktober 2022 dan Januari 2023. Data terdiri atas 3 kelas, yaitu citra stroke hemoragik, stroke iskemik, dan otak normal, dengan jumlah masing-masing kelas sebanyak 115 citra.

Preprocessing

Tahap preprocessing merupakan langkah penting dalam keberhasilan model deep learning karena secara signifikan dapat meningkatkan kinerjanya. Tahapan preprocessing ini dilakukan untuk mewujudkan citra yang lebih baik dari data CT Scan pasien stroke untuk dikerjakan pada tahap berikutnya. Mengenai preprocessing yang dikerjakan yaitu cropping dan resize. Gambar 2 berikut adalah contoh data citra yang digunakan dalam penelitian.

Cropping

Cropping dimaksudkan agar latar belakang yang berlebih di sekitar scan otak stroke dapat dihilangkan.

Resize

Pada proses resize, semua data CT scan otak stroke pada penelitian ini ukurannya diubah menjadi 224 x 224 piksel. Ukuran tersebut dipilih menyesuaikan dengan inputan arsitektur yang dipakai pada eksperimen ini.



Gambar 2. Sampel Data Citra Original

c. Augmentasi Data

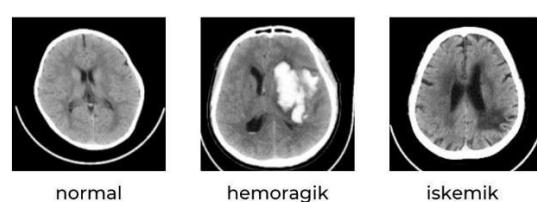
Teknik augmentasi dimaksudkan untuk memperkaya citra tanpa menghilangkan substansinya melalui proses modifikasi citra berdasarkan parameter tertentu. Selain itu, melakukan augmentasi data dapat meningkatkan akurasi model pada CNN. Pada penelitian ini penulis menggunakan beberapa parameter augmentasi, diantaranya rotation range (citra dirotasi sejauh beberapa derajat), horizontal flip (pencerminan pada citra), dan contrast (pengaturan kontras pada citra). Terdapat 1035 total data setelah dilakukan tahap augmentasi. Gambar 3-5 di bawah merupakan sampel data citra CT Scan untuk augmentasi data.



Gambar 3. Sampel Data Citra augmentasi horizontal flip



Gambar 4. Sampel Data Citra augmentasi rotation



Gambar 5. Sampel Data Citra augmentasi contrast

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

- 2 Konstruksi Dataset**
- Setelah dilakukannya tahap augmentasi data, selanjutnya dataset dibagi menjadi data pelatihan, data pengujian, dan data validasi. Penulis menggunakan perbandingan 80:20 untuk pemisahan data untuk proses pelatihan dan data pengujian. Tabel 1-3 di bawah ini merupakan statistik jumlah sampel pada pemisahan data uji, data latih dan data validasi untuk skema eksperimen pencarian model optimal.
2. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

Tabel 1. Pemisahan data latih dan data validasi 70 : 30

Citra	Kelas	Proses Pelatihan		Data Uji (20%)
		Data Latih (70%)	Data Validasi (30%)	
Asli (ori)	Stroke hemoragik	94	10	23
	Stroke iskemik	94	10	23
	Otak normal	94	10	23
	Stroke hemoragik	194	82	69
	Stroke iskemik	194	82	69
	Otaknormal	194	82	69

Tabel 2. Pemisahan data latih dan data validasi 80:20

Citra	Kelas	Proses Pelatihan		Data Uji (20%)
		Data Latih (80%)	Data Validasi (20%)	
Asli (ori)	Stroke hemoragik	74	18	23
	Stroke iskemik	74	18	23
	Otak normal	74	18	23
	Stroke hemoragik	221	55	69
	Stroke iskemik	221	55	69
	Otaknormal	221	55	69

