

IMPLEMENTASI METODE *CONVOLUTIONAL NEURAL NETWORK* (CNN) PADA PENEGANALAN OBJEK VIDEO CCTV

Vandel Maha Putra Salawazo¹, Desta Putra Jaya Gea²,
Richard Foarota Gea³, Fadhillah Azmi⁴

*Fakultas Teknologi dan Ilmu Komputer, Universitas Prima Indonesia, Indonesia
^{1,2,3} Mahasiswa Teknik Informatika, Universitas Prima Indonesia, Indonesia
Jalan Sekip Sikambang, Medan, 20111, Indonesia

e-mail: azmi.fadhillah007@gmail.com*

Abstrak

Perkembangan teknologi saat ini di bidang *computer vision* dalam beberapa tahun ini dengan penerapan jaringan saraf *convolutional* menunjukkan kinerja yang begitu signifikan dengan memiliki tingkat akurasi yang tinggi, misalnya deteksi objek yang mana baru-baru ini memiliki restorasi gambar seperti super resolusi dan video *super resolution* (VSR) mengalami peningkatan yang signifikan berkat pembelajaran mendalam dengan tujuan membantu menghasilkan aplikasi video *visual*. Penggunaan perangkat CCTV yang sudah semakin berkembang memiliki fungsi untuk mengawasi di sekitar lingkungan, misalnya untuk keamanan, sehingga banyak diterapkan pendekatan metode, salah satunya disini metode CNN (*Convolutional Neural Network*) yang mana metode ini memiliki dapat mengklasifikasikan objek berdasarkan ciri dai objek tersebut. Maka dari itu, penerapan metode yang dilakukan dapat dianalisis berdasarkan ciri tertentu dai objek yang akan dideteksi atau dikenali.

Kata Kunci: *Accurate learning, backpropagation, neural network*

1. Pendahuluan

Perkembangan teknologi saat ini di bidang *computer vision* dalam beberapa tahun ini dengan penerapan jaringan saraf *convolutional* menunjukkan kinerja yang canggih dengan tingkat akurasi yang tinggi, misalnya deteksi objek [1].

Baru-baru ini *restorasi* gambar seperti super *resolusi* mengalami peningkatan yang signifikan berkat pembelajaran mendalam dengan tujuan membantu menghasilkan aplikasi video *visual*. [2].

Super-Resolution (SR) merupakan teknik pencitraan yang mengubah gambar beresolusi rendah (LR) menjadi gambar beresolusi lebih tinggi dimana LR sebagai input kedalam jaringan algoritma SR untuk mengeksploitasi informasi internalnya sehingga menghasilkan gambar output mirip dengan resolusi tinggi (HR).[3].

Dalam pembelajaran yang mendalam, jaringan saraf konvolusi (CNN, atau ConvNet) adalah kelas jaringan saraf yang mendalam, yang paling umum diterapkan untuk menganalisis citra visual.

CNN adalah versi resmi dari *multilayer perceptrons*. *Multi layer perceptrons* biasanya merujuk ke jaringan yang sepenuhnya terhubung, yaitu, setiap *Neuron* dalam satu lapisan terhubung ke semua *Neuron* di lapisan berikutnya. "Koneksi penuh" dari jaringan-jaringan ini membuat mereka rentan terhadap

over fitting data. Secara proses regularisasi termasuk menambahkan beberapa bentuk pengukuran bobot untuk fungsi kerugian. Namun, CNN mengambil pendekatan yang berbeda terhadap regularisasi mereka mengambil keuntungan dari pola hierarkis dalam data dan mengumpulkan pola yang lebih kompleks menggunakan pola yang lebih kecil dan lebih sederhana. Oleh karena itu, pada skala keterhubungan dan kompleksitas, CNN berada di ekstrem bawah. Mereka juga dikenal sebagai *shiftinvariant* atau *space invariant artificial neural networks* (SIANN), berdasarkan arsitektur bobot bersama dan karakteristik *invarian* terjemahan.

