



ANALISIS PERFORMA ALGORITMA CONVOLUTIONAL NEURAL NETWORKS MENGUNAKAN ARSITEKTUR LENET DAN VGG16

Musthofa Galih Pradana^{1*}, Hilda Khoirunnisa²

¹Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta,

²Teknologi Rekayasa Mekatronika, Jurusan Teknik Otomasi Manufaktur dan Mekatronika,

Politeknik Manufaktur Bandung.

*musthofagalihpradana@upnvj.ac.id

Jakarta, Indonesia

Keywords:

Gender
Detection,
Convolutional
Neural
Networks,
Image
Processing,
LeNet, VGG16.

Abstract

Identifying a person's self-identity can be done by recognizing facial images, where faces can often represent a person's identity. Facial identification with technology can benefit the effectiveness efficiency and accuracy of data. This identification process can be used with the help of algorithms that will check digital images with the necessary detection results. One algorithm that can be applied in classifying and detecting gender through facial image algorithms is Convolutional Neural Networks. Convolutional Neural Network algorithms have various architectures that have advantages in each architecture. This study compared the process of identifying a person's face to obtain information in the form of gender. The models compared in this study are the LeNet model and the VGG16 model. The identification and detection process was carried out using 800 photos for data training with gender labeling data and 240 photos for testing data. A comparison of these two models is necessary to get the best final model result. The final results obtained from this study the best accuracy of both architectures was obtained in the VGG16 architecture which reached an average accuracy of 100 in several epochs compared to the VGG16 architecture at 0.925 in the 46th epoch. This is due to a Rectified Linear Unit (ReLU) on the VGG16 architecture which can minimize errors and saturation.

Article history:

Received : 22 October 2023

Revised : 19 December 2023

Accepted : 30 December 2023

Kata Kunci:

Deteksi Jenis
Kelamin,
Convolutional
Neural
Networks, Citra
Digital, LeNet,
VGG16.

Abstrak

Proses identifikasi identitas diri seseorang dapat dilakukan dengan mengenali citra wajah, dimana wajah sering dapat merepresentasikan identitas seseorang. Proses identifikasi wajah dengan teknologi dapat memberikan manfaat pada efektifitas dan efisiensi serta keakuratan data. Proses identifikasi ini dapat digunakan dengan bantuan algoritma yang akan melakukan pengecekan citra digital dengan hasil deteksi yang diperlukan. Salah satu algoritma yang dapat diterapkan dalam melakukan klasifikasi dan deteksi gender melalui citra wajah algoritma *Convolutional Neural Networks*. Algoritma *Convolutional Neural Networks* memiliki berbagai arsitektur yang mempunyai kelebihan pada masing-masing arsitekturnya. Penelitian ini melakukan perbandingan dari proses identifikasi wajah

seseorang untuk mendapatkan informasi berupa *gender*. Adapun model yang dibandingkan dalam penelitian ini adalah model *LeNet* dan model *VGG16*. Proses identifikasi dan deteksi dilakukan dengan menggunakan 800 foto untuk *training* data yang sudah memiliki data pelabelan *gender* dan 240 foto untuk data *testing*. Komparasi kedua model ini diperlukan untuk mendapatkan hasil akhir model yang terbaik. Hasil akhir yang didapatkan dari penelitian ini akurasi terbaik dari kedua arsitektur didapatkan pada arsitektur *VGG16* yang mencapai rata-rata akurasi 100 di beberapa epoch dibandingkan dengan arsitektur *VGG16* di angka 0,925 pada epoch ke 46. Hal ini disebabkan oleh keberadaan *Rectified Linear Unit (ReLU)* pada arsitektur *VGG16* yang dapat meminimalisir *error* dan saturasi.

Pendahuluan

Salah satu bagian yang dapat digunakan untuk mengidentifikasi *gender* seseorang adalah melalui identifikasi wajah. Wajah adalah bagian tubuh yang penting dan mampu merepresentasikan informasi yang sesuai dengan identitas seseorang. Informasi tersebut bisa berupa jenis kelamin seseorang, maupun juga usia yang mampu diidentifikasi dari bentuk wajah. Proses identifikasi *gender* dapat meningkatkan efektifitas dan efisiensi pengenalan, yang manfaatnya dapat berimplikasi kepada banyak hal. Contohnya untuk identifikasi kehadiran karyawan yang dengan menerapkan proses identifikasi wajah dapat membuat data kehadiran karyawan menjadi lebih akurat. Dalam melakukan identifikasi wajah, penerapan algoritma deteksi citra bisa dilakukan untuk mempermudah proses identifikasi dan pengenalan wajah seseorang berdasarkan data citra [1].

