

PENGARUH OPTIMASI DALAM KLASIFIKASI TUMOR OTAK DENGAN HIBRIDA CNN-ViT

Diterima Redaksi: 29 April 2024; Revisi Akhir: 9 Mei 2024; Diterbitkan Online: 10 Juni 2024

Ivan Christopher Sukandar¹⁾, Fetty Tri Anggraeny^{2)*}, Made Hanindia Prami Swari³⁾

^{1, 2, 3)} Fakultas Ilmu Komputer UPN "Veteran" Jawa Timur

^{1, 2, 3)} Jalan Raya Rungkut Madya, Gunung Anyar, Surabaya, Indonesia kode pos: 60294

e-mail: ivanchristopherea@gmail.com¹⁾, fettyanggraeny.if@upnjatim.ac.id²⁾, hanindia.pramiswari@gmail.com³⁾

Abstrak : Tumor merupakan suatu kelainan pada sel yang menyebabkan sel yang seharusnya hidup dan mati pada waktu tertentu menjadi tetap hidup dan tumbuh secara abnormal. Tumor dapat menyerang seluruh bagian tubuh manusia seperti otak. Dalam melakukan diagnosa terhadap penyakit tumor, dapat digunakan kecerdasan buatan untuk melakukan klasifikasi atau deteksi dengan cepat. Penggunaan kecerdasan buatan seperti Convolutional Neural Network(CNN) sudah umum digunakan hingga masuk ranah medis. Selain CNN juga terdapat algoritma lain seperti Vision Transformer(ViT) untuk melakukan pekerjaan serupa. Oleh karena itu, digunakanlah hibrida dari kedua metode tersebut untuk menunjang keunggulan masing-masing algoritma dan mendapatkan hasil yang memuaskan. Hasil dari hibrida CNN-ViT yang digunakan dalam penelitian ini mendapatkan hasil tertinggi dengan Optimasi Adam Learning Rate 0.001% dengan akurasi 94%, recall dan f1 score sebesar 94% serta presisi sebesar 95%.

Kata Kunci—Hibrida, Tumor Otak, Kecerdasan Buatan

Abstract : A tumor is an abnormality in cells that causes cells that should live and die at a certain time to remain alive and grow abnormally. Tumors can affect all parts of the human body such as the brain. In diagnosing tumors, artificial intelligence can be used to perform classification or detection quickly. The use of artificial intelligence such as Convolutional Neural Network (CNN) has been commonly used to enter the medical field. Besides CNN, there are also other algorithms such as Vision Transformer (ViT) to do similar work. Therefore, a hybrid of the two methods is used to support the advantages of each algorithm and get satisfactory results. The results of the CNN-ViT hybrid used in this study obtained the highest results with an Adam Optimization with Learning Rate of 0.001% with an accuracy of 94%, recall and f1 score of 94% and precision of 95%.

Keywords— Hybrid, Brain Tumor, Artificial Intelligence

I. PENDAHULUAN

TUMOR adalah sebuah benjolan abnormal yang dapat tumbuh di semua bagian tubuh manusia, tak terkecuali otak. Mutasi abnormal pada suatu sel menyebabkan suatu sel yang harusnya hidup dan mati pada kurun waktu tertentu menjadi berkembang biak secara abnormal dan menyebabkan tumor otak.[1] Penyakit tumor otak, tergolong ke dalam penyakit yang memerlukan penanganan yang cepat dan tepat. Penyakit ini sendiri mempengaruhi setidaknya 250.000 orang di seluruh dunia dengan tingkat harapan hidup atau disebut “5 year survival rate” sebesar 36% menurut *American Society of Clinical Oncology*. [2]

Harapan hidup para penderita tumor otak, dapat ditingkatkan apabila penderita mendapatkan hasil diagnosa yang cepat dan tepat. Dengan diagnosa yang cepat dan tepat, penanganan dapat segera dilakukan oleh dokter dan pasien dapat terhindar dari hal-hal yang tidak diinginkan seperti kematian, kerusakan jaringan otak dan kehilangan fungsi tubuh. Tumor otak sendiri memiliki berbagai jenis yang berbeda seperti *meningioma*, *glioma*, dan *pituitary tumor*. Oleh karena itulah diperlukan diagnosa yang tepat dikarenakan tiap jenis tumor memerlukan penanganan yang berbeda pula.[3]

Pada saat ini, dalam melakukan diagnosa suatu penyakit, terdapat berbagai hal yang dapat menjadikan landasan dalam melakukan diagnosa. Hal tersebut seperti diagnosa menggunakan data berupa citra ataupun teks. Dalam melakukan diagnosa atau klasifikasi penyakit menggunakan teks, diperlukan suatu sistem pakar yang menggunakan *artificial intelligence* yang menggabungkan basis pengetahuan (*knowledge based*). [4]

Penelitian yang menggunakan sistem pakar untuk melakukan deteksi penyakit dengan data berbasis teks pernah dilakukan pada tahun 2019. Pada penelitian ini, diterapkan suatu sistem pakar untuk mendeteksi penyakit yang diakibatkan oleh virus dengan menggunakan *certainty factor*(CF). Selain CF, terdapat 29 *rule* yang digunakan sebagai *evidence* dalam penelitian ini. Selain melakukan pendeteksian penyakit dengan menggunakan teks, terdapat cara lain yaitu dengan data berupa gambar.[4]

Dalam melakukan klasifikasi penyakit dengan menggunakan gambar, terdapat beberapa algoritma yang dapat digunakan seperti *Convolutional Neural Network*(CNN). CNN merupakan salah satu algoritma atau arsitektur dalam *machine learning* yang digunakan untuk memproses dan menganalisis data berupa citra atau data visual. CNN sendiri ditemukan oleh Yann LeCun 1989 dan pada awalnya diberi nama *LeNet*. Pada saat itu, penelitian yang menggunakan CNN terbilang cukup sedikit karena kurangnya *hardware* terutama *Graphics Processing Unit* (GPU) yang digunakan untuk pemrosesan grafis. Akan tetapi, karena keberhasilan *AlexNet* dalam “*ImageNet Large Scale Visual Recognition Challenge*” 2012 penggunaan CNN menjadi cukup populer.[5]

Penelitian mengenai penggunaan CNN dalam melakukan klasifikasi pernah dilakukan pada tahun 2023 dengan menggunakan arsitektur *alexnet* dan *resnet34* dengan memanfaatkan *transfer learning*. Dalam penelitian tersebut, penyakit yang dideteksi adalah penyakit yang ada pada daun kentang. Penelitian tersebut mendapatkan hasil yang cukup memuaskan dengan tingkat akurasi hingga 99%. Selain menggunakan algoritma CNN, juga terdapat algoritma lain yang dapat digunakan untuk melakukan klasifikasi data citra yaitu *Vision Transformer*. [6]

Penggunaan algoritma CNN sendiri sudah masuk ke ranah medis untuk melakukan deteksi pada suatu penyakit berbasis gambar. hal tersebut dikarenakan kapasitas CNN untuk mempelajari representasi yang kompleks dan canggih melalui pendekatan berbasis data. [7] Oleh karena itu, penggunaan CNN sendiri dalam melakukan klasifikasi sudah menjadi hal yang umum digunakan. Selain penggunaan CNN terdapat beberapa algoritma lain yang dapat digunakan salah satunya adalah *Vision Transformer*(ViT).

