

Penggunaan Augmentasi Data pada *Convolutional Neural Network* (CNN) untuk Meningkatkan Akurasi Pengenalan Wajah Dosen

Maria D. Elsoin¹, Alfian Nara Weking², Dominikus Boli Watomakin³

Program Studi Teknik Informatika, Institut Keguruan dan Teknologi Larantuka, Indonesia

Corresponding Author: Maria D. Elsoin e-mail: elsoinnila@gmail.com

Published: July, 2025

ABSTRAK

Pengenalan wajah adalah salah satu bentuk penerapan teknologi kecerdasan buatan yang mempunyai peran yang sangat penting dalam berbagai bidang, diantaranya sistem pengenalan wajah dalam bidang pendidikan. Penelitian ini bertujuan untuk meningkatkan akurasi pengenalan wajah dosen pada Convolutional Neural Network (CNN) dengan menggunakan teknik augmentasi data. Teknik augmentasi yang digunakan ialah rotasi, pencahayaan, translasi, flipping, dan scaling untuk meningkatkan variasi dataset tanpa menambah data baru. Adapun arsitektur Convolutional Neural Network (CNN) yang digunakan adalah GoogLeNet (InceptionV3), dan proses pelatihan dilakukan dengan menggunakan augmentasi data dan tanpa augmentasi data. Dataset yang digunakan pada penelitian ini adalah sebanyak 40 kelas dosen, masing-masing melalui proses cropping dan resize. Hasil evaluasi menunjukkan bahwa model Convolutional Neural Network (CNN) tanpa augmentasi menghasilkan akurasi 100%, sedangkan model Convolutional Neural Network (CNN) dengan augmentasi mencapai akurasi 86.67%. Walaupun demikian, model dengan augmentasi lebih robust terhadap variasi data dan mampu menghindari overfitting. Penelitian ini menyimpulkan bahwa penggunaan augmentasi data dapat meningkatkan generalisasi model dan layak diterapkan untuk pengenalan wajah di lingkungan pendidikan.

Kata kunci : Convolutional Neural Network, Augmentasi Data, Pengenalan Wajah, GoogLeNet, InceptionV3

ABSTRACT

Face recognition is one form of application of artificial intelligence technology that has a very important role in various fields, including face recognition systems in education. This study aims to improve the accuracy of lecturer face recognition on Convolutional Neural Network (CNN) using data augmentation techniques. The augmentation techniques used are rotation, lighting, translation, flipping, and scaling to increase dataset variation without adding new data. The Convolutional Neural Network (CNN) architecture used is GoogLeNet (InceptionV3), and the training process is carried out using data augmentation and without data augmentation. The dataset used in this study was 40 classes of lecturers, each through a cropping and resizing process. The evaluation results showed that the Convolutional Neural Network (CNN) model without augmentation produced 100% accuracy, while the Convolutional Neural Network (CNN) model with augmentation achieved 86.67% accuracy. However, the model with augmentation is more robust to data variation and is able to avoid overfitting. This study concludes that the use of data augmentation can improve model generalization and is feasible to be applied to face recognition in educational environments.

Keywords : Convolutional Neural Network, Data Augmentation, Face Recognition, GoogLeNet, InceptionV3

© The Author(s). 2021 Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0>

PENDAHULUAN

Dalam bidang pengolahan citra dan kecerdasan buatan, pengenalan wajah telah menjadi subjek yang sangat penting karena memberikan solusi yang efisien dan akurat dalam berbagai aplikasi, terlebih pada bagian keamanan dan manajemen kehadiran pada bidang pendidikan seperti universitas. Dalam hal ini, pengenalan wajah dosen memiliki kemampuan untuk mempermudah proses absensi serta meningkatkan keamanan kampus (Al, 2015)

Sistem absensi manual dengan menggunakan tanda tangan pada buku absensi seperti pada kampus Institut Keguruan Dan Teknologi Larantuka, sering kali menimbulkan berbagai masalah diantaranya ketidakakuratan data kehadiran dan penggunaan waktu absensi yang tidak efisien. Dengan menerapkan teknologi sistem pengenalan wajah ini, diharapkan masalah-masalah tersebut dapat diminimalisasikan.