Proses identifikasi menggunakan citra digital dapat digunakan menggunakan bantuan beberapa pendekatan algoritma untuk mendeteksi citra seperti *Convolutional Neural Network*. Algoritma *Convolutional Neural Network* sendiri memiliki banyak arsitektur yang menyusun, diantaranya seperti *LeNet*, *AlexNet*, *VGG* [2]. Beberapa arsitektur yang menyusun *CNN* tentu memiliki kelebihan dan kekurangan masing-masing. Pada penelitian ini akan diujikan data pengenalan wajah menggunakan dataset wajah untuk mengidentifikasi *gender* seseorang melalui wajah. Proses identifikasi yang menggunakan algoritma *CNN* dari S. Bunrit, yang menerapkan *autoencoder* menghasilkan kinerja yang meningkat dari data yang telah dilatih

sebelumnya [3]. Penelitian dari Atul Sharma dalam melakukan klasifikasi kendaraan dengan menggunakan ukuran batch 64. Penelitian ini mendapatkan akurasi 94% untuk 3 kelas yang digunakan dalam dataset *cifar-10* [4]. Klasifikasi mengenai citra forensik pada penelitian rujukan dengan hasil akhir kerangka baru yang terbentuk adalah *FusionNET* yang arsitekturnya merupakan gabungan dari *1D CNN* dan *2DCNN* [5]. Deteksi jenis kelamin dari citra wajah pernah diterapkan oleh Dani Devito dalam tulisannya ekstraksi ciri dengan menggunakan metode *histogram of oriented gradients* dengan hasil akhir identifikasi masing-masing *gender* adalah untuk pria sebesar 97,83% dan wanita hingga 95,92% [6]. Klasifikasi wajah lainnya menerapkan Viola Jones dalam metode Jaringan Syaraf Tiruan menghasilkan akurasi terbaik di 100% dengan data latih sebanyak 37 dan data uji sebanyak 19 [7].

Deteksi *gender* melalui pola dapat diterapkan dengan *K-Nearest Neighbor* dan *Local Binary Pattern* yang menghasilkan waktu eksekusi terbaik di 63,261 detik [8]. Penanganan deteksi yang dilakukan proses manipulasi dapat diatasi dengan penggunaan mekanisme *cosine similarity* [9]. Penelitian klasifikasi dalam konteks *medical imaging dataset* oleh Agostina J. Larrazabal menghasilkan kesimpulan dalam penelitiannya bahwa terdapat penurunan kinerja yang konsisten untuk *gender* yang kurang terwakili ketika keseimbangan minimum tidak terpenuhi, dan perlu diatasi dengan penerapan algoritma dalam menangani ketidakseimbangan *gender* [10]. Selain dari sisi *gender* identifikasi citra dapat pula menghasilkan informasi usia dari data citra,

penerapan algoritma CNN dalam prediksi ini menghasilkan keberhasilan deteksi pada 7 sampel foto untuk gender, dan untuk deteksi umur memiliki beberapa kesalahan disebabkan kualitas foto yang kurang memadai [11].

Peningkatan dari algoritma CNN sendiri dapat dilakukan dalam beberapa aspek seperti akurasi lebih baik dan arsitektur menjadi lebih ringan. Peningkatan akurasi dapat dilakukan dengan meningkatkan kualitas gambar, dan ringannya arsitektur dapat ditemukan dengan menggantikan beberapa modul dalam CNN dengan blok residu terbalik [12]. Permasalahan lain dalam CNN seperti *overfitting* dapat diatasi dengan kerangka HSIC baru berdasarkan 2D-3D yang disederhanakan CNN diimplementasikan melalui kerjasama antara CNN 2D dan lapisan konvolusi 3D [13]. Perbandingan algoritma selain CNN dalam melakukan klasifikasi citra memiliki berbagai jenis, salah satunya SVM. Komparasi yang diteliti oleh Sai Yeshwanth Chaganti menyatakan akurasi terbaik dari beberapa skenario uji yang diterapkan adalah algoritma CNN dengan hasil sebesar 93,7% dan SVM sebesar 82% [14].