Vision Transformer(ViT) merupakan sebuah algoritma dengan pendekatan baru dalam pemrosesan citra yang menggunakan mekanisme *attention* dari *transformer* untuk menggantikan penggunaan konvolusi yang umum dalam jaringan saraf konvensional. Pendekatan ini memungkinkan ViT untuk mencapai kinerja yang superior dalam pengolahan citra, sambil memungkinkan proses pelatihan yang lebih paralel dan membutuhkan waktu yang lebih sedikit.[8]

Model hibrida sendiri membawa suatu hal yang menjanjikan dalam melakukan klasifikasi citra medis. Hal tersebut dikarenakan dengan adanya model hibrida diharapkan dapat meningkatkan akurasi suatu sistem diagnostik.[7] Penelitian terkait model hibrida sendiri pernah dilakukan oleh Liu, dkk yang menghasilkan *PHTrans* sebuah model hasil hibrida *CNN* dan *Transformer* secara paralel untuk menangkap kedua fitur global dan lokal guna mencapai kinerja segmentasi yang unggul.[9] Oleh karena itulah, penelitian kali ini akan membahas mengenai pengaruh optimasi terhadap klasifikasi tumor otak dengan hibrida CNN-ViT.

II. TINJAUAN PUSTAKA

A. Penelitian Terdahulu

Berbagai penelitian terkait penggunaan CNN, ViT ataupun metode hibrida sudah sering dilakukan dengan berbagai studi kasus berbeda. Pada tahun 2023 dilakukan suatu penelitian yang dilakukan oleh Trong-Thuan Nguyen, Tam V. Nguyen dan Minh-Triet Tran yang menggunakan metode hibrida *CNN* dan *Vision Transformer* untuk melakukan klasifikasi pada citra *X-ray* pada paru-paru untuk mendeteksi COVID-19. Pada penelitian tersebut data yang digunakan adalah data sekunder yang bernama *Chest X-ray dataset* dan *Clean-CC-CCII* yang memiliki total data setidaknya tiga ratus ribuan gambar. Pendekatan yang dilakukan dalam penelitian ini adalah dengan menggabungkan kekuatan dari model *CNN* dan *ViT*. Model *ViT* digunakan untuk menangkap informasi kontekstual global dan mempelajari representasi dari data visual, sementara model *CNN* unggul dalam menangkap pola lokal dan mengekstrak detail-detail rumit dari gambar medis. Model yang mereka gunakan dalam *CNN* sendiri adalah *Densenet-201* sedangkan untuk *ViT* menggunakan *Token-to-Token-ViT*(T2T-ViT).

Hasil yang didapat dari penelitian ini sendiri bisa dibilang cukup mengejutkan, dengan metode yang diusulkan hasil akurasi yang didapat cukup besar dibandingkan metode *T2T* 88.51%, *DenseNet-201* 93.18% sedangkan metode yang diusulkan mendapat akurasi 98.96%. [7]

Penelitian lain pernah dilakukan pada tahun 2020 dengan melakukan hibrida CNN-ELM, penelitian ini dilakukan dengan melakukan ekstraksi fitur dengan menggunakan CNN yang kemudian hasil output akan dimasukkan kedalam lapisan tersembunyi ELM. Selain itu, pada penelitian ini juga dilakukan beberapa skenario pengujian dengan mengatur jumlah layer konvolusi yang ada pada CNN serta jumlah node yang digunakan dalam ELM. Beberapa skenario tersebut dilakukan dengan nilai 6 filter, 8 filter dan 10 filter pada CNN serta jumlah node 4000, 5000, dan 6000 pada ELM. Hasil yang didapat dari penelitian ini apabila dibandingkan dengan metode pada penelitian terdahulu yang terkait mengalami peningkatan. Metode hibrida CNN-SVM memiliki akurasi 76%, ELM 86% dan metode yang diusulkan penulis memiliki tingkat akurasi 91,4%. Selain itu, tingkat presisi yang didapatkan dengan metode yang diusulkan penulis adalah 91.5% dengan *recall* 91.5% dan *f1-score* 91.5% [10]

Penelitian lain terkait klasifikasi tumor otak pernah dilakukan pada tahun 2022 dengan menggunakan CNN arsitektur *MobileNetV2*. Dataset yang digunakan terdiri dari 3000 data yang terbagi ke dalam 4 kelas yaitu meningioma, glioma, pituitary tumor dan tanpa tumor. Dari data tersebut, dilakukan pembagian data 80% data latih, 10% data validasi dan 10% data uji. Penelitian tersebut melakukan berbagai skenario pengujian seperti pengujian optimizer antara Adam dan RMSProp, pengujian batch size berukuran 32, 64 dan 128 serta pengujian terhadap epoch 10, 30, 50 dan 100. Hasil dari pengujian yang dilakukan, didapatkan bahwa epoch terbaik adalah 30, batch size terbaik adalah 32 dan optimizer terbaik adalah adam. Hasil yang didapat adalah tingkat akurasi sebesar 88.64%, precision 90%, recall 89% dan f1-score 89%. [3]

Pada tahun 2021 dilakukan penelitian terkait klasifikasi 3 kelas tumor otak yang dilakukan oleh Neluun, Sellapan, Abdul, Iftikhar dan Madini dengan berbasis *fine-tuned* model dan *ensemble method*. Data yang digunakan dalam penelitian tersebut adalah data tumor otak yang memiliki jumlah 3064 *MRI* yang nantinya dikonversi “.png”. Pada penelitian tersebut model yang diusulkan oleh peneliti adalah dengan menggunakan *fine-tuned inception v3* dan *fine-tuned xception* model sebagai model yang digunakan dalam melakukan pemrosesan pada gambar. Metode *ensemble* yang diterapkan oleh peneliti terletak pada *classifier* atau pengklasifikasi yang digunakan. Beberapa metode yang digunakan sendiri adalah *Support Vector Machine*(SVM), *Random Forest*(RF) dan *K-Nearest Neighbors*(KNN). Hasil output *feature vector* dari model yang mereka buat, nantinya akan dimasukkan kedalam setiap algoritma *classifier* yang digunakan. Lalu setelah itu, barulah dilakukan *majority voting* untuk menentukan kelas yang diprediksi. Hasil dari penelitian ini sendiri adalah tingkat akurasi yang didapatkan adalah sebesar 94.34% untuk *inception-v3-ensemble* dan 93.79 untuk *xception-ensemble*. [11]

