Convolutional Neural Network for Colorization of Black and White Photos

Andrew Fiade [1]**, M. Ikhsan Tanggok** [1]**, Rizka Amalia Putri** [1]**, Luigi Ajeng Pratiwi** [2]**, Siti Ummi Masruroh** [1]
[1]UIN Syarif Hidayatullah, Jakarta, Indonesia
[2]Advanced Institute of Science & Technology, Daejeon, South Korea

Article Info

ABSTRACT

Nowadays, people tend to capture moments by taking photographs. Different photo features are utilized to record diverse information that people want to preserve. However, photos with digital images in black and white produce suboptimal information, necessitating an image processing approach to convert them to color photos. To address this issue, the author seeks to alter black and white photos into color photos utilizing a Convolutional Neural Network (CNN) method. The study employs Atlas 200 DK hardware and Ascend 310 processor. The data comprises 32 black and white photos in .jpg format as training data. Six experimental situations are conducted with varying numbers of black and white photos in each trial, with 81 black and white photos used for experimentation. The models generated by this study successfully produce color photos with suitable color results, anticipating the potential color of each pixel in the image. This research suggests that the trend of artificial intelligence can be employed to modify photo colors based on color prediction.

Corresponding Author:

Siti Ummi Masruroh, +6285714072654,
Faculty of Sains and Technology, Department of Informatics Engineering,
UIN Syarif Hidayatullah, Jakarta, Indonesia,
Email: ummi.masruroh@uinjkt.ac.id.

1. INTRODUCTION

An image plays an important role in human life today, such as documentation, authentic evidence of an event, and the disclosure of one's expression. Due to the large amount of information that can be retrieved from an image, Artificial Intelligence plays a role in image processing. Image processing is a method performed on an image to obtain more perfect results so that the information in the image is seen more clearly [1].

Artificial Intelligence works according to an algorithm provided by a computer system at the time of manufacture. The algorithm contains a framework from AI to process various types of data. Artificial intelligence (AI) was developed by combining computer science, logic, biology, psychology, philosophy, and many other fields. It has produced astounding results in fields like speech recognition, image processing, natural language processing, automatic proof theorems, and intelligent robots. [2]. Deep learning applications have demonstrated outstanding performance across various application domains, particularly in image classification, segmentation, and object detection [3]. In contrast to conventional machine learning, which requires a special method for feature extraction before processing it on the network model, deep learning performs feature extraction at the network layer[4].

Photos with black and white color can be changed by image processing. This process is done to improve image quality, one of which is improving the quality of contrast and color to obtain optimal information. This research tries to convert black and white photos into color photos by classifying and predicting the color of each pixel through image processing using the Convolutional Neural Network (CNN) method.

The Convolutional Neural Network (CNN) method is the most widely used image processing method. CNN is a development of Multi-Layer Perceptron (MLP) and is one of the algorithms from Deep Learning. The Convolutional Neural Network method has the most significant results in digital image recognition. This result is because CNN is implemented based on an image recognition system in the human visual cortex [5]. To increase classification accuracy, CNN-based algorithms can directly extract "learned" features from the relevant problem's raw data. Contrast this with traditional Machine Learning (ML) techniques, which frequently undergo preprocessing processes before using fixed and hand-crafted features that are not only unsatisfactory but frequently demand a high level of computational complexity. [6]. Research [7] shows that CNN can perform vehicle color recognition by 73.00% for training and 76.94% for image data testing. These results indicate that the model can recognize the color of the vehicle quite well. Research [8–10] shows that CNN can classify fruit, control quality, and detect fruit with an accuracy rate of more than 96%.

There are known algorithms for classifying images other than CNN, namely K-Nearest Neighbors (KNN) and Support Vector Machine (SVM). K-Nearest Neighbor, or KNN as it is sometimes abbreviated, is a technique that uses a supervised algorithm to classify fresh testing data according to the vast majority of classes in the KNN. This algorithm's goal is to categorize new objects using their properties and training data [11, 12]. Meanwhile, Support Vector Machine (SVM) is a supervised learning method used for classification. The SVM classification model works by trying to separate it from each class or label by making the margins as wide as possible. So SVM is a supervised learning method commonly used for Support Vector Regression or Support Vector Classification classification [13, 14].

