

Deteksi Ekspresi Wajah Menggunakan *Visual Geometry Group-16 Layer Convolutional Neural Network*

Rosalia Arum Kumalasanti^{*1}, Calvin Rooyen Marcellino Christy²

¹Universitas Sanata Dharma

²Universitas Sanata Dharma

E-mail: rosalia.santi@usd.ac.id¹

Abstrak. Era digital saat ini telah membawa kemajuan yang cukup pesat. Salah satu data digital yang hingga saat ini masih dikembangkan adalah dataset biometric. Biometrik merupakan data yang diambil dari fisik seseorang dan menjadi identitas yang otentik. Biometrik menjadi salah satu data yang mahal dan sensitive karena terkait dengan identitas diri. Salah satu dataset yang menarik untuk diteliti adalah ekspresi wajah. Ekspresi wajah sering digunakan dalam mengidentifikasi perasaan seseorang dan ekspresi yang tersirat. Pemahaman terkait ekspresi wajah secara komputerisasi masih minim dimanfaatkan terutama di area pelayanan masyarakat. Penelitian ini akan membawa dataset ekspresi wajah sebagai objek untuk diidentifikasi menggunakan Convolutional Neural Network (CNN). Beberapa model pada CNN telah menawarkan kemampuan yang unggul dalam melakukan identifikasi. Penelitian ini akan menggunakan model VGG-16 pada CNN. Pengujian dilakukan dengan optimizer Adam. Hasil penelitian menunjukkan bahwa pada model VGG-16 (*Visual Geometry Group-16*) dengan Adam, diperoleh optimizer Adam, mencapai akurasi terbaik yaitu 91%. Model terbaik ada pada learning rate 0,001, batch size 16, AUC 0,96, time 39m dan learning rate 0,0001, batch size 31, AUC 0,96, time 35m. Hasil ini menunjukkan bahwa VGG-16 dengan optimizer Adam paling optimal untuk deteksi ekspresi wajah dalam mengklasifikasi ekspresi senang dan takut, sehingga dapat direkomendasikan untuk pengembangan sistem pengenalan ekspresi wajah yang akurat dan handal

Kata kunci: Convolutional Neural Network; Ekspresi Wajah; VGG-16; optimizer; loss; AUC; accuracy

Abstract Abstract. The current digital era has brought quite rapid advancements. One of the digital datasets that is still being developed is the biometric dataset. Biometric data is taken from a person's physical characteristics and becomes an authentic identity. Biometrics is one of the expensive and sensitive data types because it is related to personal identity. Salah satu dataset yang menarik untuk dipelajari adalah ekspresi wajah. Facial expressions are often used to identify a person's feelings and implied expressions. Understanding of facial expressions thru computerization is still minimally utilized, especially in the area of public services. This research will use a facial expression dataset as the object to be identified using Convolutional Neural Network (CNN). Several models on CNN have offered superior capabilities in performing identification. This research will use the VGG-16 model on CNN. The testing was conducted with the Adam optimizer. The research results show that with the VGG-16 model using Adam, the Adam optimizer achieved the best accuracy of 91%. The

best models were found at a learning rate of 0.001, batch size of 16, AUC of 0.96, time of 39 minutes, and a learning rate of 0.0001, batch size of 31, AUC of 0.96, time of 35 minutes. These results indicate that VGG-16 with the Adam optimizer is the most optimal for facial expression detection in classifying happy and fearful expressions, and thus can be recommended for the development of an accurate and reliable facial expression recognition system.

Keywords: Convolutional Neural Network; Expression Image; VGG-16; optimizer; loss; AUC; accuracy

1. Pendahuluan

Era digital saat ini telah memberikan dampak yang luar biasa bagi kehidupan manusia. Salah satu perkembangan yang paling menonjol saat ini adalah pembelajaran mesin yang telah memunculkan aplikasi-aplikasi di berbagai bidang seperti pengolahan citra, pengolahan Bahasa alami, dan analisis data. Seiring dengan berjalannya waktu, pembelajaran mesin telah berkembang dan melahirkan algoritma yang mahir dalam memproses data kompleks. Salah satu data yang cukup kompleks adalah biometric. Aktivitas yang melibatkan biometrik sebagai identitas pribadi juga menjadi hal sensitive di kehidupan saat ini. Biometrik merupakan data yang diambil dari fisik atau perilaku seseorang yang dapat digunakan sebagai identitas. Tentu saja hal ini menjadi sesuatu yang krusial apabila ada yang menyalahgunakan data biometric tersebut. Biometrik memiliki dua sistem, yaitu perilaku dan fisiologis [1]. Wajah merupakan salah satu data biometrik yang sering digunakan, khususnya dalam pengenalan secara visual [2]. Sistem yang terintegrasi saat ini sedang banyak dikembangkan untuk memenuhi kebutuhan user. Salah satu sistem yang dikembangkan saat ini adalah pengenalan ekspresi wajah.