Berbagai metode telah dikembangkan untuk pengenalan wajah, diantaranya Metode klasik yakni Principal Component Analysis (PCA) dan Linear Discriminant Analysis (LDA) menggunakan pendekatan berbasis statistik untuk mengenali wajah (Turk & Pentland, 1991), adapun metode canggih yang telah diterapkan untuk meningkatkan akurasi pengenalan wajah, seperti Support Vector Machine (SVM) dan berbagai teknik Deep Learning, dan Convolutional Neural Networks (CNN) (Schroff et al., 2015)

Kelebihan dari metode Principal Component Analysis (PCA) dan Linear Discriminant Analysis (LDA) adalah memiliki kesederhanaan dan efisiensi komputasi, tetapi juga memiliki keterbatasan dalam penanganan variasi pose serta ekspresi wajah yang kompleks ((Zhao et al., 2019), Adapun Metode Support Vector Machine (SVM) yang merupakan metode yang lebih canggih dan dapat melakukan penanganan terhadap data non-linear, serta ketika data latih terbatas, metode ini membutuhkan pilihan fitur yang akurat serta rentan untuk mengurangi terjadinya overfitting (Liu et al.2018, 2018). Selain itu, teknik Deep Learning terutama Convolutional Neural Network, mampu mempelajari fitur yang penting secara langsung dari gambar, serta meminimalkan kebutuhan akan pra-pemrosesan yang menyeluruh (Krizhevsky et al.2012, 2012)

Dalam pengenalan wajah, penggunaan algoritma Covolutional Neural Network (CNN) tidak bebas dari masalah. Yang diperlukan oleh algoritma Convolutional Neural Network (CNN) ialah sumber daya komputasi yang tinggi dan waktu pelatihan yang lama, terutama ketika menyelesaikan dataset yang besar dan kompleks (Huang et al.2017, 2017), serta jika model Convolutional Neural Network tidak dirancang dengan baik, maka bisa menyebabkan overfitting. Perancangan yang matang dan pemilihan parameter yang tepat untuk menghasilkan hasil yang optimal dalam pengenalan wajah merupakan suatu tantangan tersendiri dalam implementasi Convolutionl Neural Network (CNN) (Srivastava et al.2014, 2014).

Beberapa solusi dapat diimplementasikan untuk mengatasi masalah ini, diataranya Augmnetasi data yang digunakan untuk memperbanyak variasi dataset pelatihan, sehingga dapat meminimalkan resiko terjadinya overfitting. Augmentasi data melibatkan beberapa teknik diantaranya rotasi, translasi, skala, serta perubahan pencahayaan pada gambar yang bertujuan untuk membuat variasi pada dataset tanpa mengumpulkan data tambahan, sehingga model Convolutional neural Network (CNN) lebih robust terhadap variasi data sebenarnya (Zhang et al.2018, 2018).

Penelitian ini bertujuan untuk mengetahui bagaimana performa model Covolutional Neural Network (CNN) dengan menggunakan augmentasi data dan juga tanpa augmentasi data, dalam pengenalan wajah dosen pada Institut Keguruan Dan Teknologi Larantuka. Fokus utama penelitian ini adalah penggunaan teknik augmentasi data pada model Convolutional Neural Network (CNN) terhadap data wajah dosen, guna meningkatkan keandalan dan efisiensi model Convolutional Neural network (CNN) terhadap data latih, dan juga mengukur akurasi Convolutional Neural Network (CNN) dalam mengenali wajah dosen dari berbagai teknik pencahayaan, pose. Oleh karena itu, penelitian ini tidak hanya sekedar bertujuan untuk memberikan kontribusi terhadap peningkatan operasional di bidang pendidikan, melainkan juga bertujuan untuk memperkaya literatur yang ada tentang implementasi kecerdasan buatan dalam operasional dunia nyata terkhususnya dalam bidang pendidikan.

Convolutional Neural Network (CNN) adalah jenis neural network yang dirancang untuk menangani data visual, seperti gambar dan video. Convolutional Neural Network (CNN) juga merupakan pengembangan dari Multilayer Perceptron (MLP) yang tergolong neural network jenis feed forward (bukan berulang). Jadi, Convolutional Neural Network ialah neural network yang didesain untuk memproses data dua dimensi. Karena kedalaman jaringan yang tinggi dan banyak diaplikasikan pada data citra, maka Convolutional Neural Network (CNN) yang juga tergolong pada jenis Deep Neural Network serta banyak pula digunakan pada data citra. (Suartika E. P, I Wayan, Wijaya Arya Yudhi, 2016).