Research [15] shows that with many data, it produces an accuracy rate of 82%. Researchers tried again using the CNN algorithm, and the same number of datasets resulted in an accuracy rate of 93.57%. This stands as a testimony to the increased potential of deep learning techniques over the more traditional machine learning techniques. Another study [16, 17], which made comparisons in image processing techniques, showed that CNN was the model with the best performance.

Because CNN is a method of developing Multi-Layer Perceptron (MLP), it has a similar way of working; it is just that in CNN, each neuron is represented in 2 (two) dimensions, while in MLP, it is only one dimension. Therefore, CNN has a greater amount of dimensions compared to MLP [16].

Based on previous research using various algorithms in classifying images, this research will use the CNN algorithm to predict the possible colors of each pixel and uses the Atlas 200 DK hardware and Ascend 310 processor in this study. This research not only classifies images but also changes the image's color. This conversion is based on an image that has no color and is taken from the internet, which will be processed to produce a color image. The model will be made to recognize objects in photos as well as real colors that match these objects by classifying and predicting colors according to the objects in the photo so that they can change colors from black and white to images with colors that match the predictions.

This study consists of 4 discussion sections. Section 1 contains an introduction to the background and research to be conducted. Section 2 contains the methods used in the research, section 3 is the analysis stage and the results obtained, and section 4 contains the conclusions of this research.

2. RESEARCH METHOD

2.1. Method of Collecting Data

In this study, the method used to collect data is literature study and observation. The researcher collects data and information from books, journals, or relevant written works to add references according to the problem being studied so that they can assist the writer in completing the research. In addition to the literature study, the authors also obtained the desired image data via the internet. After getting the desired image, the image will be processed for dataset purposes during training and testing on the system. Figure 1 shows the research flow.
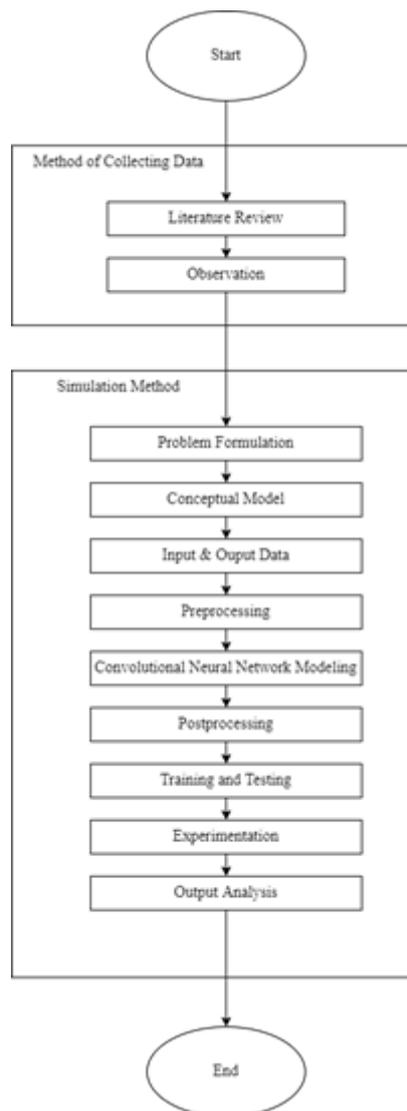

Figure 1. Research flow

2.2. Simulation Method

1. *Problem Formulation*: In this study, the author focuses on black and white photos and how black and white photos are converted into appropriate color photos by classifying and predicting the possible colors in each pixel to produce visually attractive images. To make these changes, the authors use the Convolutional Neural Network method as an implementation model, which consists of several convolution layers and is suitable for use on and using the Ascend 310 processor in this study.