Pada Tahun 2022 dilakukan penelitian terkait penggunaan optimasi SGD, RMSprop dan Adam dalam melakukan klasifikasi hewan dengan menggunakan CNN. Penelitian tersebut menggunakan 5000 gambar dari beberapa hewan seperti beruang, gajah, orang utan, harimau dan berbagai hewan lainnya. Pembagian data dalam penelitian tersebut adalah sebesar 80% untuk pelatihan dan 20% untuk *testing* atau pengujian. Dalam penelitian tersebut, juga dilakukan skenario terkait penggunaan berbagai *learning rate* seperti 0.1%, 0.01% dan 0.001% dalam melakukan pengujian yang dilakukan. Hasil yang didapat dalam penelitian tersebut adalah optimasi SGD hanya mendapatkan hasil yang cukup baik pada *learning rate* 0.1% dengan tingkat akurasi terbaik sebesar 74.80%, presisi 73%, *recall* 82% dan *f1 score* 76%. Sedangkan untuk optimasi Adam dan RMSprop yang diusulkan mendapatkan hasil yang sangat buruk pada *learning rate* 0.1% dan mendapatkan hasil yang paling optimal pada *learning rate* 0.001%. Hasil terbaik RMSprop *learning rate* 0.001% adalah akurasi sebesar 80.60%, presisi 85%, *recall* 82% dan *f1 score* 84%. Untuk Adam *learning rate* 0.001% mendapatkan hasil terbaik dengan akurasi sebesar 89.81% presisi 87%, *recall* 82% dan *f1 score* 84%. Oleh karena itu, berdasarkan penelitian tersebut didapatkan bahwa optimasi terbaik adalah Adam, RMSprop dan terakhir SGD. Selain itu, tingkat *learning rate* dianggap cukup berpengaruh dalam suatu model karena mempengaruhi hasil yang didapat oleh model tersebut. [12]

B. Convolutional Neural Network(CNN)

Convolutional Neural Network(CNN) dalam *deep learning* adalah kelas *Artificial Neural Network(ANN)* yang umum diterapkan dalam bidang pengolahan citra untuk analisis citra visual.[13] CNN menggunakan konsep konvolusi, yaitu teknik pemrosesan data dengan menggunakan filter atau *kernel* pada seluruh bagian citra. Setiap *kernel* digunakan untuk menangkap fitur pada citra seperti garis, kontur, dan warna. Teknik pemrosesan ini dilakukan secara iterasi atau berulang-ulang pada setiap lapisan CNN sehingga fitur yang dideteksi semakin kompleks dan abstrak pada lapisan yang lebih tinggi. Nantinya fitur-fitur tersebut akan ditempatkan didalam *feature map*. [14]

Dalam CNN terdapat 3 komponen utama yaitu konvolusi, *pooling*, dan *fully connected layers*. Lapisan konvolusi adalah lapisan yang digunakan untuk mengekstrak fitur-fitur pada citra dengan menggunakan *kernel*. Proses konvolusi sendiri dilakukan dengan operasi matematika(matriks) pada setiap *pixel* citra dengan *kernel* yang digeser. Setelah konvolusi, dilakukan proses *pooling* untuk mengurangi dimensi citra dan meningkatkan *robustness*(kekokohan) terhadap perubahan skala dan orientasi. Langkah akhir adalah *fully connected layers* yang digunakan untuk melakukan klasifikasi citra yang telah diolah melalui neuron-neuron dari lapisan sebelumnya. [14]

C. Vision Transformer

Vision Transformer(ViT) adalah arsitektur pengolahan citra yang menggunakan dasar dari arsitektur *transformer*. Sebelumnya arsitektur tersebut digunakan dalam bidang *Natural Language Processing(NLP)*. Sedangkan sebelum arsitektur ini ada, terdapat arsitektur yang berhubungan dengan visi komputer menggunakan metode *Convolutional Neural Network(CNN)*. Penelitian pertama mengenai *Transformer* dalam ranah visi komputer sendiri ada pada penelitian yang dilakukan oleh Dosovitskiy, dkk yang membahas mengenai penggunaan modifikasi *Transformer* agar dapat digunakan untuk melakukan pemrosesan citra. Modifikasi yang dilakukan adalah dengan membagi gambar menjadi beberapa bagian atau yang disebut dengan nama *patch* dan memberikan urutan penyematan linear dari bagian ini sebagai *input* ke *Transformer*. [15], [16]

D. Optimasi Adam dan RMSprop

Optimasi Adam adalah algoritma yang digunakan untuk menggantikan prosedur *stochastic gradient descent* dan digunakan untuk memperbarui bobot jaringan secara iteratif berdasarkan data latih. Adam merupakan metode pengoptimalan yang populer digunakan dalam *deep learning*, algoritma ini mampu mendapatkan hasil yang optimal dan memiliki waktu komputasi yang lebih baik. Pada penerapannya optimasi adam mampu memberikan hasil perbaruan bobot yang lebih baik dibandingkan metode-metode optimasi yang ada seperti *Root Mean Square Propagation(RMSprop)*, *Adaptive Gradient Algorithm(Adagrad)* dan *Stochastic Gradient Descent(SGD)*. Hal tersebut dikarenakan optimasi adam memanfaatkan Adagrad untuk mempertahankan *learning* per-parameter untuk peningkatan kinerja model dan memanfaatkan RMSProp yang diadaptasi berdasarkan besaran rata-rata gradien bobot yang dimiliki oleh model.[17]

Optimasi RMSprop adalah modifikasi dari algoritma AdaGrad. Optimasi ini unggul dengan pengaturan *nonconvex* yang mengubah total dari gradien ke dalam *moving average* eksponensial.[18] RMSprop atau *Root Mean Square* membagi gradien dengan akar kuadrat rata-rata. Geoffrey Hinton penemu RMSprop menggunakan stokastik untuk *mini-batch*. Serupa dengan Adam, *learning rate* adaptif digunakan saat pelatihan oleh.[19]