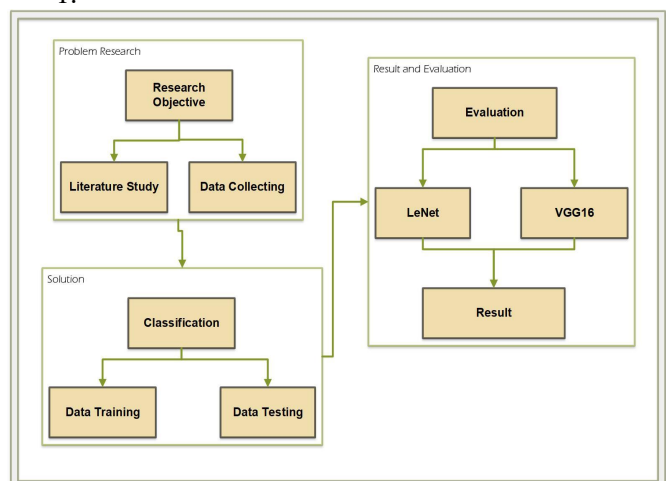
Penerapan CNN sering diterapkan pada gambar hiperspektral yang memuat informasi spasial, namun identifikasi gambar hiperspektral sering kali kurang akurat jika memiliki sample yang kecil. Hal ini dapat diatasi dengan fungsi segmentasi semantik ke dalam klasifikasi hiperspektral dalam algoritma CNN dengan *Multiscale Information Fusion* yang mendapatkan hasil akurasi klasifikasi pada masing-masing dataset dengan jumlah 3 dataset sebesar 96,09%, 97%, dan 96,56% [15]. Algoritma CNN sendiri memiliki beberapa *optimizer* yang dapat digunakan dalam proses klasifikasi, seperti ADAM dan SGDM, penelitian yang membandingkan kedua *optimizer* ini pada dataset padi menghasilkan akurasi yang lebih baik pada ADAM *Optimizer* dengan akurasi 95,33% pada epoch ke 10 [16]. Kombinasi algoritma selain CNN seperti *Multi Layer Perceptron* juga menghasilkan kombinasi yang baik seperti pada penelitian yang melakukan deteksi untuk luka akut dan kronis dengan nilai ketepatan akurasi terbaik pada 91,2% [17]. Penelitian yang akan dilakukan adalah membandingkan performa dari

arsitektur LeNet dan VGG16 dalam melakukan klasifikasi dan deteksi gender berdasarkan citra.

Metode

Adapun alur penelitian dilakukan adalah dilakukan menjadi 3 bagian untuk memudahkan proses penemuan masalah penelitian seperti berikut:

1. *Problem Research*
Pada *stage* ini dilakukan dengan melakukan studi literatur terhadap penelitian relevan yang sudah dilakukan sebelumnya dan melakukan data collecting dengan hasil data yang didapatkan dalam bentuk citra deteksi wajah beserta gender yang sudah diidentifikasi. Adapun data yang digunakan dalam penelitian ini adalah 800 foto untuk training data dan 240 foto untuk *testing*. Pembagian ini dilakukan dengan memperbanyak data training dibandingkan data testing untuk meminimalisir terjadinya *overfitting* dan menghasilkan hasil akhir yang lebih akurat.
2. *Solution*
Bagian *solution* ini memuat proses klasifikasi yang digunakan dalam melakukan *training* terhadap 800 foto data dan melakukan pengetesan terhadap 240 foto yang perlu dilakukan proses pembagian ke dalam kelas.
3. *Result and Evaluation*
Stage terakhir adalah dalam menemukan hasil percobaan. Proses yang dilakukan adalah dengan menggunakan dua arsitektur dalam CNN yakni LeNet dan juga VGG16 dengan hasil perbandingan kinerja keduanya dalam melakukan deteksi citra. Adapun detailnya ditunjukkan pada Gambar 1.