2. *Conceptual Model*: In this stage, the conceptual model of research is described by designing a physical topology, namely, carrying out processes from a laptop that runs applications and programs that have been built from a laptop. This study uses the Atlas 200 DK hardware and Ascend 310 processor, and the output will be sent back to the laptop, as shown in Figure 2.
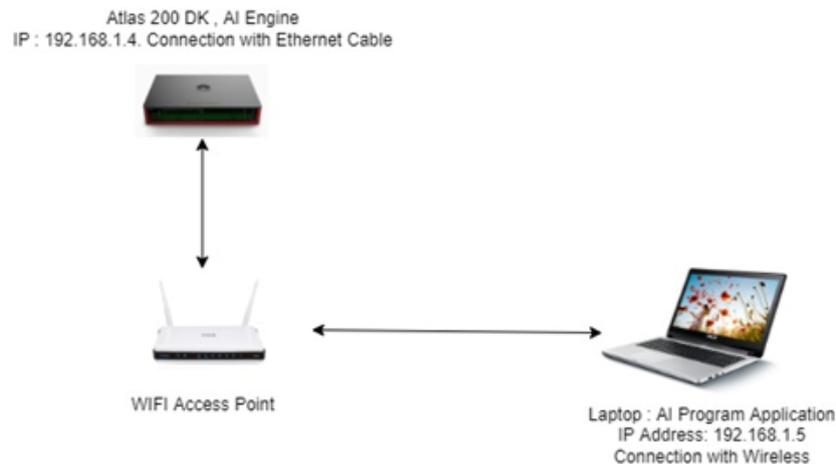

Figure 2. Physical topology

3. *Input dan Output Data*: Photos that have been obtained based on what has been described in the research methodology are collected and will be used as data in the input process. The collected photos are in .jpg format (Join Photographic Group) and put in a file called "data" in the colorization folder in the ascend project. The photos that become the dataset will go through a color-changing process which is carried out using the Convolutional Neural Network (CNN) model by passing through several convolution layers. The resulting output is a color photo according to the color prediction on the photo object.

4. *Preprocessing*: In this stage, the photos that have been collected are named according to the object or theme of the photo and then entered into a data file located in the colorization folder. Photos will undergo a process of resizing according to the size set in the system.

5. *Convolutional Neural Network Modeling*: The Convolutional Neural Network method has the most significant results in digital image recognition. This result is because CNN is implemented based on an image recognition system in the human visual cortex [5]. The workings of CNN itself are transforming the original image that has been obtained into several layers of image pixel values for classification. By adjusting the parameters, the convolutional layer uses filters as a feature extractor to extract high-level features such as horizontal or vertical edges from the input image [17]. The input, hidden, and output layers are typically the three layers of a convolutional neural network. The hidden layer is a neuron layer with a complex multi-layer nonlinear structure that comprises a convolution layer and a sub-sampling layer. The input layer is the original image without any changes. The output layer is the result of classifying the features. [18]

In this study, images that have been collected in one folder and underwent resizing will enter the CNN stage by passing through the four main layers, which will be explained later, and undergo a feature extraction process up to the classification stage.

CNN utilizes the convolution process by moving a convolution kernel (filter) with a certain size to an image. CNN consists of 4 (four) main layers, namely the convolution layer, the activation layer, the pooling layer, and the fully connected layer. In its application, CNN utilizes the convolution process so that it can only be used on two-dimensional data such as images and processes it into several layers. Several stages were carried out in building the CNN model, among others:

    a. *Convolution Layer*

        The convolution layer is the main process of CNN. This layer applies filters consisting of neurons and is arranged like the pixel values in the photo [19]. This layer employs weighted filters to separate characters from objects like colors, curves, or edges [20]. The purpose of convolution on photos is to produce a feature map which will later be used in the activation layer.

    b. *Activation Layer*

       The feature map obtained will be included in the activation function in the Activation Layer. The values contained in the feature map will be changed in a certain range according to the activation function. Based on this, the activation map representing higher-level features will be the outcome of another convolution operation [21]. This study uses the ReLU activation function.

    c. *Pooling Layer*

       The pooling Layer is used to reduce the dimensions of the feature map without losing important information in it, commonly known as downsampling. The output of a cluster of neurons at one layer is combined with another single neuron in the following layer by this layer [22]. The pooling layer is used to reduce overfitting and speed up computation because fewer parameters need to be changed after passing through it [23].

    d. *Fully Connected Layer*

       The preceding step is fed into the Fully Connected layer to determine which features are most associated with a certain class. This layer requires a long training time. The results obtained from the pooling layer will be used in the fully connected layer. In this layer, the data obtained from the activity of the previous layer is converted into one-dimensional data. After transforming, the data will be processed so that it can be classified. The sigmoid activation function follows the Fully Connected Layer and completes the classifier function [24]. For classification purposes, fully connected layers "flatten" the network's 2D spatial information into a 1D vector that represents image-level features [25].