E. Hibrida CNN-ViT

Metode hibrida adalah sebuah metode yang menggabungkan dua atau lebih algoritma *machine learning* atau *deep learning* yang bertujuan untuk mendapatkan hasil yang lebih baik dan optimal. Dengan menggabungkan algoritma-algoritma tersebut dapat meningkatkan kekuatan dan fleksibilitas dari suatu model yang dibuat. Selain itu, juga mengurangi kelemahan yang ada pada masing-masing algoritma tersebut. Penggunaan hibrida sendiri juga sudah masuk ke ranah medis, karena menjanjikan hasil yang lebih baik dibandingkan algoritma biasa.[20]

Hibrida CNN-ViT sendiri adalah penggabungan antara dua algoritma *deep learning* yang cukup populer. Algoritma CNN sendiri adalah algoritma yang populer dalam memproses data gambar

sedangkan untuk *ViT* merupakan pengembangan dari *transformer* yang awalnya digunakan untuk NLP. Penelitian terkait hibrida CNN-*ViT* pernah dilakukan oleh Trong-Thuan Nguyen, dkk pada citra *x-ray* untuk klasifikasi COVID-19. Pada penelitian tersebut Trong-Thuan Nguyen beranggapan bahwa DenseNet-201 dan T2T-*ViT* memiliki kekuatan masing-masing dimana DenseNet mampu mempelajari *local features* dan T2T-*ViT* mampu menggunakan mekanisme *self-attention* untuk memodelkan hubungan antara *image patches* dan untuk *long-range dependencies* serta memahami konteks gambar secara global. Dalam penelitian tersebut, *feature map* atau fitur yang berhasil diekstrak dari masing-masing metode digabungkan sehingga didapatkan keuntungan dari DenseNet dan T2T-*ViT* secara bersamaan. Selain itu, dengan menggabungkan kedua metode tersebut model yang dihasilkan pun menjadi lebih kokoh.[7]

F. Confusion Matrix

Dalam mengukur kinerja prediktif dari suatu sistem pengklasifikasi yang dibuat diperlukan suatu metode penilaian, salah satunya adalah metode *confusion matrix*. *Confusion matrix* dapat menunjukkan hasil prediksi pengklasifikasi dengan membandingkan hasil klasifikasi oleh sistem yang dibuat dan hasil klasifikasi yang sebenarnya. Peran *confusion matrix* adalah membuktikan performa suatu model yang telah dibuat yaitu dengan nilai akurasi, presisi, *recall* dan *f1-score* berdasarkan tiap nilainya.[21]

Tabel 1. Confusion Matrix

(n)	P	N
P	True Positive(TP)	False Positive(FP)
N	False Negative(FN)	True Negative(TN)

Pada Tabel 1. nilai pengukuran matriks adalah *True Positive*(TP), *True Negative*(TN), *False Positive*(FP), dan *False Negative*(FN). Setiap probabilitas peristiwa yang semestinya positif(P) dan setiap probabilitas yang semestinya negative(N). Nilai-nilai tersebut akan digunakan dalam metrik-metrik akurasi, presisi, *recall* dan *f1-score*. [22]

$$Akurasi = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

Akurasi adalah pengujian berdasarkan kedekatan nilai prediksi dengan nilai yang sebenarnya. Dalam kata lain, akurasi adalah rasio prediksi benar (positif dan negatif) keseluruhan data. [22]

$$Presisi = \frac{TP}{TP + FP} \quad (2)$$

Presisi adalah metode pengujian dengan perbandingan jumlah informasi yang bersangkutan dengan jumlah informasi yang terambil oleh sistem baik berhubungan maupun tidak. [22]

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

Recall adalah metode pengujian dengan membandingkan jumlah informasi bersangkutan yang didapatkan oleh sistem dengan seluruh informasi relevan pada kumpulan data, baik yang diambil oleh sistem maupun tidak. [22]

$$F1 - score = \frac{2 \times precision \times recall}{precision + recall} \quad (3)$$

F1 score adalah perbandingan dari rata-rata *precision* dan *recall* yang dilakukan pembobotan. Rentang yang dihasilkan oleh *F1 score* sendiri adalah antara 0 hingga 1. Jika skor yang didapatkan baik, maka model klasifikasi dianggap memiliki presisi dan *recall* yang baik. [22]

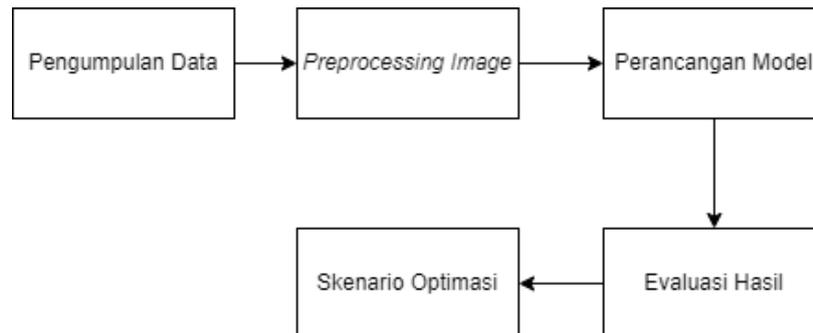
III. METODE PENELITIAN

Dalam penelitian ini terdapat beberapa tahapan yang perlu dilakukan untuk mendukung keberhasilan penelitian ini. Selain itu, dalam menunjang penelitian ini terdapat spesifikasi *hardware* dan *software* yang digunakan dalam menjalankan program yang dapat dilihat pada Tabel 2.

Tabel 2. Spesifikasi Perangkat

Jenis	Spesifikasi yang digunakan
OS	Windows 11
CPU	Intel core i5-8265u @1.60GHz(8 CPUs)
Graphics	Intel UHD Graphics 620
RAM	12GB(4GB + 8GB)
Software	Jupyter Notebook(Python 3.8.8)

Pada Tabel 2. Dapat dilihat spesifikasi *hardware* dan software yang digunakan dalam penelitian ini. Selanjutnya adalah tahapan-tahapan yang dilakukan dalam penelitian ini terbagi menjadi beberapa tahapan yang dapat dilihat pada Gambar 1.

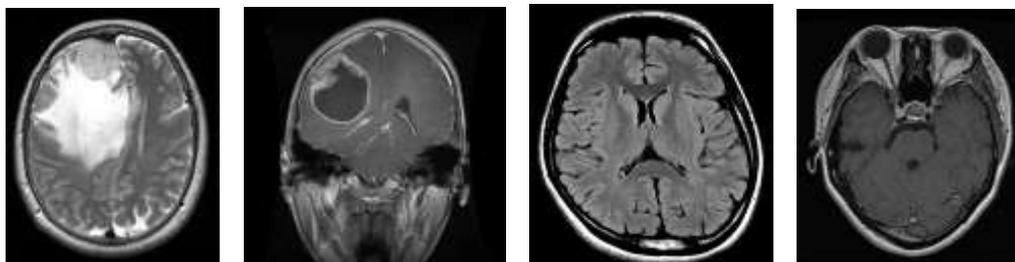


Gambar 1. Alur Penelitian

Pada Gambar 1 dapat dilihat alur penelitian ini dimulai dari pengumpulan data yang diambil melalui *website kaggle*, *preprocessing image*, perancangan model, evaluasi hasil dan skenario optimasi yang dijalankan.