Gambar 1. Alur Penelitian

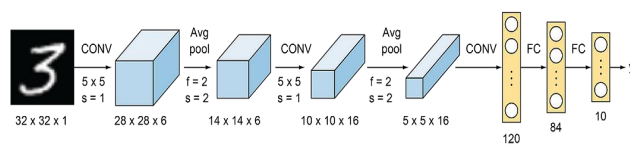
1. Convolutional Neural Networks (CNN)

CNN merupakan jenis struktur jaringan saraf tiruan yang khusus dirancang untuk tugas-tugas yang melibatkan data dalam bentuk gambar atau grid data seperti data spasial. Berikut ini adalah beberapa prinsip dasar yang membentuk dasar teori CNN : [18]

- a. *Convolution Layer* : Komponen utama dalam CNN ini digunakan untuk menemukan ciri-ciri pada gambar dengan menerapkan filter (*kernel*) pada data gambar masukan.
- b. *Pooling Layer* : Lapisan ini digunakan untuk mengurangi dimensi peta ciri yang dihasilkan oleh lapisan konvolusi. Pooling dapat dilakukan dengan metode seperti *max pooling* atau *average pooling* untuk mengambil nilai tertinggi atau rata-rata.
- c. *Fully Connected Layer* : Setelah proses konvolusi dan pooling, lapisan sepenuhnya terhubung digunakan untuk menghubungkan hasil dari lapisan-lapisan sebelumnya dalam sebuah jaringan saraf tiruan sepenuhnya terhubung.
- d. *Activation Functions* : Fungsi aktivasi seperti ReLU (*Rectified Linear Unit*) digunakan dalam CNN untuk memperkenalkan non-linearitas ke dalam model.

2. LeNet

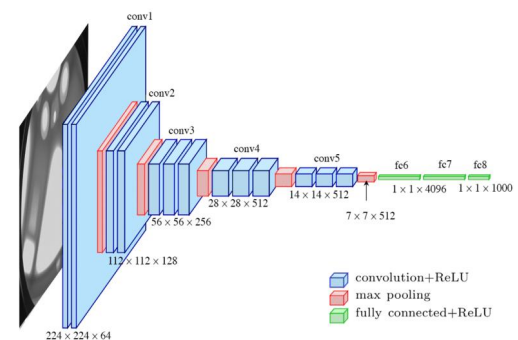
LeNet merupakan arsitektur yang dikembangkan pada tahun 1990-an oleh Yann LeCun, Leon Bottou, Yosuha Bengio, dan Patrick Haffner untuk pengenalan karakter tulisan tangan dan cetakan mesin sebagai *LeNet-5*. Arsitekturnya jelas dan mudah dipahami. Gambar masukan berupa skala abu-abu dengan dimensi 32*32*1, diikuti oleh dua pasang lapisan konvolusi dengan langkah 2 dan lapisan penggabungan rata-rata dengan langkah 1. Terakhir, lapisan terhubung sepenuhnya dengan aktivasi *Softmax* di lapisan keluaran [19]. Adapun arsitekturnya ditunjukkan pada Gambar 2.



Gambar 2. Arsitektur LeNet-5

3. VGG16

Arsitektur VGG 16 dicetuskan oleh Simonyan dan Zisserman. Desainnya terdiri dari filter konvolusional 3*3, lapisan Penggabungan Maks 2*2 dengan langkah 1, dan *padding* dijaga konstan untuk mempertahankan dimensi. Jaringan memiliki total 16 lapisan, dengan gambar input berformat RGB dengan dimensi 224*224*3, diikuti oleh 5 pasang Konvolusi (filter: 64, 128, 256.512.512) dan *Max Pooling*. Keluaran lapisan ini disuplai ke dalam tiga lapisan yang terhubung sepenuhnya dan fungsi *softmax* di lapisan keluaran. Total ada 138 Juta parameter di *VGG Net* [19]. Adapun arsitekturnya ditunjukkan pada Gambar 3.



Gambar 3. Arsitektur VGG16

Hasil

Penelitian ini melakukan hasil deteksi citra untuk mendapatkan informasi berupa gender dari seseorang. Adapun contoh data yang digunakan dalam penelitian ini diambil dari data sekunder seperti pada Tabel 1.