6. *Postprocessing*: At the postprocessing stage, photos that have undergone processing in the CNN model will be resized again according to the original size of the input photo before being changed in the preprocessing stage.

3. RESULT AND ANALYSIS

3.1. Training and Testing

    At this stage, the photos that have been collected will be tested on the model that has been made. The testing process uses 32 black and white photos measuring 16x32 pixels as the test data and produces 32 color output photos. Figure 3 shows an example of a photo that became test data from this study and has not been modified:

Figure 3. Testing sample

    To produce output in the form of result data obtained from the test data, researchers use the Conv2D layer as a layer in the image processing process which is used to get filters from images or photos. Figure 4 shows the process for the Convolutional Neural Network algorithm in this study:
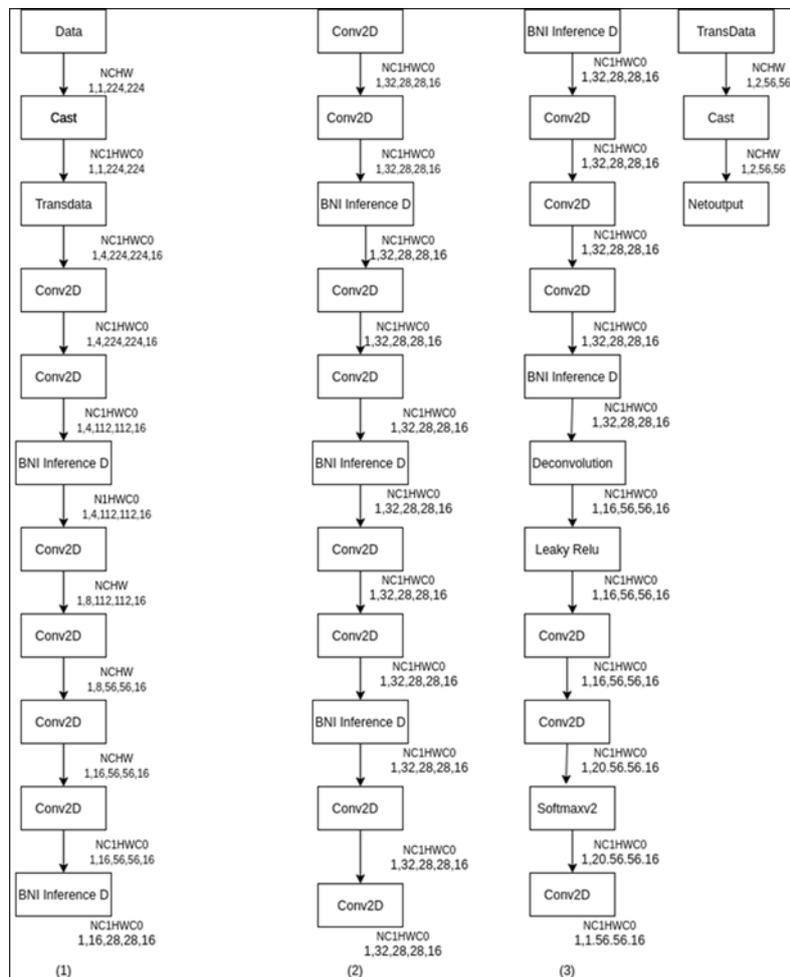
Figure 4. CNN algorithm process works in the system

Data in the form of photos that have been selected for testing will first enter the casting stage, where in that stage, the conversion process will be carried out according to system requirements. After going through the casting stage, the input in the form of a photo will be equalized in size first. The input photo is 32 x 16 in size, which is the length x width of an image (photo). The first convolution process is carried out with input dimensions in Conv2D with input dimensions of 32x16x3. Then enter the second convolution process by continuing the size of the previously produced photo. Each layer has a different size in getting a feature. The convolution process results in a feature map of the input data (photos).