A. Pengumpulan Data

Pengumpulan data dilakukan dengan mencari data-data terkait studi kasus yang diteliti, yakni data citra *MRI* tumor otak. Data tumor otak yang digunakan dalam penelitian ini merupakan data sekunder yang diambil melalui *website kaggle* dengan nama dataset *Brain Tumor Classification (MRI)* yang dapat diakses melalui <https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri>. Data tersebut berjumlah 3264 dan data tersebut berekstensi “.jpg” yang terbagi menjadi 926 gambar *glioma tumor*, 937 gambar *meningioma tumor*, 500 gambar *no tumor* dan 901 gambar *pituitary tumor*. Data tersebut memiliki ukuran 512x512 *pixel* dan *channel* warna RGB. Dalam penelitian kali ini, dari 3264 data tersebut hanya digunakan 500 data dari tiap kelasnya. Dataset tersebut awalnya dilakukan pengunduhan kemudian dimasukkan ke dalam folder yang sudah dibuat. Sampel gambar dapat dilihat pada Gambar 2.

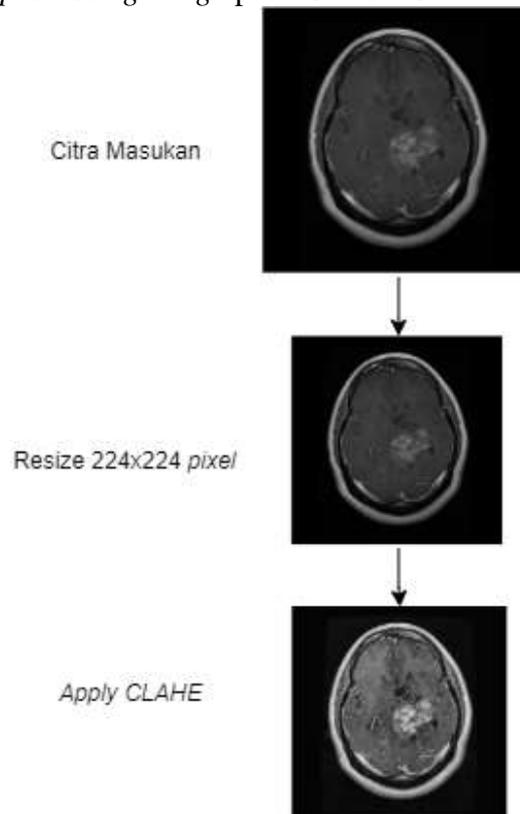


Gambar 2. Sampel Data Tumor meningioma, glioma, normal dan pituitary tumor

Pada Gambar 2 dapat dilihat sampel data yang digunakan pada penelitian ini. Dari kiri ke kanan secara berurutan ditampilkan data dari tiap kelas *meningioma*, *glioma*, *no tumor* dan *pituitary tumor*. Data-data tersebut nantinya akan dilakukan *preprocessing* data pada tahap selanjutnya.

B. *Preprocessing Image*

Tahapan *preprocessing image* dilakukan setelah melakukan pengumpulan data tumor yang diambil melalui *website kaggle*. Tahapan *preprocessing image* ini terbagi ke dalam beberapa tahapan yang dapat dilihat pada alur *preprocessing image* pada Gambar 3.

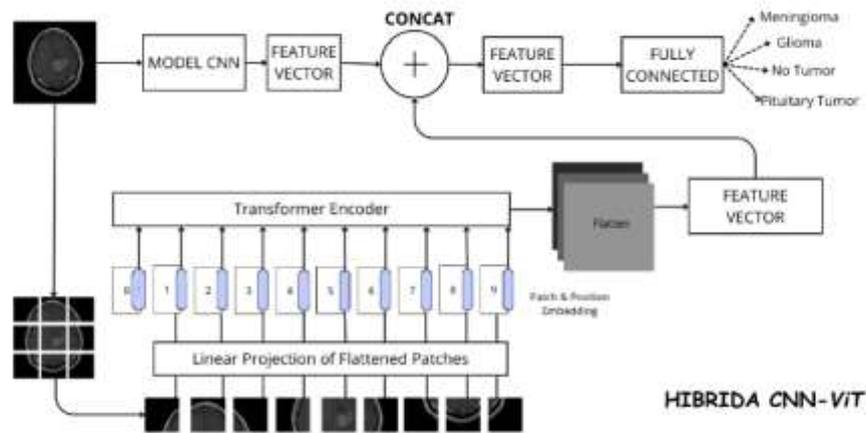


Gambar 3. *Preprocessing Image*

Pada Gambar 3 dapat dilihat tahapan yang harus dilalui oleh gambar dalam *preprocessing image*. Gambar yang dimasukkan awalnya dilakukan *resizing* menjadi berukuran yang sama yaitu *224x224 pixel*. Setelah itu gambar dilakukan konversi menjadi *grayscale* sebelum nantinya dilakukan *Contrast Limited Adaptive Histogram Equalization (CLAHE)*. Barulah setelah itu dikembalikan menjadi gambar dengan *channel RGB*. Hal ini dikarenakan model hanya dapat menerima gambar dengan *channel* warna RGB.

C. *Perancangan Model*

Model hibrida yang digunakan pada penelitian ini terdiri dari 2 bagian utama yang penting yaitu CNN dan ViT. Untuk lebih jelasnya dapat dilihat pada Gambar 4. Penggunaan CNN dan ViT ini dimaksudkan untuk mengambil *feature vector* yang nantinya akan digabungkan sebelum dilakukan klasifikasi.



*Ilustrasi Gambar Diatas Belum Melewati *Preprocessing Image*

Gambar 4. Hibrida CNN-ViT

Pada Gambar 4 dapat dilihat bahwa nantinya model yang dibangun akan menerima masukan berupa 2 gambar. Masing-masing gambar tersebut akan diproses menggunakan model CNN *custom* yang sudah dibuat dan gambar lain dimasukkan ke dalam *vision transformer*. Nantinya tiap *feature vector* masing-masing model akan dilakukan *concat* sebelum akhirnya dihubungkan ke *fully connected* untuk dilakukan klasifikasi.