Tabel 1. Data Image

No	Image	Gender
1		Female

Dataset 1



Dataset 2



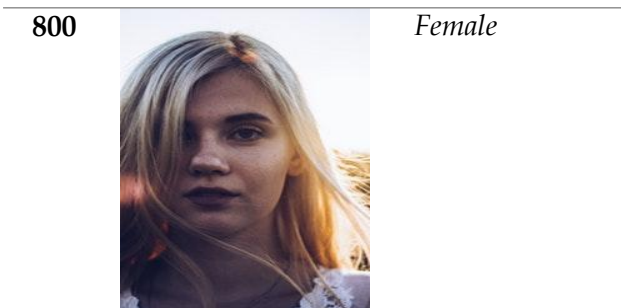
Dataset 3



Dataset 4



Dataset 5



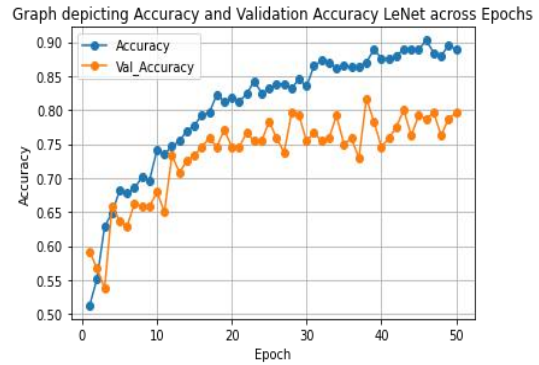
Dataset 6

Dalam skenario yang dilakukan, proses klasifikasi dan identifikasi dibagi menjadi 2 arsitektur CNN yakni *LeNet* dan *VGG16*. 2 arsitektur tersebut sama-sama dilakukan proses

epoch sebanyak 50 kali. Adapun representasi hasil dari keduanya adalah berikut ini :

1. *LeNet*

Adapun hasil akurasi dan validasi akurasi yang didapatkan dari arsitektur *LeNet* ditunjukkan pada Gambar 4.



Gambar 4. Akurasi dan Validasi *LeNet*

Secara grafik, hasil menunjukkan bahwa semakin tinggi epoch maka grafik akurasi cenderung meningkat dan mulai stabil pada kisaran epoch ke 20. Adapun hasil detail dari grafik epoch terhadap akurasi dan validasi akurasi ditunjukkan pada Tabel 2.

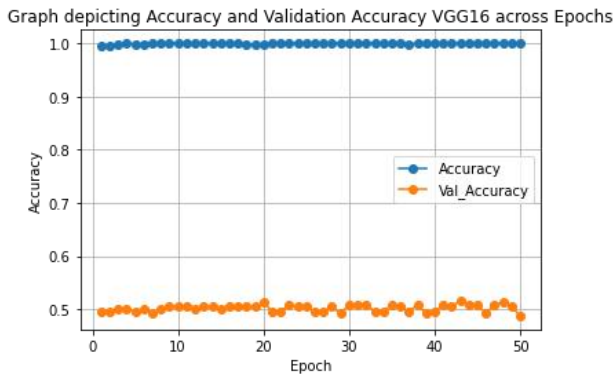
Tabel 2. Detail Akurasi *LeNet*

Epoch	Akurasi	Val_Akurasi
1	0.5125	0.5917
2	0.5525	0.5667
3	0.6288	0.5375
4	0.6488	0.6583
46	0.9025	0.7875
...		
50	0.8888	0.7958

Adapun hasil akurasi terbaik pada epoch ke 46 dengan nilai akurasi 0,925.

2. *VGG16*

Adapun hasil akurasi dan validasi akurasi yang didapatkan dari arsitektur *VGG16* ditunjukkan pada Gambar 5.



Gambar 5. Akurasi dan Validasi VGG16

Secara grafik, hasil menunjukkan bahwa kestabilan hasil dari akurasi dan validasi akurasi sangat baik untuk arsitektur VGG16 dengan hasil detail ditunjukkan pada Tabel 3.