Teknologi pengenalan wajah telah muncul sebagai aplikasi utama pembelajara mesin yang sangat mempengaruhi langkah-langkah keamanan dan pengalaman pengguna. Pengenalan wajah sering dikaitkan dengan ekspresi yang tersirat. Dalam hal ini ekspresi wajah menjadi variabel penting dalam mengidentifikasi suatu keadaan yang sedang dialami seseorang. Tidak jarang ekspresi wajah ini menjadi focus penelitian yang dapat memberikan nilai lebih dari sekedar pengenalan wajah.

Penelitian ini dilakukan dalam rangka untuk memberikan pemahaman tentang sistem visual manusia yang kompleks dan untuk membedakan berbagai ekspresi wajah dengan akurasi yang tinggi. Sistem akan bekerja mulai dari deteksi, penangkapan karakter dan pencocokan ekspresi wajah. Pengenalan wajah hingga saat ini masih menjadi topik yang menarik untuk diteliti dan dikembangkan di bidang Computer Vision (CV). Computer Vision merupakan bidang ilmu computer yang bekerja pada visual, identifikasi, dan memiliki proses yang mirip seperti mata manusia [3].

Penelitian ini akan menggunakan Convolutional Neural Network (CNN) untuk mengklasifikasikan ekspresi wajah. Convolutional Neural Network (CNN) merupakan salah satu algoritma yang populer saat ini, terutama keandalannya dalam mengklasifikasikan objek berupa citra [4] [5].

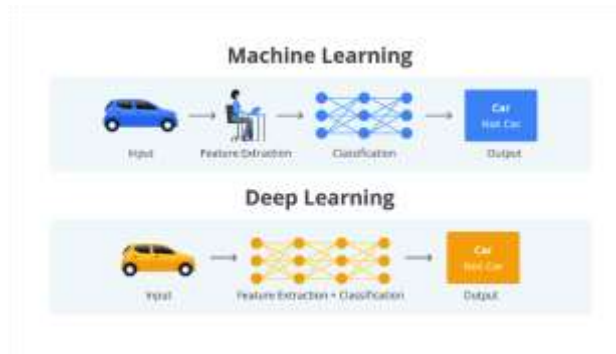
2. Metode

2.1. Ekspresi wajah

Data citra wajah sebagai dataset akan digunakan dalam penelitian ini dalam rangka untuk memberikan pengenalan ekspresi. Ekspresi wajah merupakan fenomena kompleks dimana kontur wajah dapat berubah bergantung dengan keadaan emosi seseorang. Perubahan halus pada wajah sering kali menjadi indikator kuat ketika seseorang sedang mengalami perubahan emosi, seperti bahagia, sedih marah, takut dan lain-lain. Di bidang psikologi, terdapat tujuh ekspresi dasar yang diakui secara universal, seperti takut, senang, netral, sedih, jijik, takut dan terkejut. Setiap ekspresi memiliki ciri khas tersendiri yang dapat membantu dalam identifikasi.

2.2. Deep Learning

Deep learning merupakan subbidang dari machine learning yang menggunakan arsitektur jaringan saraf tiruan dengan lapisan-lapisan yang digunakan untuk mengekstrak fitur dari data secara hierarki. Hal ini memungkinkan model mempelajari representasi data yang kompleks tanpa memerlukan ekstraksi fitur manual. Perbedaan machine learning dan deep learning dapat dilihat seperti gambar 1 di bawah ini.