Adapun penelitian yang relevan dengan penelitian ini adalah penelitian yang dilakukan oleh Zufar dan Setiyono (2016) dengan judul "Convolutional Neural Networks for Real-Time Face Recognition." Tujuan dari penelitian tersebut adalah untuk mengonstruksi model Convolutional Neural Networks (CNN) sebagai model pengenalan wajah manusia dari berbagai sisi, menginisialisasi parameter jaringan yang sesuai guna mengurangi kesalahan dalam proses pelatihan model CNN, serta mengimplementasikan model tersebut untuk mengenali wajah manusia secara real-time. Secara khusus, tujuan dari penelitian ini adalah untuk mengetahui performa model Convolutional Neural Network (CNN) dengan menggunakan augmentasi data dan tanpa augmentasi data dalam pengenalan wajah dosen pada Institut Keguruan dan Teknologi Larantuka. Dengan demikian, diharapkan penelitian ini dapat memberikan kontribusi yang signifikan dalam pengembangan teknologi, khususnya pada sistem pengenalan wajah di institusi pendidikan.

METODE PENELITIAN

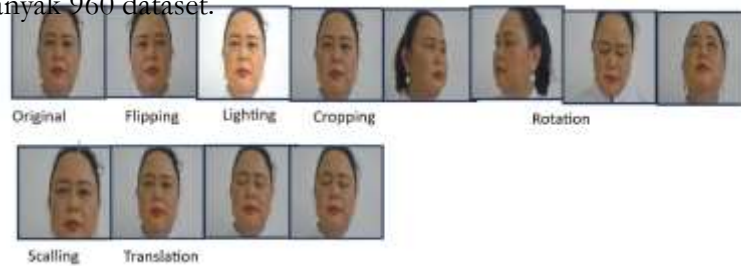
Metode pengumpulan data dalam penelitian ini dilakukan melalui beberapa cara. Pertama, dilakukan pengambilan citra wajah dosen yang telah diterapkan teknik augmentasi data untuk meningkatkan variasi dan jumlah data latih. Selain itu, peneliti juga melakukan studi literatur dengan menggunakan berbagai referensi seperti jurnal, e-book, dan hasil penelitian terdahulu yang relevan dengan topik penelitian. Studi literatur ini dimaksudkan sebagai alat perbandingan guna memperoleh hasil yang lebih valid dan mendalam. Peneliti juga memperluas pencarian referensi dengan menelusuri sumber-sumber tambahan dari internet untuk mendukung keakuratan dan efisiensi hasil penelitian. Dalam proses pengumpulan data citra, digunakan kamera Nikon D3200 dengan resolusi tinggi sebesar 6016 x 4000 piksel. Namun, karena adanya penyesuaian otomatis dari perangkat, resolusi awal citra yang digunakan sebelum proses pemotongan (cropping) adalah 1155 x 768 piksel. Dari citra awal tersebut, kemudian dilakukan proses cropping untuk mengambil fokus area wajah dengan ukuran 224 x 224 piksel, yang merupakan ukuran standar pada arsitektur model Convolutional Neural Network (CNN) seperti GoogleNet. Ukuran ini dipilih untuk mempermudah proses pelatihan model. Setelah itu, dilakukan tahap pengubahan ukuran (resize) menjadi 160 x 160 piksel agar sesuai dengan arsitektur CNN yang digunakan dalam penelitian ini.

Metode penelitian yang diterapkan dalam studi ini terdiri dari beberapa tahapan, yaitu plan, analyze, design, dan test and refine. Pada tahap perencanaan (plan), peneliti merumuskan masalah dan tujuan dari penelitian. Selanjutnya, pada tahap analisis (analyze), dilakukan pengumpulan data citra wajah dosen yang telah diterapkan teknik augmentasi seperti pencahayaan, rotasi, translasi, flipping, scaling, dan cropping untuk meningkatkan keragaman data latih. Tahap berikutnya adalah desain (design), yaitu perancangan arsitektur model Convolutional Neural Network (CNN) dengan menggunakan platform TensorFlow/Keras. Terakhir, pada tahap pengujian dan penyempurnaan (test and refine), dilakukan proses pelatihan dan pengujian model menggunakan data yang telah diaugmentasi. Pada tahap ini juga dilakukan pengukuran akurasi dan evaluasi kinerja model agar sistem pengenalan wajah dosen dapat berfungsi secara optimal dan andal.