The feature map that has been obtained will be entered into the activation function contained in the Activation Layer, where in this study, the ReLu function is used to change the value of features in a range. The activation function used in this study is the ReLU function which works by changing the value of the features within a range. The next data will undergo a pooling process which is carried out to reduce the dimensions of the feature map without losing the important information in it.

The next stage is the fully connected layer using previously obtained results and converting the data into one-dimensional data. After transforming, the data will be processed so that it can be classified. The softmax stage is carried out with a function to get the classification results. The results that will be output will be in the form of photos that have color and are obtained from the results of color predictions on objects in the image (photo). These results will be resized according to the initial size of the photo. After resizing, the photo will be output in the "out" folder and adjusted to the photo's name.

After the training and testing process has been carried out, the results of the test data will be obtained in Figure 5:

Figure 5. Result of testing image

3.2. Experimentation

In this study, the authors carried out the experimental stage with a total of 6 (six) experiments which were differentiated based on the number of input photos that were used as a dataset in 1 process using the CNN method. The following is a description of each experiment.

1. *First Experiment*: The author carried out the first experiment using 1 (one) input photo. The result obtained is a color photo with a processing time of 20 seconds to produce the photo, which can be seen in Table 1.

Table 1. First experiment

Number	Input Photo	Result	Time
1			20 Second

2. *Second Experiment*: The author conducted the second experiment using 5 (five) input photos. The result obtained is a color photo with a processing time of 36 seconds to produce the photo, which can be seen in Table 2.

Table 2. Second experiment

Number	Input Photo	Result	Time
2			36 Second

3. *Third Experiment*: The author conducted the third experiment using 10 (ten) input photos. The result obtained is a color photo with a processing time of 31 seconds to produce the photo.
4. *Fourth Experiment*: The author conducted the fourth experiment using 15 (fifteen) input photos. The result obtained is a color photo with a processing time of 27 seconds to produce the photo.
5. *Fifth Experiment*: The fifth experiment was conducted by the author using 20 (twenty) input photos. The result obtained is a color photo with a processing time of 30 seconds to produce the photo.
6. *Sixth Experiment*: The author conducted the sixth experiment using 30 (thirty) input photos. The result obtained is a color photo with a processing time of 62 seconds to produce the photo.

3.3. Output Analysis

Based on the results of the research that has been done, after the photos are processed using the CNN method, it can be seen that the model created successfully processed 81 black and white input photos into output photos with color. Each experiment conducted by the author takes a different time. Apart from being caused by the amount of data, the time difference also depends on internet speed. Table 3 shows a list of the results of carrying out 6 (six) trial scenarios:

Table 3. List of Times to do Scenarios

Number	Experimentation	Number of Photos	Processing Time
1	First Experiment	1 Photo	20 Second
2	Second Experiment	5 Photo	36 Second
3	Third Experiment	10 Photo	31 Second
4	Fourth Experiment	15 Photo	27 Second
5	Fifth Experiment	20 Photo	30 Second
6	Sixth Experiment	30 Photo	62 Second

3.4. Comparison With Other Colorization Methods

By using user-guided methods using Scribbler [26, 27] and Real-Time User-Guided Colorization [28], in converting black and white colors, the researcher is required to determine the color to be displayed himself so that the resulting image has a predetermined color. Meanwhile, using CNN can automatically generate colors without determining the color first. Research [29], using interactive deep colorization cannot perform coloring correctly if the model cannot provide image data that can be studied. Research [30] uses Tag2Pix to colorize images by entering the desired color tag in certain parts of the object.

Table 4. Results Comparison

Technique	Accuracy Rate	Notes	Sources
Traditional Machine Learning	82%	With A Lot Of Data	[15]
CNN Algorithm	93.57%	With The Same Number Of Datasets	[15]
CNN vs. Other Image Processing Techniques	Best Performance	According To Study	[31, 32]

4. CONCLUSION

Based on the discussion of the research that has been done, it can be concluded that this study using the Convolutional Neural Network (CNN) method, successfully converted black and white photos into color photos that match the predicted color of the object in the photo. The results obtained show that the CNN algorithm is appropriate to use as a method for performing image processing. In each experiment that was carried out, the results showed that the model succeeded in processing images for different lengths of time and as shown in the results table. The length of time required in 1 (one) process is influenced by the speed of the author's network when carrying out each experiment so that the number of photos entered does not become a factor in the processing time. Different test algorithms and software can be used with more data for further research. They can compare results using different parameters that vary, such as the level of color sharpness and the level of difference in the results of the activation function. This algorithm can be developed for black-and-white images defective / damaged so that they can be repaired so that the quality of the image is damaged it gets better.