D. Evaluasi Hasil

Tahapan selanjutnya dalam penelitian ini adalah dengan melakukan evaluasi hasil. Evaluasi hasil ini dilakukan dengan menggunakan *confusion matrix* dan *classification report*. Dengan menggunakan 2 metode tersebut nantinya akan didapatkan tingkat akurasi, presisi, *recall* dan *f1 score*. Evaluasi terakhir yang dijalankan adalah memuat skenario pengujian yang ada ke dalam satu tabel. Hal ini dilakukan untuk membandingkan akurasi pada skenario mana yang memiliki hasil yang paling baik.

E. Skenario Optimasi

Dalam penelitian ini terdapat 4 skenario optimasi yang digunakan dalam penelitian ini. Sedangkan untuk pembagian data yang dilakukan dalam penelitian ini adalah 80% data latih, 10% data uji dan 10% data validasi. Untuk skenario dapat dilihat pada Tabel 2.

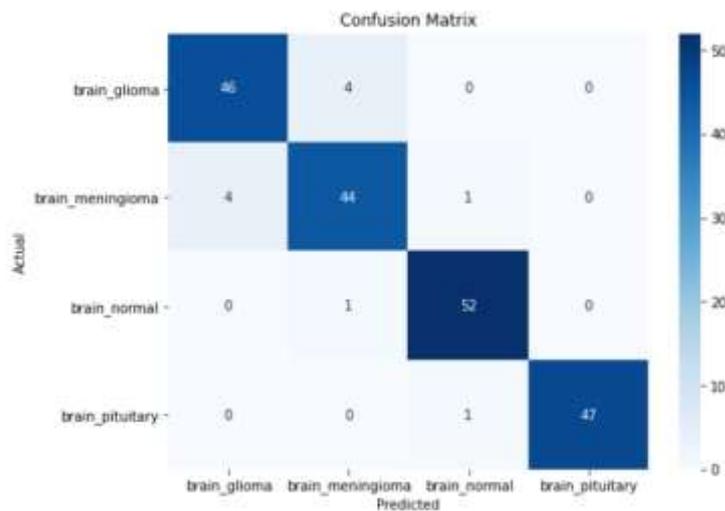
Tabel 3. Skenario Optimasi

No.	Metode	Optimasi	Learning Rate
1.	CNN-ViT	Adam	0.001%
2.			0.0001%
3.		RMSprop	0.001%
4.			0.0001%

Pada Tabel 2 dapat dilihat skenario optimasi yang dilakukan dalam penelitian ini. Tujuan dari skenario tersebut adalah untuk menentukan skenario mana yang mendapatkan hasil terbaik diantara tiap optimasi yang digunakan.

IV. HASIL DAN PEMBAHASAN

Hasil yang didapat dari penelitian ini berupa *confusion matrix* dan *classification report* berdasarkan tiap skenario yang sudah dijelaskan. Hasil dari *confusion matrix* skenario Adam *learning rate* 0.001% dapat dilihat pada Gambar 5.



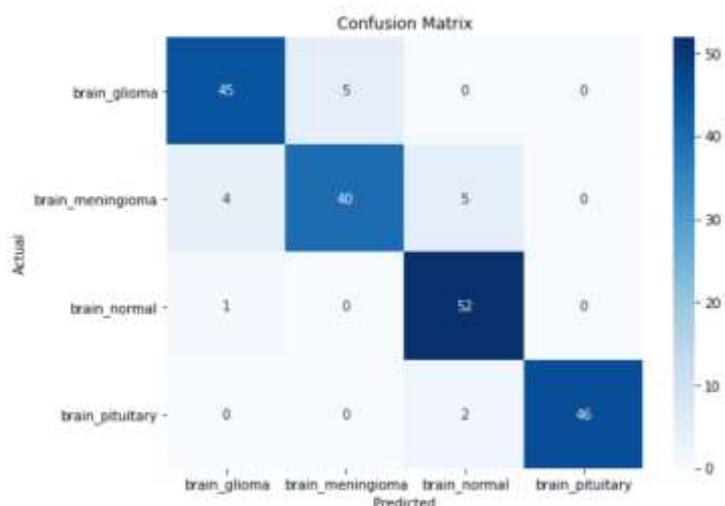
Gambar 5. Confusion Matrix Skenario Adam 0.001%

Setelah mendapatkan hasil *confusion matrix* pada Gambar 5, selanjutnya didapatkan hasil *classification report* skenario Adam dengan *learning rate* 0.001% dapat dilihat pada Tabel 4.

Tabel 4. Classification Report Skenario Adam 0.001%

	Precision	Recall	F1-score	Support
<i>Brain Glioma</i>	0.92	0.92	0.91	50
<i>Brain Meningioma</i>	0.90	0.90	0.90	49
<i>Brain Normal</i>	0.96	0.98	0.97	53
<i>Brain Pituitary</i>	1.00	0.98	0.99	48
<i>Accuracy</i>			0.94	200
<i>Macro avg.</i>	0.95	0.94	0.94	200
<i>Weighted avg.</i>	0.95	0.94	0.95	200

Pada Tabel 3 dapat dilihat tingkat akurasi yang didapatkan adalah sebesar 94% dengan tingkat *recall* dan *f1 score* sebesar 94%. Sedangkan untuk *precision* hasil yang didapatkan adalah sebesar 95%. Langkah selanjutnya adalah melakukan skenario berikutnya yaitu Adam 0.0001%. Hasil *confusion matrix* dari skenario Adam 0.0001% dapat dilihat pada Gambar 6.



Gambar 6. Confusion Matrix Skenario Adam 0.0001%

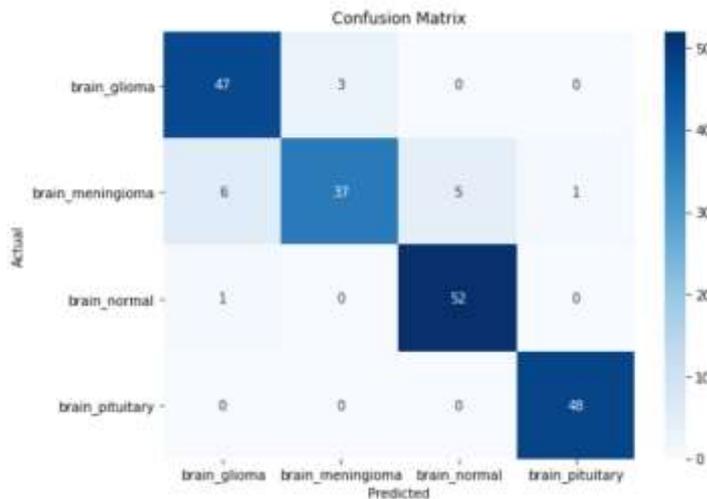
Setelah mendapatkan hasil *confusion matrix* pada Gambar 6, selanjutnya didapatkan hasil *classification report* dari skenario Adam 0.0001% dapat dilihat pada Tabel 5.