PEMBAHASAN

Dataset yang digunakan ialah wajah dosen. Pada proses pengambilan gambar wajah dosen itu juga secara langsung menggunakan teknik augmentasi data yang bertujuan untuk menghasilkan data gambar wajah dosen yang lebih banyak dan bervariasi, sehingga model dapat lebih *robust*. Teknik pengambilan gambar wajah dosennya sesuai dengan teknik augmentasi, dengan rincian satu orang

dosen dengan 1 variasi dengan 2 kali pemotretan, jadi untuk 12 variasi sebanyak 24 kali pemotretan. Misalnya variasi pencahayaan (cahaya terang sebanyak dua kali, cahaya redup sebanyak 2 kali), rotasi (dari arah depan 2 kali, kiri dua kali, kanan dua kali, atas dua kali, bawah dua kali). Jadi total gambar pengambilan gambar wajah dosen yang berjumlah 40 orang dengan menerapkan teknik augmentasi data adalah sebanyak 960 dataset.



Gambar 1. Data Wajah Dosen Dengan Teknik Augmentasi

1. *Preprocessing Data*

Preprocessing data merupakan suatu langkah pengolahan data untuk menghasilkan kualitas citra yang bagus. *preprocessing data* disini terdiri dari 2 langkah yakni melakukan *cropping* untuk mendapatkan ukuran atau bentuk baru dengan cara mengubah atau memotong sebagian gambar. Setelah proses *cropping* dilakukan, dilanjutkan dengan proses *resize* yakni hanya mengubah ukuran tetapi tetap mempertahankan seluruh gambar. Tahapan *cropping* dan *resize* dapat dilihat pada gambar berikut:

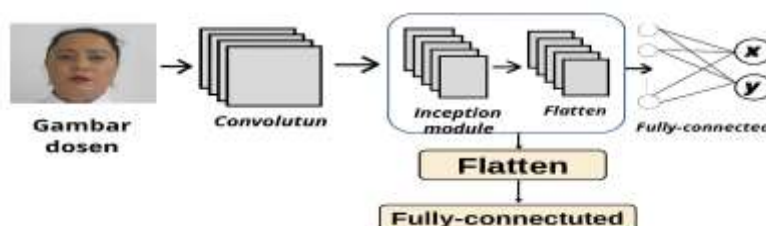


Gambar 2. Hasil *Preprocessing* dan *Resized*

2. *Desain Model Arsitektur Convolutional Neural Network*

Dalam penelitian ini model arsitektur CNN yang digunakan adalah *GoogLeNet (InceptionV3)*. Arsitektur ini kemudian akan dilatih menggunakan data citra wajah dosen dengan menerapkan teknik augmentasi data, seperti perubahan pencahayaan dan translasi posisi.

Selama pelatihan, digunakan *optimizer Adam* dengan laju pembelajaran (*learning rate*) sebesar 0,001, serta fungsi kerugian *categorical cross-entropy* untuk meminimalkan kesalahan klasifikasi. Proses pelatihan juga menggunakan data validasi untuk memantau kinerja model secara bertahap dan mencegah *overfitting*, sehingga model tetap mampu melakukan generalisasi terhadap data baru.



Gambar 3. Arsitektur *Convolutional Neural Network* secara Umum

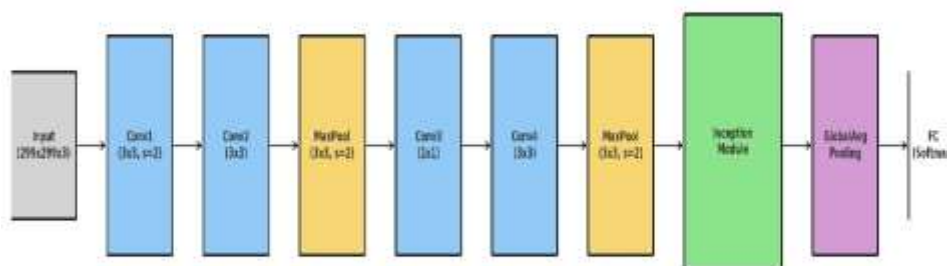
3. Pembagian Data

Setelah dilakukan preprocessing, dataset itu kemudian dibagi menjadi 3 bagian untuk membantu selama proses pelatihan dan evaluasi. Dataset pelatihan terdiri dari 70% yang kemudian digunakan untuk mengoptimalkan parameter model. Sedangkan dataset validasi terdiri dari 15%, yang nantinya akan digunakan selama proses pelatihan untuk memantau kinerja dan untuk mencegah *overfitting*, dan untuk dataset pengujian juga sebesar 15% yang digunakan untuk menguji kinerja akhir model pada wajah.