Jaringan *konvolusional* diilhami oleh proses biologisdalam hal pola konektivitas antara *Neuron* menyerupai organisasi korteks visual hewan. *Neuron* kortikal individual merespons rangsangan hanya di daerah terbatas bidang visual yang dikenal sebagai bidang reseptif. Bidang reseptif dari neuron yang berbeda sebagian tumpang tindih sehingga menutupi seluruh bidang visual. CNN menggunakan prapemrosesan yang relatif sedikit dibandingkan dengan algoritma klasifikasi gambar lainnya. Ini berarti bahwa jaringan mempelajari filter yang dalam algoritma tradisional direkayasa oleh tangan. Kemandirian ini dari pengetahuan sebelumnya dan upaya manusia dalam desain fitur adalah keuntungan utama. Mereka memiliki aplikasi dalam pengenalan gambar dan video, sistem pemberi rekomendasi, klasifikasi

gambar, analisis gambar medis, dan pemrosesan bahasa alami.[8].

Jaringan saraf *convolutional* terdiri dari input dan lapisan output, serta beberapa lapisan tersembunyi. Lapisan tersembunyi CNN biasanya terdiri dari serangkaian lapisan *convolutional* yang berbelit-belit dengan perkalian atau produk titik lainnya. Fungsi aktivasi umumnya merupakan lapisan RELU, dan kemudian diikuti oleh konvolusi tambahan seperti pooling layer, sepenuhnya terhubung lapisan dan lapisan normalisasi, disebut sebagai lapisan tersembunyi karena input dan output mereka ditutupi oleh fungsi aktivasi dan konvolusi akhir. Konvolusi terakhir, pada gilirannya, seringkali melibatkan *backpropagation* untuk lebih akurat menimbang produk akhir.

Meskipun lapisan-lapisan tersebut secara sehari-hari disebut sebagai *konvolusi*, ini hanya dengan *konvensi*. Secara matematis, ini secara teknis merupakan produk titik geser atau korelasi silang. Ini memiliki signifikansi untuk indeks dalam matriks, karena mempengaruhi bagaimana berat ditentukan pada titik indeks tertentu.

2. Metodologi Penelitian

Prosedur kerja yang dilakukan dalam penelitian ini adalah:

a. Observasi lapangan

Observasi lapangan adalah satu cara teknik pengambilan data dengan cara mengamati objek penelitian secara langsung dilapangan untuk memperoleh data yang aktual. Nursid Sumaatmadja (1988:105) mengemukakan bahwa Secara geografis, gejala dan masalah geografi ada dan terjadi secara langsung dilapangan. Untuk mendapatkan data geografi yang aktual dan langsung, kita harus melakukan observasi lapangan.

Teknik pengumpulan data dan observasi digunakan apabila penelitian berkenaan dengan perilaku manusia, proses kerja, gejala-gejala alam dan bila responden yang diamati tidak terlalu besar.

b. Studi kepustakaan

Penulis menggunakan teknik ini bertujuan untuk melakukan penelusuran berbagai teori, menelaah dan mengaplikasikannya terhadap pelaksanaan dan analisis penelitian. Masri Singarinbun dan Sofyan Efendi (1989:70) mengemukakan manfaat yang diperoleh dari penelusuran perpustakaan yaitu sebagai berikut :

- 1) Menggali teori-teori dasar dan konsep yang telah ditemukan oleh para ahli terdahulu, mengikuti perkembangan

penelitian dalam bidang yang akan diteliti, memperoleh orientasi yang lebih luas mengenai topik yang dipilih. memanfaatkan data skunder, dan menghindarkan duplikasi penelitian.