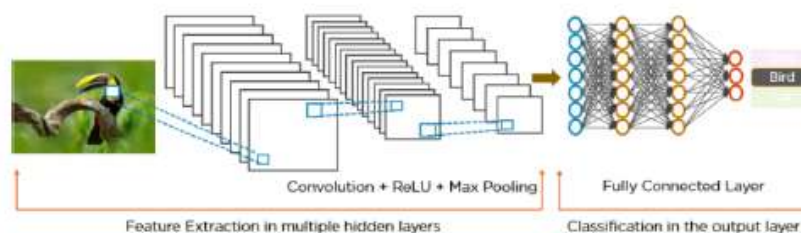


Gambar 1. Perbedaan alur Machine Learning dan Deep Learning

Usaha untuk mempercepat proses di dalam sistem, perlu menambahkan NVIDIA CUDA Deep Neural Network (cuDNN). Hal ini bertujuan untuk mengoptimalkan operasi-operasi seperti convolutional, pooling, activation functions dan normalization. cuDNN bekerja sebagai lapisan optimalisasi di atas CUDA. Pustaka ini akan otomatis diterapkan pada kernel ipynb ketika pustaka sudah diinstalasi pada framework yang akan digunakan dan device memiliki diskrit GPU (Graphic Processing Unit)

2.3. CNN

CNN memiliki sistem yang mirip dengan proses visual manusia, terutama dalam struktur optimasi yang tinggi dalam memproses karakter objek. Pengenalan objek ini tentu saja membutuhkan data latih dengan proses pembelajaran terbimbing. CNN sebagai model deep learning yang sukses dan memiliki kemampuan untuk mempelajari fitur secara hierarkis dan generalisasi yang lebih baik dari fitur yang dibuat secara manual [6]. Pada umumnya CNN terdidi dari empat komponen, yaitu Convolution Layer, ReLU Layer (Activation Function), Pooling Layer, Fully Connected Layer dan setiap komponen tersebut seperti pada gambar 2 di bawah ini [6].



Gambar 2. Struktur CNN

2.4. VGG-16

CNN dikenal sebagai metode yang efektif dalam pemrosesan citra. Beberapa model berbasis CNN yang populer dalam pengenalan ekspresi wajah adalah VGG-16 dan ResNet-50 [7].

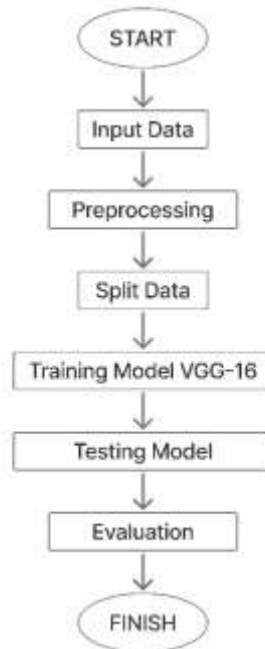
VGG-16 merupakan salah satu model vision terbaik yang memiliki convolution layer dengan filter 3x3, max pool dengan filter 2x2, memiliki 2 fully connected layer. Pengaturan convolution layer dan lapisan maxpool diterapkan secara seragam di seluruh arsitektur. Seperti Namanya, arsitektur ini memiliki 16 layer aktif dan memiliki jumlah parameter yang cukup besar yaitu 138 juta parameter [8]. Arsitektur VGG-16 dapat dilihat pada gambar 3 di bawah ini.



Gambar 3. Arsitektur VGG-16

3. Hasil dan Pembahasan

Penelitian ini akan berjalan dalam beberapa tahapan, seperti input data, preprocessing, split data, load model, training model, model evaluation, testing model dan evaluasi. Bagian ini menjelaskan kronologis penelitian, meliputi desain penelitian, prosedur penelitian, cara pengujian dan akuisisi data, dll. Langkah-langkah penelitian ini dipastikan sesuai dengan tujuan penelitian seperti yang dituliskan pada bagian Pendahuluan. Deskripsi jalannya penelitian harus didukung referensi, sehingga penjelasan dapat diterima secara ilmiah. Penggunaan gambar, tabel, dan persamaan dapat dilihat pada penjelasan berikutnya.