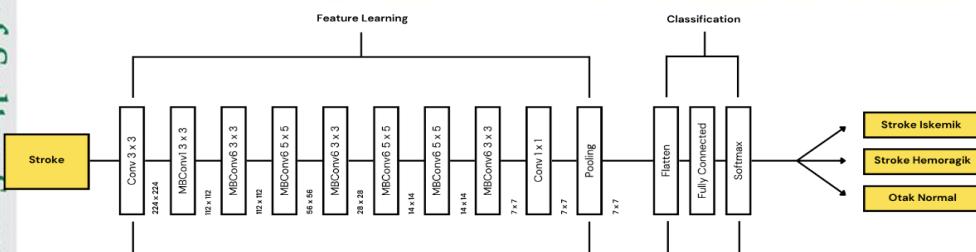
Tabel 3. Pemisahan data latih dan data validasi 90:10

Citra	Kelas	Proses Pelatihan		Data Uji (20%)
		Data Latih (90%)	Data Validasi (10%)	
Asli (ori)	Stroke hemoragik	83	9	23
	Stroke iskemik	83	9	23
	Otak normal	83	9	23
	Stroke hemoragik	249	27	69
	Stroke iskemik	249	27	69
	Otaknormal	249	27	69

2.3 Arsitektur EfficientNet-B0

Arsitektur EfficientNet-B0 merupakan arsitektur deep learning dengan metode CNN (convolutional neural network). Metode CNN mengungguli banyak metode machine learning dalam hal pengklasifikasian sebuah objek pada data gambar [18]. Hal ini dikarenakan penghitungan komputer yang lebih kuat, kumpulan data yang banyak dan juga cara melatih jaringan yang mendalam [19].

Penelitian ini menggunakan EfficientNet-B0 sebagai arsitektur. Pada tes ImageNet, arsitektur EfficientNet-B0 itu sendiri mengungguli banyak arsitektur lain seperti DenseNet, Inception, dan ResNet, maka dari itu EfficientNet sangat cocok dan disarankan untuk pengklasifikasian [20]. Selain itu, EfficientNet juga menggunakan parameter yang sedikit sehingga dapat menghemat dalam penggunaan memori. Walaupun model EfficientNet-B0 termasuk kecil, proses task klasifikasi nya tidak melibatkan akurasi sehingga EfficientNet menghasilkan performa yang sangat baik [21]. Gambar 6 di bawah ini merupakan ilustrasi dari Arsitektur EfficientNet-B0.



Gambar 6. Ilustrasi EfficientNet-B0

Dalam penelusuran model terbaik dalam tahap pelatihan (training), deep learning membutuhkan optimasi hyperparameter. Berdasarkan studi [17] menunjukkan bahwa dengan menggunakan hyperparameter method dapat memberikan dampak kinerja yang optimum pada suatu model. Dengan memilih parameter yang tepat, pengoptimalan kinerja algoritma dapat dilakukan dengan baik. Pada penelitian ini penulis menggunakan beberapa parameter seperti learning rate, batch size, dense, dan optimizer untuk kombinasi hyperparameter optimization.

4 Skema Eksperimen dan Evaluasi

Eksperimen yang dilakukan dalam penelitian ini dikerjakan menggunakan Google Colabotary dengan runtime GPU yang terkoneksi dengan Google Drive sebagai tempat penyimpanan dalam melakukan proses penelitian dengan algoritma CNN arsitektur EfficientNet-B0. Perangkat yang dipakai untuk melakukan pengujian yaitu menggunakan bahasa pemrograman Python versi 3.9.

Pencarian model optimal dimulai dari sebuah arsitektur standard sebagai baseline, dengan parameter-parameter yang ditentukan secara empiris. Kemudian parameter-parameter tersebut divariasikan untuk dipilih yang memberikan hasil paling optimal di antara variasi pilihannya. Parameter yang dioptimasi ditelusuri mulai dari optimizer, ukuran batch data pelatihan, learning rate, kedalaman dense layer, dan komposisi data sampel untuk proses pelatihan (data latih dan data uji), dengan pilihan sebagai berikut:

```
optimizer = {adam, RMSProp}
learning rate = {0.1, 0.01, 0.001, 0.0001}
batch size = {32, 64}
dense layer = {64, 256}
proporsi data latih dan validasi = {90: 10, 80: 20, 70: 30}
komposisi data = {ori, aug, ( ori + aug)}
```

(1)

Pengujian menggunakan confusion matriks dengan metrik pengukuran yang digunakan untuk pencarian model optimal adalah Akurasi dan F1-Score. Sedangkan hasil pengujian terhadap data uji, akan dihitung pula precision, recall dan F1-score-nya, menggunakan library sklearn. Hasil confusion matrix juga ditampilkan menggunakan library dari sklearn.