- 2) Mamfaat lain yang sering dilupakan adalah melalui penelusuran dan penelaahan kepustakaan, dapat dipelajari bagaimana cara mengungkapkan buah pikiran secara sistematis, kritis dan ekonomis.

c. Studi Literatur

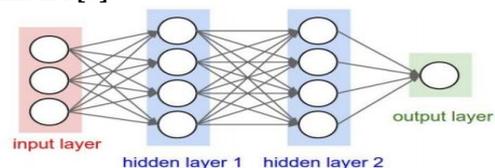
Studi literatur adalah salah satu studi yang digunakan untuk mencari dan mengkaji dasar-dasar teoritis, yaitu dengan cara mempelajari buku, majalah, Koran, laporan dokumen dan literatur lainnya yang berkaitan dengan penelitian. Berdasarkan penjelasan diatas maka penulis menggunakan teknik studi literatur ini untuk mencari dan melengkapi dasar-dasar teoritis yang diperlukan.

d. Analisis Hasil Pengujian

e. Menyusun laporan tugas akhir.

2.1. Convolutional Neural Network

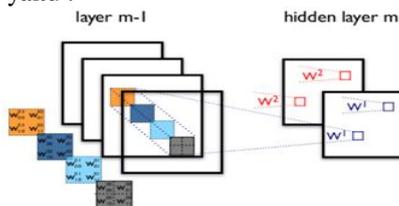
Metode dengan menggunakan *Concolutional Neuwral Network* (CNN) merupakan salah satu metode dari *machine learning* yang merupakan pengembangan dari *Multi Layer Perceptron* (MLP) yang mana dirancang untuk mengolah atau membuat data dari dua dimensi. CNN juga merupakan salah satu jenis metode dari *Deep Neural Network* yang dikarenakan didalamnya memiliki tingkat jaringan dan mempunyai banyak penerapan yang dilakukan di dalam citra. Metode CNN terdiri dari dua metode yaitu klasifikasi yang menggunakan *feedward* dan tahap pembelajarannya menggunakan *backpropagation*. Prinsip kerja dari metode ini mempunyai kesamaan pada metode MLP, tetap di dalam metode CNN setiap neuronnya disajikan dalam bentuk dua dimensi yang mana tidak sama seperti pada metode MLP yang setiap neuron hanya memiliki ukuran satu dimensi.[6].



Gambar 1. Arsitektur dari metode *Multilayer Perceptron* (MLP)

Di dalam metode MLP yang dapat dilihat pada gambar di atas yang memiliki *i layer* dengan masing-masing *layer* memiliki *ji*

neuronal. Di dalam metode MLP menerima input data pada satu dimensi dan merambatkan data tersebut ke jaringan hingga menghasilkan output. Setiap hubungan antar neuron pada dua layer yang bertetangga memiliki nilai bobot satu dimensi yang menentukan kualitas dari mode tersebut. Di setiap data input di layernya dilakukan perhitungan dengan nilai bobot yang ada, kemudian hasil perhitungan akan ditransformasikan dengan menggunakan perhitungan secara non-linear yang mana itu disebut fungsi aktivasi. Data yang dirambatkan ke dalam metode CNN adalah data dengan dua dimensi, sehingga perhitungan yang dilakukan secara linear dan dengan menggunakan parameter bobot pada CNN yang berbeda. Perhitungan dengan secara linear di dalam metode CNN menggunakan perhitungan konvolusi, dengan bobot yang tidak lagi satu dimensi saja, tetapi sudah berbentuk empat dimensi yang merupakan kumpulan dari berbagai kernel konvolusi seperti gambar berikut. Dimensi pada bobot di dalam metode CNN yaitu :