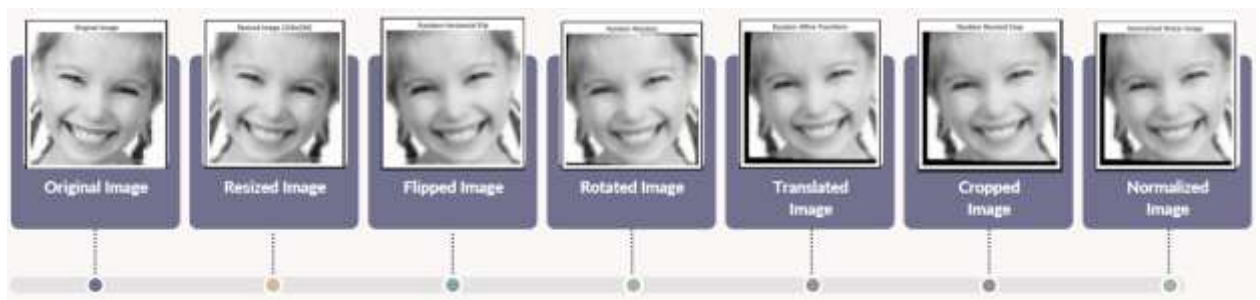
Gambar 4. Alur Penelitian

Input Data

Dataset yang digunakan pada penelitian ini diambil dari situs Kaggle, dan diakses pada tahun 2025. Data yang digunakan merupakan citra ekspresi wajah dengan total 14.110 gambar. Data dibagi pada dua label yaitu ekspresi senang dan takut. Jumlah data ekspresi senang adalah 8.989 dan ekspresi takut adalah 5.121.

Preprocessing

Langkah berikutnya setelah pengumpulan data adalah preprocessing. Preprocessing merupakan tahap untuk mempersiapkan kualitas dataset sebelum diproses lebih lanjut. Tahapan Preprocessing ini meliputi resize, augmentasi, random rotation, random Affine Transformation, Random Resized Crop, ToTensor dan Normalize. Resize untuk data input ini berukuran 256x256 piksel, yang kemudian akan dikenai augmentasi random horizontal flip dan random rotation. Dataset dibagi menjadi 3 bagian, yaitu data training, testing, dan validation. Pembagian ini bertujuan untuk memastikan bahwa hasil atau akurasi yang diberikan benar-benar sudah sesuai.



Gambar 5. Alur *Preprocessing*

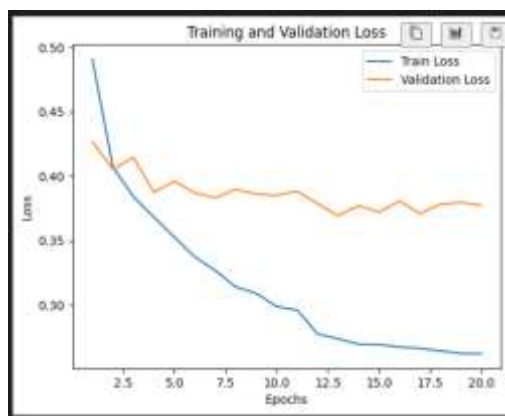
Training

Langkah berikutnya merupakan langkah utama dari penelitian ini yaitu training VGG-16. Proses convolution akan dilakukan menggunakan filter 3x3 dan stride 1. Gambar 5 di bawa ini merupakan filter yang digunakan dalam penelitian.

-1	0	1
-1	0	1
-1	0	1

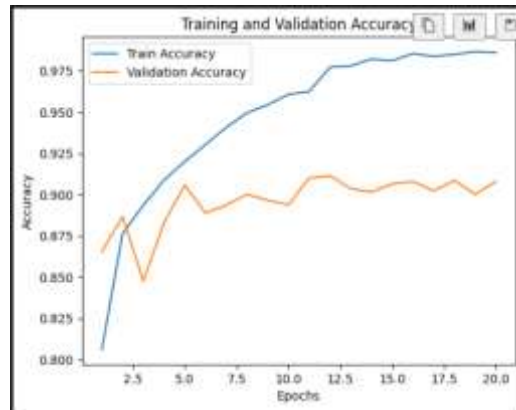
Gambar 5 . Filter 3x3

Model terbaik didapat dengan *batch size* 32, *learning rate* 0,0001, epoch 20 dengan nilai AUC 0,96. Pemodelan ini membutuhkan waktu sekitar 35 menit. Grafik Loss VGG-16 dapat dilihat pada gambar 2.



Gambar 6. Grafik Loss VGG-16, Adam

Gambar 6 menunjukkan bahwa nilai *train loss* yang turun secara stabil dari 0,48 hingga 0,2 dengan nilai epoch 20. *Validation Loss* turun dari 0,44 hingga 0,38.



Gambar 7. Grafik Akurasi VGG-16, Adam

Gambar 7 menunjukkan akurasi yang memiliki tren yang cukup baik yaitu dari angka 0,81 menuju 0,98.