Dari persamaan (1), penelusuran model optimal dimulai dari baseline untuk masing-masing komposisi data pelatihan, dengan komposisi parameter: optimizer Adam, learning rate 0.001, batch size 32 dan kedalaman dense layer 256. Hasil pelatihan yang mencapai validasi akurasi terbaik disimpan dalam suatu file model. Selanjutnya, optimizer diubah menjadi RMSProp. Validasi akurasi terbaik dicatat, hasil pelatihan juga disimpan dalam file model yang lain. Kemudian dibandingkan, apabila Adam memiliki akurasi tertinggi, maka dipilih optimizer Adam untuk percobaan parameter berikutnya. Demikian sebaliknya.

Percobaan ini diulang terus dengan teknik pencarian parameter optimal yang sama, untuk parameter selanjutnya seperti dalam persamaan (1). Model yang terbaik untuk data pelatihan Ori, dipilih sebagai model final untuk sumber data Ori. Percobaan ini dilakukan kembali untuk data pelatihan yang bersumber dari data Aug. Model terbaik dari data Aug, dipilih sebagai model final untuk sumber data Aug.

Dari kedua set eksperimen di atas, dilihat kecenderungan proporsi data latih dan data validasi terbaik, yang ternyata diberikan oleh komposisi data 90:10. Untuk pengujian komposisi data Ori+Aug, maka hanya dilakukan untuk perbandingan data 90:10 saja, karena dari hasil tersebut dapat disimpulkan, semakin banyak data yang digunakan untuk pelatihan (data latih), semakin baik hasil akurasi validasinya.

3. HASIL DAN PEMBAHASAN

3.1 Penelusuran Model Optimal

Hasil pencarian parameter terbaik dari baseline yang digunakan, disajikan dalam Tabel 4-6. Dari skema eksperimen pertama, dilakukan pengujian dengan proporsi data latih dan data validasi berasal dari data Ori saja. Hal ini dituliskan dalam tabel 4, dengan parameter terpilih yang paling optimal adalah komposisi latih-validasi (90:10), kedalaman dense layer 64, dan batch size, learning rate dan optimizer berturut-turut 32, 0.1 dan Adam. Hasil validasi akurasi tertinggi yang dipilih dari penggunaan data latih Ori adalah 81.84% (Eksperimen nomor 7). Meskipun di ketiga porsi memiliki akurasi yang sama, mengingat jumlah data Ori sangat sedikit, maka dipilih porsi 90:10 agar variasi data latih yang terlihat selama training lebih banyak. Model dengan parameter ini akan diujikan pada data test di pengujian performa di bagian selanjutnya.

Tabel 4. Hasil penelusuran model terbaik untuk sumber data pelatihan dari data Ori saja.

Nomor Expt	Porsi Split Data Latih:Validasi	Dense	Batch Size	Learning Rate	Optimizer	Akurasi data validasi	Epoch
1	90:10	256	32	0.001	Adam	77.78%	28

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Nomor Exp.	Porsi Split Data Latih:Validasi	Dense	Batch Size	Learning Rate	Optimizer	Akurasi data validasi	Epoch
Hak Cipta Dilindungi Undang-Undang	80:20	256	32	0.001	RMSProp	77.78%	15
		256	32	0.1	Adam	77.78%	30
		256	32	0.01	Adam	77.78%	9
		256	32	0.0001	Adam	77.78%	10
		256	64	0.1	Adam	77.78%	2
		64	32	0.1	Adam	81.84%	15
		256	32	0.001	Adam	79.63%	3
		256	32	0.001	RMSProp	79.63%	23
		256	32	0.1	RMSProp	81.84%	5
		256	32	0.01	RMSProp	77.78%	7
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.	70:30	256	32	0.0001	RMSProp	79.63%	17
		256	64	0.1	RMSProp	79.63%	20
		64	32	0.1	RMSProp	79.63%	6
		256	32	0.001	Adam	79.01%	13
		256	32	0.001	RMSProp	81.48%	4
		256	32	0.1	RMSProp	76.54%	5
		256	32	0.01	RMSProp	79.01%	12
		256	32	0.0001	RMSProp	76.54%	21
		256	64	0.001	RMSProp	74.07%	4
		64	32	0.001	RMSProp	76.54%	29