4. Penerapan *Convolutional Neural Network* Pada Model

Arsitektur model *CNN* yang digunakan dalam penelitian ini mengacu pada *GoogLeNet (InceptionV3)*, yang dimulai dengan input citra wajah dosen berukuran 299×299 piksel dalam format RGB. Gambar ini diproses melalui beberapa lapisan *convolutional* yang berfungsi untuk mengekstraksi fitur visual dasar seperti tepi, tekstur, dan pola pencahayaan. Setelah itu, fitur-fitur tersebut diringkas melalui proses *max pooling* untuk mengurangi dimensi dan kompleksitas komputasi tanpa kehilangan informasi penting.

Ciri utama *InceptionV3* terletak pada *Inception Module* yang menggabungkan beberapa jalur konvolusi paralel dengan ukuran kernel berbeda (1×1 , 3×3 , 5×5) dan *pooling*, sehingga mampu menangkap fitur multiskala secara simultan. *Output* dari modul ini kemudian diratakan dan diteruskan ke *fully connected layer* yang bertugas melakukan klasifikasi wajah ke dalam kelas dosen yang sesuai model dilatih menggunakan fungsi aktivasi *ReLU* dan *softmax* pada *layer output*, dengan *loss function categorical cross-entropy* untuk meminimalkan kesalahan prediksi selama proses pelatihan.



Gambar 4. Arsitektur *GoogleNet InceptionV3*, untuk Pengenalan Wajah

5. Proses Pelatihan Model

Model *CNN* dilatih menggunakan *optimizer Adam* dengan *learning rate* 0,001, yang bertugas menyesuaikan bobot secara bertahap agar kesalahan klasifikasi (*loss*) semakin kecil. Proses pelatihan dilakukan selama beberapa *epoch*, di mana setiap epoch model belajar dari seluruh data latih.

Selama pelatihan, digunakan data validasi untuk memantau kinerja model pada data yang belum pernah dilihat sebelumnya. Hal ini penting untuk mencegah *overfitting*, yaitu kondisi di mana model terlalu cocok terhadap data latih namun gagal mengeneralisasi pada data baru.

25/25 [=====] - 159s 6s/step - loss: 3.8226 - accuracy: 0.0623 - val_loss: 3.1400 - val_accuracy: 0.1495

Epoch 2/30

25/25 [=====] - 153s 6s/step - loss: 3.2663 - accuracy: 0.1777 - val_loss: 2.6439 - val_accuracy: 0.3098

Evaluasi Model

a. Akurasi Model

Hasil akurasi dari kedua model dapat dilihat pada tabel berikut:

Tabel 1. Hasil Evaluasi Model Augmented dan Non_Augmented

Model	Jumlah Data Uji	Akurasi (%)
Dengan Augmentasi	120	86,67
Tanpa Augmentasi	80	100

Model tanpa augmentasi menunjukkan akurasi sempurna (100%), sedangkan model dengan augmentasi menunjukkan akurasi 86.67%. Namun, hasil ini harus dianalisis lebih dalam karena akurasi tinggi pada model tanpa augmentasi bisa jadi disebabkan oleh *overfitting* terhadap data latih, sedangkan model dengan augmentasi memiliki kemampuan generalisasi yang lebih baik, terlihat dari variasi data yang lebih kompleks.

b. Presisi dan Recall

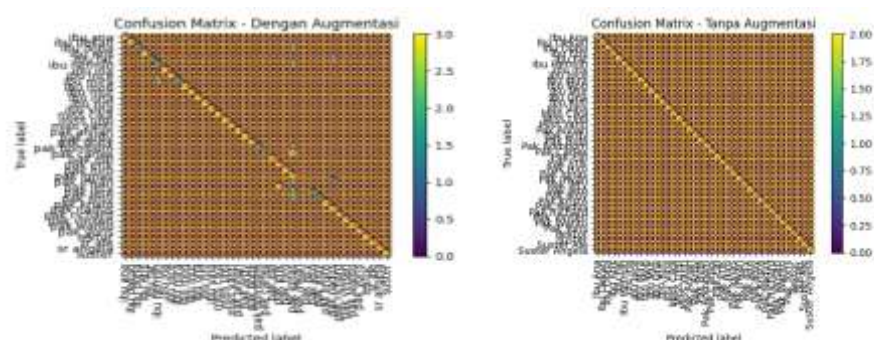
Selanjutnya menghitung metrik presisi, recall.