Tabel 5. Classification Report Skenario Adam 0.0001%

	Precision	Recall	F1-score	Support
--	------------------	---------------	-----------------	----------------

<i>Brain Glioma</i>	0.88	0.92	0.99	50
<i>Brain Meningioma</i>	0.87	0.84	0.85	49
<i>Brain Normal</i>	0.89	0.94	0.92	53
<i>Brain Pituitary</i>	1.00	0.94	0.97	48
<i>Accuracy</i>			0.91	200
<i>Macro avg.</i>	0.91	0.91	0.91	200
<i>Weighted avg.</i>	0.91	0.91	0.91	200

Pada Tabel 4 dapat dilihat tingkat akurasi yang didapatkan adalah sebesar 91% dengan tingkat *recall* dan *f1 score* sebesar 91%. Sedangkan untuk *precision* hasil yang didapatkan adalah sebesar 91%. Langkah selanjutnya adalah melakukan skenario berikutnya yaitu RMSprop 0.001%. Hasil *confusion matrix* dari skenario RMSprop 0.001% dapat dilihat pada Gambar 7.



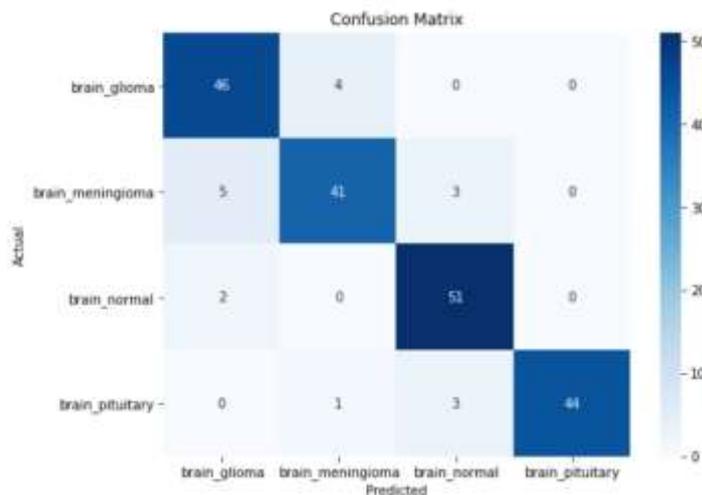
Gambar 7. Confusion Matrix Skenario RMSProp 0.001%

Setelah mendapatkan hasil *confusion matrix* pada Gambar 7, selanjutnya didapatkan hasil *classification report* dari skenario RMSProp 0.001% dapat dilihat pada Tabel 6.

Tabel 6. Classification Report Skenario RMSprop 0.001%

	Precision	Recall	F1-score	Support
<i>Brain Glioma</i>	0.87	0.94	0.90	50
<i>Brain Meningioma</i>	0.93	0.76	0.83	49
<i>Brain Normal</i>	0.91	0.98	0.95	53
<i>Brain Pituitary</i>	0.98	1.00	0.99	48
<i>Accuracy</i>			0.92	200
<i>Macro avg.</i>	0.92	0.92	0.92	200
<i>Weighted avg.</i>	0.92	0.92	0.92	200

Pada Tabel 5 dapat dilihat tingkat akurasi yang didapatkan adalah sebesar 92% dengan tingkat *recall*, *precision* dan *f1 score* sebesar 92%. Langkah selanjutnya adalah melakukan skenario terakhir yaitu RMSprop 0.0001%. Hasil *confusion matrix* dari skenario RMSprop 0.0001% dapat dilihat pada Gambar 8.



Gambar 8. Confusion Matrix Skenario RMSprop 0.0001%

Setelah mendapatkan hasil *confusion matrix* pada Gambar 8, selanjutnya didapatkan hasil *classification report* dari skenario RMSprop 0.0001% dapat dilihat pada Tabel 7.

Tabel 7. Classification Report Skenario RMSprop 0.0001%

	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	<i>Support</i>
<i>Brain Glioma</i>	0.87	0.92	0.89	50
<i>Brain Meningioma</i>	0.89	0.84	0.86	49
<i>Brain Normal</i>	0.89	0.96	0.93	53
<i>Brain Pituitary</i>	1.00	0.92	0.96	48
<i>Accuracy</i>			0.91	200
<i>Macro avg.</i>	0.91	0.91	0.91	200
<i>Weighted avg.</i>	0.91	0.91	0.91	200

Pada Tabel 6 dapat dilihat tingkat akurasi yang didapatkan adalah sebesar 91% dengan tingkat *recall*, *precision* dan *f1 score* sebesar 91%. Langkah terakhir adalah menggabungkan semua skenario yang telah dilakukan ke dalam 1 tabel. Hasil dari evaluasi akhir dapat dilihat pada Tabel 8.

Tabel 8. Evaluasi Akhir Tiap Skenario

Metode	Optimasi	<i>Learning Rate</i>	Akurasi
CNN-ViT	Adam	0.001%	94%
		0.0001%	91%
	RMSprop	0.001%	92%
		0.0001%	91%

Dari Tabel 8 dapat dilihat bahwa skenario yang mendapatkan hasil tertinggi ada pada skenario Adam 0.001% dengan tingkat akurasi 94%. Kemudian tingkat akurasi tertinggi kedua didapatkan oleh RMSprop 0.001%. Sedangkan untuk Adam dan RMSprop dengan *learning rate* 0.0001% mendapatkan tingkat akurasi yang sama yaitu sebesar 91%.

Tabel 9. Perbandingan Dengan Penelitian Terdahulu

Metode	Akurasi	Presisi	<i>Recall</i>	<i>F1 Score</i>
CNN-ViT (diusulkan)	94%	95%	94%	94%
CNN-ELM [10]	91.4%	91.5%	91.5%	91.5%
CNN(MobileNetV2) [3]	88.64%	90%	89%	89%

Pada Tabel 9 dapat dilihat bahwa metode yang diusulkan oleh penulis mendapatkan hasil tertinggi diantara penelitian terdahulu. Akan tetapi, perlu dicatat bahwa hasil yang didapatkan berbeda tergantung dataset yang digunakan, model yang digunakan dan metode pembagian data yang dilakukan. Oleh karena itu, perlu dilakukan penelitian lebih lanjut dengan parameter yang sudah disetarakan.