Pada penelusuran data latih yang bersumber dari data Aug, seperti terlihat pada Tabel 5, diperoleh model terbaik dengan parameter kedalaman dense layer, batch size, learning rate dan optimizer secara berturut-turut adalah 256, 32, 0.001 dan Adam. Sedangkan porsi data latih dan data validasinya adalah 90:10 (Eksperimen nomor 26). Meskipun ada beberapa akurasi data validasi yang mencapai 100%, namun tetap harus dipilih 1 model saja untuk pengujian akhir di bagian selanjutnya.

Tabel 5. Hasil penelusuran model terbaik untuk sumber data pelatihan dari data Aug saja.

Nomor Exp.	Porsi Split Data Latih:Validasi	Dense	Batch Size	Learning Rate	Optimizer	Akurasi data validasi	Epoch
22	90:10	256	32	0.001	Adam	100.00%	9
23		256	32	0.001	RMSProp	100.00%	2
24		256	32	0.1	Adam	100.00%	7
25		256	32	0.01	Adam	100.00%	6
26		256	32	0.0001	Adam	100.00%	23
27		256	64	0.0001	Adam	100.00%	16
28		64	32	0.0001	Adam	100.00%	13
29		256	32	0.001	Adam	91.52%	6
30		256	32	0.001	RMSProp	92.12%	21
31		256	32	0.1	RMSProp	91.52%	16
32	80:20	256	32	0.01	RMSProp	91.52%	6
33		256	32	0.0001	RMSProp	90.91%	20
34		256	64	0.001	RMSProp	92.73%	5
35		64	64	0.001	RMSProp	91.52%	12
36		256	32	0.001	Adam	92.28%	18
37		256	32	0.001	RMSProp	92.68%	18
38		256	32	0.1	RMSProp	91.87%	28
39		256	32	0.01	RMSProp	92.28%	21
40		256	32	0.0001	RMSProp	92.28%	27
41		256	64	0.001	RMSProp	93.09%	22
42		64	64	0.001	RMSProp	92.28%	26

Terakhir, penelusuran data latih bersumber dari data gabungan Ori+Aug, seperti terlihat pada Tabel 6. Bedasarkan pengujian terpisah data latih Ori saja atau Aug saja, kecenderungan model terbaik adalah pada prosi data latih:validasi 90:10. Maka pada penelusuran data latih gabungan ini, hanya dilakukan pada porsi data pelatihan 90:10. Model terbaik dengan parameter kedalaman dense layer, batch size, learning rate dan optimizer secara berturut-turut adalah 64, 32, 0.001 dan Adam (Eksperimen nomor 48). Meskipun juga ada beberapa akurasi

data validasi yang mencapai 100%, namun tetap harus dipilih 1 model saja untuk pengujian akhir di bagian selanjutnya.

Tabel 6. Hasil penelusuran model terbaik untuk sumber data pelatihan dari gabungan data Ori+Aug.

Nomor Exp	Porsi Split Data Latih:Validasi	Dense	Batch Size	Learning Rate	Optimizer	Akurasi data validasi	Epoch
43	90:10	256	32	0.1	Adam	99.07%	11
44		256	32	0.01	Adam	100.00%	11
45		256	32	0.001	Adam	100.00%	20
46		256	32	0.0001	Adam	100.00%	13
47		256	64	0.001	Adam	100.00%	17
48		64	32	0.001	Adam	100.00%	20

2 Pengujian Performa Model

Dari ketiga model optimal yang dipilih berdasarkan sumber data pelatihan sebagaimana Tabel 4-6, dilakukan pengujian terhadap data test yang belum pernah terlihat pada saat training. Hasil yang diukur adalah akurasi dan F1-score untuk ketiga kelas klasifikasi. Data test yang diujikan, ada 3 kelompok data, yaitu data Ori saja (69 citra), data Aug saja (207 citra), dan data gabungan Ori+Aug (276 citra). Hasil akurasi dan F1-score dari ketiga kelompok pengujian data test ini juga disajikan pada Tabel 7 dan Tabel 8.