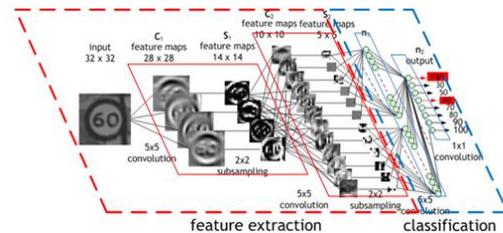


Gambar 2. Proses metode CNN

Metode CNN adalah salah satu pengembangan yang lebih lanjut dari metode MLP yang dikarenakan menggunakan metode yang mirip dengan dimensi yang lebih banyak. Di dalam metode CNN, input layer yang digunakan sebelumnya bukanlah bentuk 1 dimensi melainkan bentuk dari dua dimensi. Apabila dianalogikan dengan fitur-fitur dari wajah manusia, layer pertama adalah penggambaran goresan-goresan yang berbeda arah yang mana pada layer kedua fitur tersebut seperti bentuk mata, hidung, dan mulut mulai terlihat, proses tersebut dikarenakan penggabungan dari layer pertama yang masih berupa goresan-goresan, di dalam layer ketiga akan terbentuk kombinasi fitur-fitur mata hidung, dan mulut yang mana nantinya akan disimpulkan dengan wajah orang tertentu bahkan dapat kemungkinan sudah dapat diidentifikasi hasilnya.

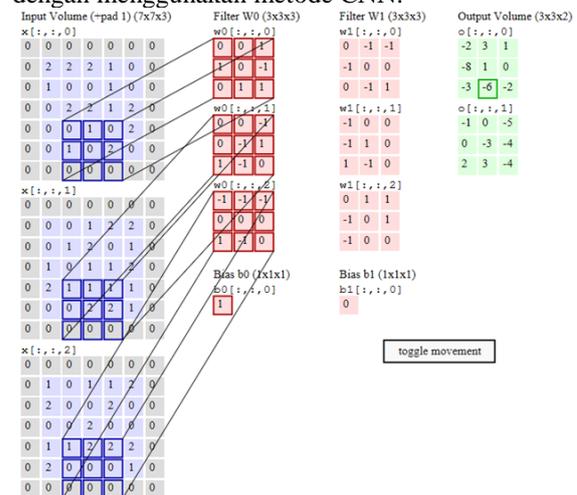
Sama halnya dengan metode neural network pada umumnya, metode CNN mempunyai beberapa hidden layer dari suatu input yang memiliki vektor tunggal. Di dalam input yang merupakan citra digital yang

dijadikan ke dalam vektor tunggal. Di dalam hidden layer, terdapat beberapa neuron yang layaknya memiliki empat fitur pemetaan yaitu C11 yang di dalam gambar tersebut. Di dalam neuron-neuron C1 dihubungkan dengan neuron di dalam S1, dan seterusnya begitu. Layer terakhir yang terhubung dengan hidden layer sebelumnya disebut dengan output layer dan disajikan hasil akhir ke dalam klasifikasi kelas.



Gambar 3. Arsitektur jaringan CNN

Berikut adalah gambaran langkah-langkah yang digunakan dalam proses perhitungan dengan menggunakan metode CNN:



Gambar 4. Proses perhitungan metode CNN

2.2. CCTV (Closed Circuit Television)

CCTV atau *Closer Circuit Television* (CCTV) pertama kalinya ditemukan oleh Walter Brunch dan digunakan oleh tim pelaksana untuk peluncuran roket V-2 yang dilakukan pada tahun 1942 di Jerman. Pertama kalinya juga diproduksi oleh perusahaan Siemens AG yang mana pada saat itu digunakan untuk memonitoring proses peluncuran roket V-2 agar dapat diketahui apakah memiliki fungsi yang baik atau tidak.[7].



Gambar 5. Model CCTV

Dulunya, hasil rekaman CCTV masih memiliki kualitas yang sangat jelek yang dikarenakan memiliki resolusi yang rendah, dan masih belum memiliki warna aliasa masih hitam putih. Pada sistem CCTV pada saat itu dihubungkan dengan kabel koaksial. Maka jika menggunakan 5 buah kamera CCTV, maka dibutuhkan 5 monitor untuk mengawasinya.

Saat ini, teknologi kamera sudah semakin pesat berkembangnya sehingga kamera pada CCTV jauh lebih baik dari tahun ke tahun dari pada sebelumnya yang mana sudah berbasis komputer dan DVR (*Digital Video Recorder*).