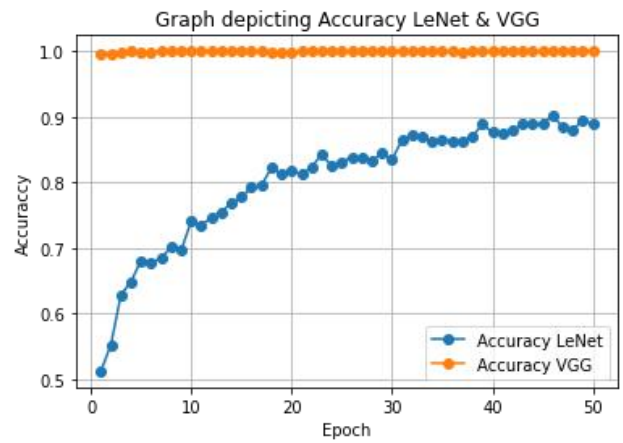
Tabel 3. Detail Akurasi VGG16

Epoch	Akurasi	Val_Akurasi
1	0.9950	0.4958
2	0.9962	0.4958
3	0.9975	0.5000
4	1.0000	0.5000
5	0.9987	0.4958
....		
50	1.0000	0.4875

Adapun hasil akurasi terbaik pada epoch ke 4 dengan nilai akurasi 1. Hasil dari akurasi model VGG16 mendapatkan nilai akurasi 1 sebanyak 40 kali dari 50 kali epoch yang dilakukan.

3. Komparasi Hasil

Hasil dari perbandingan kedua model arsitektur CNN yakni LeNet dan VGG16 divisualisasikan pada Gambar 6.



Gambar 6. Perbandingan Akurasi LeNet dan VGG

Hasil di atas menunjukkan nilai akurasi yang lebih stabil dimiliki oleh VGG16 dan LeNet yang cenderung meningkat selama proses penambahan kuantitas epoch.

Kesimpulan dan Saran

Hasil akhir dari penelitian ini yang melakukan komparasi dari 2 model arsitektur CNN dalam melakukan klasifikasi citra wajah untuk mendeteksi *gender* seseorang mendapatkan kesimpulan sebagai berikut:

- Akurasi model terbaik dari arsitektur *LeNet* adalah 0,925 yang terjadi pada *epoch* ke 46.
- Akurasi model terbaik dari arsitektur *VGG16* adalah 1,000 yang terjadi pada sejumlah *epoch*, yakni sebanyak 40 *epoch* dari total 50 kali *epoch*.
- Hasil terbaik dari kedua arsitektur didapatkan pada arsitektur *VGG16* yang mencapai rata-rata akurasi 1.000 di beberapa *epoch*.
- Perbedaan arsitektur keduanya terutama keberadaan *Rectified Linear Unit (ReLU)* pada arsitektur *VGG16* menjadikan nilai akurasi menjadi jauh lebih baik sesuai dengan fungsi *ReLU* yang meminimalisir *error* dan saturasi.

Adapun saran yang dapat dikembangkan untuk penelitian lanjutan adalah dengan melakukan optimasi pada model *VGG16* dengan meneliti pengaruh nilai *strides* terhadap tingkat keberhasilan dari sebuah model dalam melakukan klasifikasi.