Skenario pengujian ini menggunakan optimizer yaitu Adam dan didapat model terbaik. Parameter yang menjadi focus pada penelitian ini adalah learning rate, Batch size, AUC, Accuracy dan time. Tabel scenario pengujian dapat dilihat pada table 1 di bawah ini

Tabel 1 Skenario Pengujian

VGG-16				
Learning rate	Batch size	AUC	Accuracy	Time (menit)
0,0001	16	0.96	90,5%	40m
0,001	16	0.96	91%	39m
0,0001	32	0.96	91%	35m
0,001	32	0.95	90%	35m
0,0001	64	0.96	90%	96m
0,001	64	0.96	90%	113m

4. Kesimpulan

Berdasarkan hasil penelitian di atas, dapat disimpulkan bahwa VGG-16 menunjukkan performa yang baik. Skenario pengujian yang dilakukan dengan menggunakan optimizer *Adam*, mencapai akurasi terbaik yaitu 91%. Model terbaik ada pada *learning rate* 0,001, *batch size* 16, *AUC* 0,96, *time* 39m dan *learning rate* 0,0001, *batch size* 31, *AUC* 0,96, *time* 35m.

5. Ucapan Terima Kasih

Penulis yang ingin mengucapkan terima kasih atas bantuan atau dorongan dari rekan kerja, pekerjaan khusus oleh staf teknis atau dukungan keuangan dari suatu organisasi harus menyebutkannya di bagian Ucapan Terima Kasih.

Referensi

Jurnal Konstelasi menggunakan style referensi IEEE. Disarankan membuat referensi dengan tools Mendeley [1] dengan jumlah referensi minimal 15 dan 80% dari referensi tersebut harus berasal dari jurnal. Sangat

disarankan merujuk pada referensi yang berusia kurang dari 10 tahun kecuali referensi tersebut dianggap sangat penting meskipun berusia lebih dari 10 tahun. Nomor referensi diurutkan berdasarkan urutan kemunculannya dalam teks [2]. Dalam teks, referensikan referensi dengan nomor referensinya, seperti pada [3], Jangan menggunakan “Ref[3]” atau “referensi[3]” kecuali pada awal kalimat seperti “Referensi [1] menyatakan...”

Et. al digunakan jika nama penulis lebih dari enam. Paper yang belum dipublikasikan, bahkan jika sudah dikumpulkan untuk publikasi harus disitasi sebagai “unpublished”[4]. Paper yang telah diterima untuk dipublikasikan harus disitasi sebagai “in press”[5].

Referensi

- [1] S. Dhanny, D. P. Andikha, S. Kezia dan F. Jenisa, “Implementasi Convolutiional Neural Network untuk Facial Recognition,” *Media Informatika*, vol. 20, no. 2, pp. 66-79, 2021.
- [2] Q. Sun dan A. Redei, “Knock Knock, Who's There : Facial Recognition using CANN-based Classifier,” *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 1, pp. 9-16, 2022.
- [3] P. L. Prasanna, D. R. Lavanya, T. Sasidhar dan B. S. Babu, “Image Classification using Convolutional Neural Network,” *International Journal of Emerging Trends in Engineering Research*, vol. 8, no. 10, pp. 6816-6820, 2020.
- [4] F. F. Abdullah dan S. Agustin, “Penerapan Biometric Fase Recognition Menggunakan Metode Convolutional Neural Network pada Aplikasi Berbasis Android,” *Informatic and Computational Intelligent Journal*, vol. 6, no. 1, pp. 1-11, 2024.
- [5] J. Desai, U. Pedhekar, A. Sarode dan R. Ramteke, “Face Detection System using Convolutional Neural Network Review,” *International Journal of Novel Research and Development*, vol. 9, no. 4, pp. 134-43, 2024.
- [6] S. Bhairnallykar, A. Prajapati, A. Rajbhar dan S. Mujaawar, “Concolutional Neural Network (CNN) for Image Detection,” *International Research Journal of Engineering and technology (IRJET)*, vol. 7, no. 11, pp. 1239-1243, 2020.
- [7] T. Elbedary, A. Khalil, M. M. Saafan dan H. Mostafa, “Palm-print Recognition Based on Deep Residual Networks,” *International Journal of Telecommunication*, vol. 4, no. 2, pp. 1-18, 2024.
- [8] K. Simonyan dan A. Zisserman, “Very Deep Convolutional Networks for Large Scale Image Recognition,” dalam *ICLR*, New York, 2015.