V. KESIMPULAN DAN SARAN

Berdasarkan penelitian yang sudah dilakukan, dapat disimpulkan bahwa optimasi dan tingkat *learning rate* yang digunakan mempengaruhi tingkat akurasi yang didapatkan pada klasifikasi. Optimasi terbaik dalam penelitian ini adalah Adam dengan *learning rate* 0.001% karena mendapatkan akurasi tertinggi yaitu sebesar 94%. Selain itu, untuk RMSprop *learning rate* 0.001% mendapatkan tingkat akurasi 92%. Sedangkan untuk Adam dan RMSprop 0.0001% mendapatkan akurasi yang sama yaitu 91%. Dapat disimpulkan bahwa optimasi yang paling optimal dalam penelitian adalah Adam dengan *learning rate* 0.001%. Untuk penelitian selanjutnya diharapkan dapat menggunakan model CNN ataupun ViT yang sudah *pretrained* sebelumnya agar bisa meningkatkan akurasi yang didapatkan lebih jauh lagi.

DAFTAR PUSTAKA

- [1] M. Kristian, S. Andryana, and A. Gunayarti, “Diagnosa Penyakit Tumor Otak Menggunakan Metode Waterfall dan Algoritma Depth First Search,” *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 6, pp. 11–24, Jun. 2021.
- [2] K. N. Deeksha, M. Deeksha, A. V Girish, A. S. Bhat, and H. Lakshmi, “Classification of Brain Tumor and its types using Convolutional Neural Network,” in *2020 IEEE International Conference for Innovation in Technology (INOCON)*, IEEE, 2020, pp. 1–6.
- [3] M. N. Winnarto, M. Mailasari, and A. Purnamawati, “Klasifikasi Jenis Tumor Otak Menggunakan Arsitektur Mobile Net V2,” *Jurnal SIMETRIS*, vol. 13, no. 2, 2022.
- [4] M. Hanindia, P. Swari, G. Ngurah, and A. Mahendra, “SISTEM PAKAR SKRINING PENYAKIT YANG DISEBABKAN OLEH VIRUS MENGGUNAKAN CERTAINTY FACTOR,” *Jurnal Mantik Penusa*, vol. 3, no. 1, pp. 196–203, 2019.
- [5] D. Bhatt *et al.*, “CNN Variants for Computer Vision: History, Architecture, Application, Challenges and Future Scope,” *Electronics (Switzerland)*, vol. 10, no. 20. MDPI, Oct. 01, 2021. doi: 10.3390/electronics10202470.
- [6] K. Mochammad *et al.*, “IMPLEMENTASI ARSITEKTUR ALEXNET DAN RESNET34 PADA KLASIFIKASI CITRA PENYAKIT DAUN KENTANG MENGGUNAKAN TRANSFER LEARNING,” 2023.
- [7] T. T. Nguyen, T. V. Nguyen, and M. T. Tran, “Collaborative Consultation Doctors Model: Unifying CNN and ViT for COVID-19 Diagnostic,” *IEEE Access*, vol. 11, pp. 95346–95357, 2023, doi: 10.1109/ACCESS.2023.3307014.
- [8] A. Vaswani *et al.*, “Attention Is All You Need,” *Advances in Neural Information Processing Systems* 30, Dec. 2017.
- [9] W. Liu *et al.*, “PHTrans: Parallely Aggregating Global and Local Representations for Medical Image Segmentation,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Mar. 2022, pp. 235–244. [Online]. Available: <http://arxiv.org/abs/2203.04568>
- [10] R. Rakhman Wahid, F. Tri Anggraeny, and B. Nugroho, “Brain Tumor Classification with Hybrid Algorithm Convolutional Neural Network-Extreme Learning Machine,” 2021.
- [11] N. Noreen, S. Palaniappan, A. Qayyum, I. Ahmad, and M. O. Allassafi, “Brain Tumor Classification Based on Fine-Tuned Models and the Ensemble Method,” *Computers, Materials and Continua*, vol. 67, no. 3, pp. 3967–3982, Mar. 2021, doi: 10.32604/cmc.2021.014158.
- [12] D. Irfan and T. Surya Gunawan, “COMPARISON OF SGD, RMSProp, AND ADAM OPTIMIZATION IN ANIMAL CLASSIFICATION USING CNNs,” 2023.
- [13] A. Peryanto, A. Yudhana, and D. R. Umar, “Rancang Bangun Klasifikasi Citra Dengan Teknologi Deep Learning Berbasis Metode Convolutional Neural Network,” 2019. [Online]. Available: <https://www.mathworks.com/discovery/convolutional-neural-network.html>
- [14] J. C. Ye, “Convolutional Neural Networks,” *Geometry of Deep Learning: A Signal Processing Perspective*, pp. 113–134, 2022.

- [15] A. Dosovitskiy *et al.*, “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,” *International Conference on Learning Representations*, Oct. 2020, [Online]. Available: <http://arxiv.org/abs/2010.11929>
- [16] J. A. Figo, N. Yudistira, and A. W. Widodo, “Deteksi Covid-19 dari Citra X-ray menggunakan Vision Transformer,” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 7, no. 3, pp. 1116–1125, Mar. 2023, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [17] O. N. Putri, “Implementasi Metode Cnn Dalam Klasifikasi Gambar Jamur Pada Analisis Image Processing (Studi Kasus: Gambar Jamur Dengan Genus Agaricus Dan Amanita),” Undergraduate Thesis, Universitas Islam Indonesia, Yogyakarta, 2020.
- [18] D. Soydaner, “A Comparison of Optimization Algorithms for Deep Learning,” *Intern J Pattern Recognit Artif Intell*, vol. 34, no. 13, Dec. 2020, doi: 10.1142/S0218001420520138.
- [19] A. Wibowo, P. W. Wiryawan, and N. I. Nuqoyati, “Optimization of neural network for cancer microRNA biomarkers classification,” in *Journal of Physics: Conference Series*, Institute of Physics Publishing, Jun. 2019. doi: 10.1088/1742-6596/1217/1/012124.
- [20] S. Ahlawat and A. Choudhary, “Hybrid CNN-SVM Classifier for Handwritten Digit Recognition,” in *Procedia Computer Science*, Elsevier B.V., 2020, pp. 2554–2560. doi: 10.1016/j.procs.2020.03.309.
- [21] D. Lovell, D. Miller, J. Capra, and A. Bradley, “Never mind the metrics -- what about the uncertainty? Visualising confusion matrix metric distributions,” Jun. 2022, [Online]. Available: <http://arxiv.org/abs/2206.02157>
- [22] H. A. P. Belangi, “Komparasi Performa Algoritma Convolutional Neural Network,” Undergraduate Thesis, Universitas Pembangunan Nasional “Veteran” Jawa Timur, Surabaya, 2023.