Dari Tabel 7 dan 8 terlihat bahwa model yang menggunakan augmentasi data untuk pelatihan dapat meningkatkan hasil akurasi terhadap data uji secara signifikan dibandingkan data Ori saja. Namun, berdasarkan hasil tersebut, terlihat bahwa gabungan data Ori+Aug lebih robust terhadap pertambahan jenis data uji, yaitu dari data uji bersumber dari Ori saja, dari Aug saja, maupun gabungan keduanya, model pelatihan dari gabungan data Ori+Aug menunjukkan hasil terbaik yang konsisten.

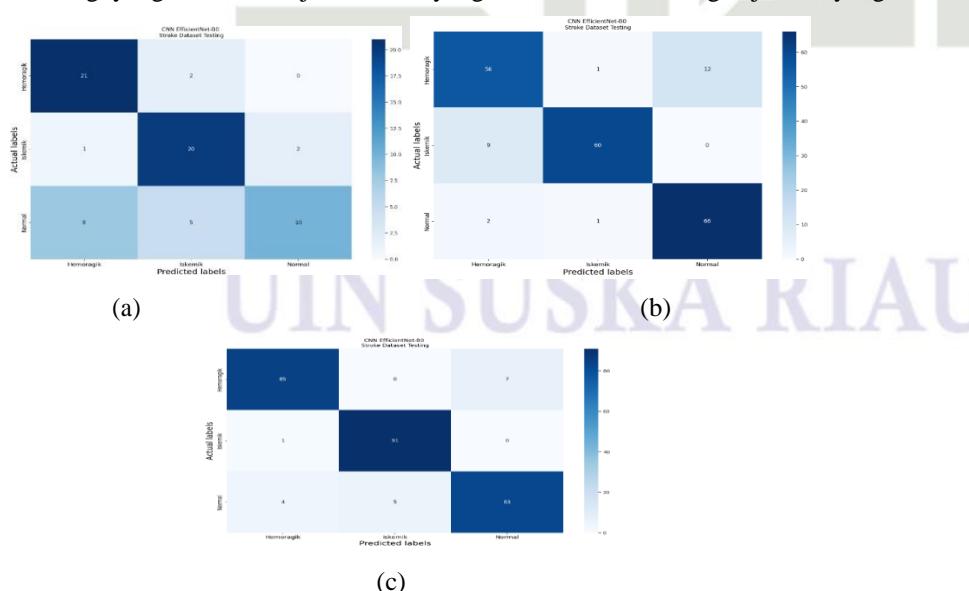
Tabel 7. Akurasi Hasil Pengujian

	Data Test		
	Ori (69)	Aug (207)	Ori+Aug (276 img)
Ori	74%	86%	83%
Aug	97%	88%	90%
Ori+Aug	97%	93%	94%

Tabel 8. F1-score Hasil Pengujian

	Data Test		
	Ori (69)	Aug (207)	Ori+Aug (276 img)
Ori	72%	86%	83%
Aug	97%	88%	90%
Ori+Aug	97%	93%	94%

Confusion matriks dari pengujian terhadap data uji, dapat dilihat pada Gambar 7 (a-c), masing-masing untuk pengujian terhadap data Ori saja (a), data Aug saja (b) dan gabungan Ori+Aug (c). Model yang digunakan untuk menampilkan confusion matriks tersebut adalah model terbaik yang dihasilkan dari pelatihan terhadap data yang bersumber dari Ori+Aug. Dari gambar confusion matriks ini, dapat terlihat kelas mana yang kurang dapat diklasifikasi dengan benar. Secara umum, model yang dikembangkan telah dapat mengenali ketiga kelas data dengan tepat dan seimbang, yang terlihat dari jumlah data yang salah klasifikasi dengan jumlah yang sedikit.