Tabel 2. Metrik Presisi, dan Recall

Nama Kelas	Precision	Recall	F1-Score	Support	Model
ibu ana	1	1	1	3	Dengan Augmentasi
ibu dona	1	0,67	0,8	3	Dengan Augmentasi
pak ignas	0,25	1	0,4	3	Dengan Augmentasi
pak jery	0	0	0	3	Dengan Augmentasi
Ibu Ana	1	1	1	2	Tanpa Augmentasi
Pak Jery	1	1	1	2	Tanpa Augmentasi
Pak Ignas	1	1	1	2	Tanpa Augmentasi
Suster Angela	1	1	1	2	Tanpa Augmentasi

c. Confusion Matrix

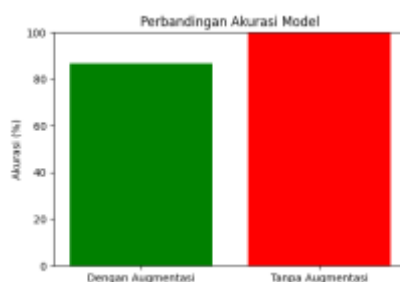
Untuk melihat prediksi kesalahan, maka digunakan Confusion Matrix.



Gambar 5. Metrik Confusion Matrix dengan Augmentasi dan Tanpa Augmentasi

d. Grafik Perbandingan Akurasi

Grafik perbandingan akurasi model CNN dengan augmentasi dan tanpa augmentasi:



Gambar 6. Perbandingan Akurasi Model CNN

KESIMPULAN

Penggunaan augmentasi data seperti pencahayaan, rotasi, scalling, dan translasi dalam penelitian ini menunjukkan bahwa berhasil meningkatkan variasi dataset wajah dosen dan membantu model CNN untuk belajar. Dan model yang dilatih tanpa augmentasi memang menunjukkan akurasi 100%, namun memiliki risiko overfitting yang tinggi karena hanya belajar dari data yang terbatas. Sebaliknya, model dengan augmentasi walaupun memiliki akurasi lebih rendah (86.67%), mampu menggeneralisasi data lebih baik. Dengan demikian, augmentasi data merupakan teknik penting dalam pelatihan model CNN untuk pengenalan wajah, karena mampu meningkatkan generalisasi model dan mengurangi ketergantungan terhadap data pelatihan terbatas.

DAFTAR PUSTAKA

- Al, P. J., Taigman, Y., Yang, M., & Ranzato, M. (2015). Deep face recognition. *British Machine Vision Conference (BMVC)*.
- Deng, J., Guo, J., Xue, N., & Zafeiriou, S. (2019). ArcFace: Additive angular margin loss for deep face recognition. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4690–4699.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4700–4704.
- Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25, 1097–1105.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single shot multibox detector. *European Conference on Computer Vision*, 21–37.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, real-time object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 779–788.
- Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 815–823.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(1), 1929–1958.
- Suartika, I. W. E. P., & Wijaya, S. R. A. Y. (2016). Klasifikasi citra menggunakan Convolutional Neural Network (CNN) pada Caltech 101. *Jurnal Teknik ITS*, 5(1), A76–A80. <http://repository.its.ac.id/48842/>
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1–9.
- Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1), 71–86.
- Zhang, X., Zou, J., He, K., & Sun, J. (2018). Accelerating very deep convolutional networks for classification and detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(10), 1943–1955.
- Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. (2019). Object detection with deep learning: A review. *IEEE Transactions on Neural Networks and Learning Systems*, 30(11), 3212–3232.
- Zufar, M., & Setiyono, B. (2016). Convolutional Neural Networks untuk pengenalan wajah secara real-time. *Jurnal Sains dan Seni ITS*, 5(2), 2337–3520. <https://repository.its.ac.id/72552/1/1212100076-Undergraduate%20Thesis.pdf>.
- Krizhevsky, A., Sutskever, I.,