Referensi

- [1] Y. Sari, *Ekstraksi Fitur dan Aplikasinya pada Citra 2D*. Lampung: Perahu Litera, 2022.
- [2] G. Nguyen *et al.*, "Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey," *Artif. Intell. Rev.*, vol. 52, no. 1, pp. 77-124, 2019, doi: 10.1007/s10462-018-09679-z.
- [3] S. Bunrit, N. Kerdprasop, and K. Kerdprasop, "Improving the representation of cnn based features by autoencoder for a task of construction material image classification," *J. Adv. Inf. Technol.*, vol. 11, no. 4, pp. 192-199, 2020, doi: 10.12720/jait.11.4.192-199.
- [4] A. Sharma and G. Phonsa, "Image Classification Using CNN," *SSRN Electron. J.*, no. Icicc, pp. 1-5, 2021, doi: 10.2139/ssrn.3833453.
- [5] I. Amerini, C. T. Li, and R. Caldelli, "Social Network Identification Through Image Classification with CNN," *IEEE Access*, vol. 7, no. c, pp. 35264-35273, 2019, doi: 10.1109/ACCESS.2019.2903876.
- [6] D. Devito, R. C. Wihandika, and A. W. Widodo, "Ekstraksi Ciri Untuk Klasifikasi Gender Berbasis Citra Wajah Menggunakan Metode Histogram of Oriented Gradients," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 8, pp. 8002-8011, 2019.
- [7] S. A. Farhan, I. J. Raharjo, and N. Kumalasari, "Identifikasi Wajah Berdasarkan Gender Dan Kelompok Usia Dengan Metode Viola Jones Dan Metode Jaringan Syaraf Tiruan Face Identification Base on Gender and Age Group With Viola Jones Method and Artificial Neuron Network Method," *Telkomuniversity*, vol. 6, no. 2, p. 8, 2019.
- [8] C. E. Magdalena, H. F. Tresna, and S. Sa'idah, "Identifikasi Gender Individu Melalui Pola Rugae Palatina Berbasis Smartphone Android Gender Types of Individuals Identification Through Rugae Palatine Based on Android Smartphone," *e-Proceeding Eng.*, vol. 8, no. 5, pp. 4991-4999, 2021.
- [9] H. Dang, F. Liu, J. Stehouwer, X. Liu, and A. K. Jain, "On the Detection of Digital Face Manipulation," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 5780-5789, 2020, doi: 10.1109/CVPR42600.2020.00582.
- [10] A. J. Larrazabal, N. Nieto, V. Peterson, D. H. Milone, and E. Ferrante, "Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 117, no. 23, pp. 12592-12594, 2020, doi: 10.1073/pnas.1919012117.
- [11] A. Arifandi, "Identifikasi dan Prediksi Umur Serta Jenis Kelamin Berdasarkan Citra Wajah Menggunakan Algoritma Convolutional Neural Network (CNN)," *J. Terap. Sains Teknol.*, vol. 4, no. 2, pp. 89-96, 2022, [Online]. Available: <https://ejournal.unikama.ac.id/index.php/jtst/article/view/6985>.
- [12] J. Qin, W. Pan, X. Xiang, Y. Tan, and G. Hou, "A biological image classification method based on improved CNN," *Ecol. Inform.*, vol. 58, no. January, p. 101093, 2020, doi: 10.1016/j.ecoinf.2020.101093.
- [13] C. Yu, R. Han, M. Song, C. Liu, and C. I. Chang, "A simplified 2D-3D CNN architecture for hyperspectral image classification based on spatial-spectral fusion," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 2485-2501, 2020, doi: 10.1109/JSTARS.2020.2983224.
- [14] N. K. Sai Yeshwanth Chaganti, Ipseeta Nanda, Koteswara Rao Pandi, Tavva GNRSN Prudhvith, "Image classification using SVM and CNN," *IEEE Xplore*, pp. 1-10, 2018, [Online]. Available: https://rpubs.com/Sharon_1684/454441.
- [15] H. Gong *et al.*, "Multiscale information fusion for hyperspectral image classification based on hybrid 2d-3d cnn," *Remote Sens.*, vol. 13, no. 12, 2021, doi: 10.3390/rs13122268.
- [16] P. H. Saputro, D. P. Wijaya, M. G. Pradana, D. L. Tyas, and W. F. Zalmi, "Comparison ADAM-optimizer and SGDM for Classification Images of Rice Leaf Disease," *Proc. - 4th Int. Conf. Informatics, Multimedia, Cyber Inf. Syst.*

ICIMCIS 2022, pp. 348–353, 2022, doi:
10.1109/ICIMCIS56303.2022.10017644.

- [17] B. Rostami, D. M. Anisuzzaman, C. Wang, S. Gopalakrishnan, J. Niezgoda, and Z. Yu, "Multiclass wound image classification using an ensemble deep CNN-based classifier," *Comput. Biol. Med.*, vol. 134, no. May, p. 104536, 2021, doi: 10.1016/j.combiomed.2021.104536.
- [18] U. Michelucci, *Advanced Applied Deep Learning: Convolutional Neural Networks and Object Detection*. Apress, 2018.
- [19] M. Elgendy, *Deep Learning for Vision Systems*. New York: Manning, 2020.