Gambar 7. Confusion Matrix pada Model terbaik (sumber pelatihan Ori+Aug)

terima kasih

- Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan,

b. Pengutipan tidak merugikan kepentingan yang wajar

UIN SUSKA RIAU

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

2

3 Analisa

Secara keseluruhan, model terpilih pada citra ori, citra aug, dan citra ori+aug diperoleh dengan porsi data latih dan data validasi 90:10, batch size 32, dan optimizer Adam. Pemilihan parameter yang tepat dapat menghasilkan performa akurasi yang baik. Pembagian porsi data latih dan data validasi dapat berpengaruh dalam meningkatkan performa akurasi. Semakin banyak data latih yang digunakan maka hasil performa akurasi dapat semakin meningkat. Adanya proses augmentasi digunakan untuk menambah data dan variasi data sehingga dapat meningkatkan generalisasi untuk mengurangi overfitting. Hasil pengujian pada citra ori+aug mendapatkan hasil performa yang tertinggi karena terdiri dari gabungan citra ori dan citra aug yang menjadikan data gambar lebih unik dibandingkan dengan tidak menggunakan augmentasi.

4. KESIMPULAN

Setelah melakukan seluruh rangkaian proses penelitian, maka dapat ditarik kesimpulan bahwa hasil performa tertinggi diperoleh pada porsi data latih dan data validasi 90:10 dengan akurasi berturut-turut 97%, 93%, 94% untuk uji citra original, citra augmentasi, dan citra gabungan (original + augmentasi). Dari hasil tersebut, maka penggabungan citra asli dan augmentasi untuk data training menunjukkan bahwa model dapat menghasilkan akurasi yang tinggi. Penelitian selanjutnya diharapkan dapat meningkatkan jumlah data yang lebih banyak dengan kualitas data yang sangat baik, dan ditambahkan dengan teknik augmentasi data lainnya.

UCAPAN TERIMAKASIH

Terima kasih kepada orang-orang yang terlibat membantu saya dalam penyusunan jurnal ini. Kepada Rumah Sakit Syafira Pekanbaru, terima kasih telah memberikan izin pengambilan data pada penelitian ini, dan tidak lupa kepada semua kolega yang ikut serta memberikan kontribusi berupa tenaga, fikiran, dan materinya.

REFERENCES

- [1] N. Permatasari, "Perbandingan Stroke Non Hemoragik dengan Gangguan Motorik Pasien Memiliki Faktor Resiko Diabetes Melitus dan Hipertensi," *Jurnal Ilmiah Kesehatan Sandi Husada*, vol. 11, no. 1, pp. 298–304, 2020, doi: 10.35816/jiskh.v10i2.273.
- [2] "Begini Cara Mengenali Gejala Stroke," Kementerian Kesehatan Republik Indonesia, Oct. 02, 2019. <https://www.kemkes.go.id/article/view/19102900002/begini-cara-mengenali-gejala-stroke.html> (accessed Mar. 14, 2023).
- [3] "Stroke, Cerebrovascular accident," WHO EMRO. <https://www.emro.who.int/health-topics/stroke-cerebrovascular-accident/index.html> (accessed Mar. 14, 2023).
- [4] M. I. Fuadi, D. P. Nugraha, and E. Bebasari, "Gambaran obesitas pada pasien stroke akut di Rumah Sakit Umum Daerah Arifin Achmad Provinsi Riau periode Januari-Desember 2019," *Jurnal Kedokteran Syiah Kuala*, vol. 20, no. 1, Apr. 2020, doi: 10.24815/jks.v20i1.18293.
- [5] R. E. Yunus, L. K. Putri, and F. Afif, "Peran CT Scan dan MRI dalam Diagnosis Stroke," Rumah Sakit Universitas Indonesia, 2022. <https://rs.ui.ac.id/umum/berita-artikel/artikel-populer/peran-ct-scan-dan-mri-dalam-diagnosis-stroke> (accessed Mar. 16, 2023).
- [6] Y. Yueniwati, *Pencitraan pada Stroke*, Cetakan Pertama. Malang: UB Press, 2016.
- [7] Md. M. Islam, S. Akter, Md. Rokunojjaman, J. H. Rony, A. Amin, and S. Kar, "Stroke Prediction Analysis using Machine Learning Classifiers and Feature Technique," *International Journal of Electronics and Communications Systems*, vol. 1, no. 2, pp. 57–62, Dec. 2021, doi: 10.24042/ijecs.v1i2.10393.
- [8] P. Govindarajan, R. K. Soundarapandian, A. H. Gandomi, R. Patan, P. Jayaraman, and R. Manikandan, "Classification of stroke disease using machine learning algorithms," *Neural Comput Appl*, vol. 32, no. 3, pp. 817–828, Feb. 2020, doi: 10.1007/s00521-019-04041-y.
- [9] K. L. Kohsasilh et al., "Analisis Perbandingan Algoritma Convolutional Neural Network dan Algoritma Multi-layer Perceptron Neural dalam Klasifikasi Citra Sampah," 2021. [Online]. Available: <http://ejournal.stmik-time.ac.id>
- [10] M. F. Naufal, "ANALISIS PERBANDINGAN ALGORITMA SVM, KNN, DAN CNN UNTUK KLASIFIKASI CITRA CUACA," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, vol. 8, Apr. 2021, doi: 10.25126/jtiik.202184553.
- [11] D. R. Pereira, P. P. R. Filho, G. H. de Rosa, J. P. Papa, and V. H. C. de Albuquerque, "Stroke Lesion Detection Using Convolutional Neural Networks," in *2018 International Joint Conference on Neural Networks (IJCNN)*, 2018, pp. 1–6, doi: 10.1109/IJCNN.2018.8489199.
- [12] J. T. Marbun, Seniman, and U. Andayani, "Classification of stroke disease using convolutional neural network," in *Journal of Physics: Conference Series*, Institute of Physics Publishing, Mar. 2018. doi: 10.1088/1742-6596/978/1/012092.
- [13] M. S. Atshan and Z. N. Nemer, "Pre-Diagnosing the Stroke Using Deep Learning," *Journal of Al-Qadisiyah for Computer Science and Mathematics*, vol. 14, no. 3, pp. 33–43, 2022, doi: 10.29304/jqcm.2022.14.3.984.
- [14] A. K. Nugroho, Dinar Mutiara Kusumo Nugraheni, Terawan Agus Putranto, I Ketut Eddy Purnama, and Mauridhi Hery Purnomo, "Classification of Ischemic Stroke with Convolutional Neural Network (CNN) approach on b-1000 Diffusion-Weighted (DW) MRI," *EMITTER International Journal of Engineering Technology*, pp. 195–216, Jun. 2022, doi: 10.24003/emitter.v10i1.694.