5. ACKNOWLEDGEMENTS

6. DECLARATIONS

AUTHOR CONTIBUTION

All authors conceived and designed the study. Andrew Fiade, M. Ikhsan Tanggok, and Rizka Amalia Putri carried out the study concept and design. In addition, Andrew Fiade, Rizka Amalia Putri, Luigi Ajeng Pratiwi, and Siti Ummi Masruroh conducted experiments analyzing the data.

FUNDING STATEMENT

COMPETING INTEREST

The authors declare that there are no conflicts of interest regarding the publication of this paper.

REFERENCES

[1] I. B. L. M. Suta, R. S. Hartati, and Y. Divayana, "Diagnosa Tumor Otak Berdasarkan Citra MRI (Magnetic Resonance Imaging)," *Majalah Ilmiah Teknologi Elektro*, vol. 18, no. 2, 2019.

[2] C. Zhang and Y. Lu, "Study on Artificial Intelligence: The State of the Art and Future Prospects," *Journal of Industrial Information Integration*, vol. 23, no. May, p. 100224, 2021.

[3] A. Dhillon and G. K. Verma, "Convolutional Neural Network: A Review of Models, Methodologies and Applications to Object Detection," *Progress in Artificial Intelligence*, vol. 9, no. 2, pp. 85–112, 2020.

[4] B. K. Triwijoyo, A. Adil, and A. Anggrawan, "Convolutional Neural Network With Batch Normalization for Classification of Emotional Expressions Based on Facial Images," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 21, no. 1, pp. 197–204, 2021.

[5] F. F. Maulana and N. Rochmawati, "Klasifikasi Citra Buah Menggunakan Convolutional Neural Network," *Journal of Informatics and Computer Science*, vol. 01, pp. 104–108, 2019.

[6] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, "1D Convolutional Neural Networks and Applications: A Survey," *Mechanical Systems and Signal Processing*, vol. 151, p. 107398, 2021.

[7] T. Setiawan and D. Avianto, "Implementasi Convolutional Neural Network untuk Pengenalan Warna Kendaraan," *Naskah Publikasi*, vol. FTIE, p. Unversitas Teknologi Yogyakarta, 2020.

[8] J. Naranjo-Torres, M. Mora, R. Hernández-García, R. J. Barrientos, C. Fredes, and A. Valenzuela, "A Review of Convolutional Neural Network Applied to Fruit Image Processing," *Applied Sciences (Switzerland)*, vol. 10, no. 10, 2020.

[9] N. M. Trieu and N. T. Thinh, "Quality Classification of Dragon Fruits Based on External Performance Using a Convolutional Neural Network," *Applied Sciences (Switzerland)*, vol. 11, no. 22, 2021.

[10] M. Tripathi, "Analysis of Convolutional Neural Network Based Image Classification Techniques," *Journal of Innovative Image Processing*, vol. 3, no. 2, pp. 100–117, 2021.

[11] A. R. Lubis, M. Lubis, and Al-Khowarizmi, "Optimization of Distance Formula in K-Nearest Neighbor Method," *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 1, pp. 326–338, 2020.

[12] A. Y. Muniar, P. Pasnur, and K. R. Lestari, "Penerapan Algoritma K-Nearest Neighbor pada Pengklasifikasian Dokumen Berita Online," *Inspiration: Jurnal Teknologi Informasi dan Komunikasi*, vol. 10, no. 2, p. 137, 2020.

[13] B. W. Kurniadi, H. Prasetyo, G. L. Ahmad, B. Aditya Wibisono, and D. Sandya Prasvita, "Analisis Perbandingan Algoritma SVM dan CNN untuk Klasifikasi Buah," *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA) Jakarta-Indonesia*, no. September, pp. 1–11, 2021.