Adapun jenis-jenis dari CCTV saat ini adalah sebagai berikut:

1. Ptz camera
2. Dome camera
3. Bullet camera
4. Box camera
5. Board camera
6. Day/night camera
7. Spy camera
8. IP camera / network camera
9. Wireless camera
10. HD (*High Definition*) camera
11. Outdoor camera
12. Varifocal camera
13. IR (*Infra Red*) camera

3. Hasil Dan Pembahasan

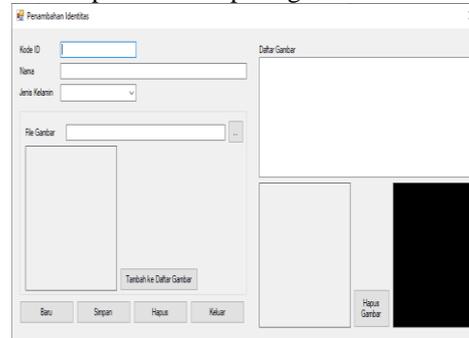
3.1. Hasil Rancangan

Tampilan awal dari aplikasi implementasi metode *Convolutional Neural Network* untuk Pengenalan Video CCTV dapat dilihat pada gambar berikut:



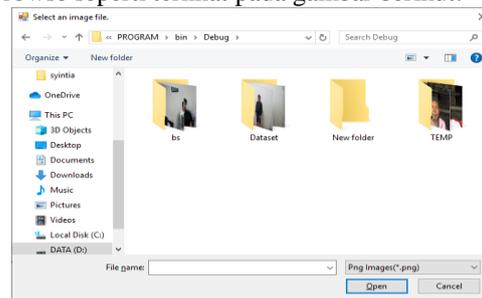
Gambar 6. Tampilan Menu Awal

Untuk melakukan proses penambahan *dataset* ke dalam sistem pengenalan wajah, maka pengguna dapat mengklik menu Dataset sehingga sistem akan menampilkan *form* Dataset seperti terlihat pada gambar berikut:



Gambar 7. Tampilan Form Dataset

Untuk menambah dataset baru, maka pengguna dapat memasukkan data kode identitas, nama identitas dan jenis kelamin. Setelah itu, pengguna dapat memilih *file input* yang akan ditambahkan untuk identitas yang bersangkutan dengan cara mengklik tombol 'Browse' sehingga sistem akan menampilkan kotak dialog browse seperti terlihat pada gambar berikut:



Gambar 8. Tampilan Form Browse

Pengguna dapat memilih *file* yang diinginkan dan mengklik tombol Open untuk membuka *file* yang dipilih. Setelah itu, klik tombol Tambah ke Daftar yang terdapat pada *form* Dataset untuk memasukkan *file* yang dipilih ke dalam daftar.

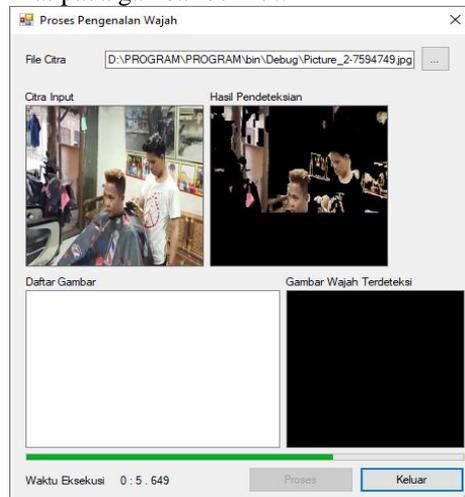
Setelah selesai memasukkan semua data yang diinginkan, maka pengguna dapat mengklik tombol Simpan untuk menyimpan data yang dimasukkan ke dalam *database*. Setelah selesai melakukan pengisian *dataset*, maka pengguna dapat mengekstrak gambar frame citra dari sebuah video. Caranya adalah dengan mengklik menu Ekstrak Video sehingga sistem akan menampilkan *form* Ekstrak Video. Pengguna dapat membuka video yang akan diekstrak, sehingga sistem akan menjalankan video yang dibuka. Pengguna dapat mengekstrak frame video yang diinginkan dan citra hasil *capture*

akan disimpan pada folder bin >> debug, seperti terlihat pada gambar berikut:



Gambar 9. Form Ekstrak Video

Setelah itu, pemakai dapat melakukan pengenalan wajah. Caranya adalah dengan mengklik menu Proses sehingga sistem akan menampilkan form Proses Pengenalan seperti terlihat pada gambar berikut:



Gambar 10. Form Proses Pengenalan Wajah

3.2. Pengujian

Pengujian akan dilakukan dengan menggunakan berbagai jenis gambar dengan jumlah wajah dan ekspresi wajah yang berbeda dalam gambar. Berikut dirincikan hasil pengujian yang dilakukan. Identitas yang terdaftar dalam *database* dapat dirincikan pada tabel 1:

Tabel 1.

Identitas yang Terdaftar dalam *Database*

No.	Identitas	Gambar Wajah
1	Richard	
2	Dest	

Hasil proses pengenalan wajah yang diperoleh dapat dirincikan pada tabel 2:

Tabel 1.

Hasil Pengujian untuk Proses Pengenalan Wajah

Citra Uji	Citra Hasil Segmentasi	Wajah Terdeteksi	Identitas	Ket
			Dest	Berhasil
			Dest	Gagal, karena wajah tidak ada di dataset, harusnya tidak dikenali
			Richard	Berhasil
			Dest	Gagal, karena wajah tidak ada di dataset, harusnya tidak dikenali
			Richard	Berhasil
			Dest	Gagal, karena wajah tidak ada di dataset, harusnya tidak dikenali
			Dest	Berhasil
			-	Berhasil
			-	Gagal

			-	Berhasil
--	--	---	---	----------

- [8] Woo, S., Jang, HD., Hwang, S & Kweon, S O. (2019). Gated Bidirectional Feature Pyramid Network for Accurate One-Shot Detection. Machine Vision and Applications. Springer. <https://doi.org/10.1007/s00138-019-01017-9>.

4. Simpulan

Setelah menyelesaikan pembuatan perangkat lunak ini, penulis dapat menarik beberapa kesimpulan, yaitu dengan menggunakan metode CNN mampu mendeteksi beberapa wajah yang terdapat pada sebuah citra digital, mengenali wajah yang terdapat pada sebuah citra digital yang merupakan citra *frame* hasil ekstrak dari video CCTV dengan tingkat akurasi sebesar 80 % untuk objek yang telah terdaftar dalam *database*, dan dapat membedakan objek yang belum terdaftar dalam *database* dengan tingkat akurasi sebesar 40 %.

5. Referensi

- [1] Feng, R., Gu, J., Qiao, Y & Dong, C. (2019). Suppressing Model Overfitting for Image Super-Resolution Network. Conference Computer Vision and Pattern Recognition (CVPR) Workshop.
- [2] Heaton, J. (2015). Artificial Intelligence for Humans: Deep learning and neural networks of Artificial Intelligence for Humans Series. Createspace Independent Publishing Platform.
- [3] Hijazi, S., Kumar, R., & Rowen, C. (2015). Image Recognition Using Convolutional Neural Networks. Cadence Whitepaper, 1–12.
- [4] Hu, F., Xia, G. S., Hu, J., & Zhang, L. (2015). Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. Remote Sensing, 7(11), 14680–14707. <https://doi.org/10.3390/rs71114680>
- [5] Katole, A. L., Yellapragada, K. P., Bedi, A. K., Kalra, S. S., & Siva Chaitanya, M. (2015). Hierarchical Deep Learning Architecture for 10K Objects Classification. Computer Science & Information Technology (CS & IT), (September), 77–93. <https://doi.org/10.5121/csit.2015.51408>
- [6] Ren, S., Girshick, R., Sun, J., He, K & Zhang, X. (2015). Object Detection Network on Convolutional Feature Maps. IEEE Transactions on Pattern analysis and Machine Intelligence. <https://doi.org/10.1109/TPAMI.2016.2601099>
- [7] Reza, A. The World of Science: <http://www.infountuksemua.weebly.com>