- [5] X. Zhu, J. Yuan, Y. Xiao, Y. Zheng, and Z. Qin, "Stroke classification for sketch segmentation by fine-tuning a developmental VGGNet16," *Multimed Tools Appl.*, vol. 79, no. 45–46, pp. 33891–33906, Dec. 2020, doi: 10.1007/s11042-020-08706-y.
- [6] P. Adam, S. A. N. Alexandropoulos, P. M. Pardalos, and M. N. Vrahatis, "No free lunch theorem: A review," in *Springer Optimization and Its Applications*, Springer International Publishing, 2019, pp. 57–82. doi: 10.1007/978-3-030-12767-1_5.
- [7] J. Wu, X. Y. Chen, H. Zhang, L. D. Xiong, H. Lei, and S. H. Deng, "Hyperparameter optimization for machine learning models based on Bayesian optimization," *Journal of Electronic Science and Technology*, vol. 17, no. 1, pp. 26–40, Mar. 2019, doi: 10.11989/JEST.1674-862X.80904120.
- [8] S. B. Jadhav, V. R. Udupi, and S. B. Patil, "Convolutional neural networks for leaf image-based plant disease classification," *IAES International Journal of Artificial Intelligence*, vol. 8, no. 4, pp. 328–341, Dec. 2019, doi: 10.11591/ijai.v8.i4.pp328-341.
- [9] Hendriyana and Y. H. Maulana, "Identifikasi Jenis Kayu menggunakan Convolutional Neural Network dengan Arsitektur Mobilenet," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 4, no. 1, pp. 70–76, 2020, Accessed: Mar. 14, 2023. [Online]. Available: <https://doi.org/10.29207/resti.v4i1.1445>
- [10] T. Samitha, "Brain Tumor Classification using EfficientNet Models," *International Research Journal of Engineering and Technology*, vol. 9, no. 8, Aug. 2022, [Online]. Available: www.irjet.net
- [11] E. Luz et al., "Towards an effective and efficient deep learning model for COVID-19 patterns detection in X-ray images," *Res. Biomed. Eng.*, vol. 38, pp. 149–122, 2022, doi: 10.1007/s42600-021-00151-6/Published.