[14] D. A. Pisner and D. M. Schnyer, *Support Vector Machine*. Elsevier Inc., 2019.

[15] S. Morris, "Image Classification Using SVM," *IEEE Xplore*, pp. 1–10, 2018.

[16] T. A. Zuraiyah, S. Maryana, and A. Kohar, "Automatic Door Access Model Based on Face Recognition using Convolutional Neural Network," vol. 22, no. 1, pp. 241–252, 2022.

[17] H. Yu, L. T. Yang, Q. Zhang, D. Armstrong, and M. J. Deen, "Convolutional Neural Networks for Medical Image Analysis: State-Of-The-Art, Comparisons, Improvement and Perspectives," *Neurocomputing*, vol. 444, pp. 92–110, 2021.

[18] Y. Tian, "Artificial Intelligence Image Recognition Method Based on Convolutional Neural Network Algorithm," *IEEE Access*, vol. 8, pp. 125 731–125 744, 2020.

[19] P. Musa, E. P. Wibowo, S. B. Musa, and I. Baihaqi, "Pelican Crossing System for Control a Green Man Light with Predicted Age," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 21, no. 2, pp. 293–306, 2022.

[20] A. Luqman Hakim and Prawito, "Rain Detection in Image Using Convolutional Neural Network," *Journal of Physics: Conference Series*, vol. 1528, no. 1, 2020.

[21] R.-s. Images, L. Hu, M. Qin, F. Zhang, Z. Du, and R. Liu, "RSCNN : A CNN-Based Method to Enhance Low-Light," *Remote Sensing, MDPI 2020*, pp. 1–13, 2020.

[22] N. Agrawal and R. Katna, *Applications of Computing, Automation and Wireless Systems in Electrical Engineering.* Springer Singapore, 2019, vol. 553.

[23] A. Ajit, K. Acharya, and A. Samanta, "A Review of Convolutional Neural Networks," *International Conference on Emerging Trends in Information Technology and Engineering, ic-ETITE 2020*, pp. 1–5, 2020.

[24] D. R. Sarvamangala and R. V. Kulkarni, "Convolutional Neural Networks in Medical Image Understanding: A Survey," *Evolutionary Intelligence*, vol. 15, no. 1, 2022.

[25] S. B. Jadhav, V. R. Udupi, and S. B. Patil, "Identification of Plant Diseases Using Convolutional Neural Networks," *International Journal of Information Technology (Singapore)*, vol. 13, no. 6, pp. 2461–2470, 2021.

[26] I. Zeger, S. Grgic, J. Vukovic, and G. Sisul, "Grayscale Image Colorization Methods: Overview and Evaluation," *IEEE Access*, vol. 9, pp. 113 326–113 346, 2021.

[27] L. Zhang, C. Li, E. Simo-Serra, Y. Ji, T. T. Wong, and C. Liu, "User-Guided Line Art Flat Filling with Split Filling Mechanism," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 9884–9893, 2021.

[28] S. Anwar, M. Tahir, C. Li, A. Mian, F. S. Khan, and A. W. Muzaffar, "Image Colorization: A Survey and Dataset," pp. 1–20, 2020.

[29] Y. Xiao, P. Zhou, Y. Zheng, and C. S. Leung, "Interactive Deep Colorization Using Simultaneous Global and Local Inputs," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2019-May, pp. 1887–1891, 2019.

[30] H. Kim, H. Y. Jhoo, E. Park, and S. Yoo, "Tag2pix: Line Art Colorization Using Text Tag With SECAT and Changing Loss," *Proceedings of the IEEE International Conference on Computer Vision*, vol. 2019-Octob, pp. 9055–9064, 2019.

[31] S. Rajpurohit, S. Patil, N. Choudhary, S. Gavasane, and P. Kosamkar, "Identification of Acute Lymphoblastic Leukemia in Microscopic Blood Image Using Image Processing and Machine Learning Algorithms," *2018 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2018*, no. Cll, pp. 2359–2363, 2018.

[32] B. K. Hatuwal, A. Shakya, and B. Joshi, "Plant Leaf Disease Recognition Using Random Forest, KNN, SVM and CNN," *Polibits*, vol. 62, no. May, pp. 13–19, 2020.

[This page intentionally left blank.]