ISSN : 1978-8282, Online ISSN: 2655-4275

Classify Tomato Fruit Images Using Convolutional Neural Network (CNN) Method

Fauziah[1], Bayu Kumoro Yakti[*2], Ragiel Hadi Prayitno[3], Tjahjo Dwi Nurti[4], Nur Azizah[5]
[1,2,3] Information Technology Doctoral Program, Gunadarma University, Indonesia
[4] Industrial Engineering, Gunadarma University, Indonesia
[5] Faculty of Science and Technology, University of Raharja, Indonesia
E-mail: [1]fauziah87@staff.gunadarma.ac.id, [*2]**bayuyakti@staff.gunadarma.ac.id**, [3]ragielhp@staff.gunadarma.ac.id, [4]dwinurti@staff.gunadarma.ac.id, [5]nur.azizah@raharja.info

Abstract

The agricultural and forestry product processing industry is developing very rapidly. The stage in the process of processing plant products is selecting products according to their quality, for example fruit ripeness. The process of selecting agricultural and horticultural products is often largely based on human perception of the color composition of fruit images (manual selection). The weakness of manual fruit classification is greatly influenced by the subjectivity of the classifier. To reduce subjectivity and manual methods, a Convolutional Neural Network (CNN) deep learning method is needed. Tomatoes are chosen for image classification based on color diversity, aiding accurate recognition through distinct ripeness variations. The aim of this research is to detect and recognize tomato images and determine its accuracy value by applying the CNN Deep Learning method. The first stage is preparing the required tomato image data set. The second stage is preprocessing and sorting the tomato image. The third stage is model formation and system training. The last is to carry out system testing. This research uses 14 tomato images which are used as testing data from 56 tomato images used in the training dataset. Testing tomatoes produces an average data testing accuracy value of 97%.

Keywords *— CNN, Deep Learning, Tomato*

1. INTRODUCTION

The agricultural and forestry product processing industry is developing very rapidly[1]. The stage in the process of processing plant products is selecting products according to their quality, for example fruit ripeness. Progress in selecting agricultural and horticultural products is often largely based on human perception of the color composition of images (fruit). The manual method is based on direct observation of the fruit being assessed. The weakness of manual fruit classification is greatly influenced by the subjectivity of the classifier, so that under certain conditions the classification process is not specific. Determination using this method has several weaknesses, especially the relatively long time required and the production of various products. Limited human vision, level of fatigue and differences in perception of fruit quality[2]. Evolved science and digital image processing technology enable agricultural and horticultural products to be sorted automatically using image processing algorithms. This technology can reduce subjectivity issues in human classifiers, increase efficiency, and produce more consistent products[3].

Based on existing facts, this research uses tomatoes as an object to classify ripe tomatoes (red tomatoes) and unripe tomatoes (green tomatoes)[4]. The ripeness of tomatoes can be detected via digital images by implementing the Deep learning method[5]. Deep learning is a field of machine learning that implements learning methods. The use of deep learning can be based on several factors such as the use of quite a lot of training data to be able to learn the characteristics of the training data, making the development of parallel computers for fairly high performance and structural design networks and training strategies that have developed significantly[6] [7]. The advantage of using deep learning is that users can collect elements that make it possible to make judgments[8].

Convolutional Neural Network (CNN) is a type of artificial neural network algorithm that is commonly used to classify and recognize objects in data in the form of images[9]. CNN network is created with the assumption that the input used is an image. This network has a special layer called the convolution layer, where in this layer an input image will be processed based on predetermined filters. Each layer will produce a pattern from several parts of the image which will be easier to classify, thus making the learning function more efficient to implement[10] [11]. CNN convolution is a mathematical operation on two functions which then produces a third function. CNNs use convolution as a replacement for general matrix multiplication. This operation is used at least once in each layer[12] [13].

Previous research that has been carried out discusses the accuracy of tomato ripeness using several algorithms. Shinta A, Sukemi conducted research on the classification of tomato ripeness using the K-Nearest Neighbor (KNN) algorithm based on the skin color of the fruit. The k values used in this research are 1, 3, 5, 7, and 9 to test searching for Euclidean distance on images with a size of 512x512 pixels. Research carried out proves that the Euclidean distance k=3 has a percentage value of 92%. Based on the level of accuracy, the color feature k=3 shows the best k value in classifying tomato ripeness levels [14]. Seyed M N, Amir T, Abdolabbas J conducted research to classify ripe tomatoes based on color, size and hardness of fruit using a fuzzy algorithm. The research results show that the fuzzy algorithm is able to group fruit correctly, with an accuracy level ranging from 87,4% to 93,3% depending on the fuzzy algorithm configuration used. The best results were obtained by using the zmf and sigmf fuzzifier methods, as well as the gbellmf fuzzifier as a fuzzifier and mom as a defuzzifier [15].

Conclusions from previous studies show that classification of tomato fruit ripeness can be done with a good level of accuracy using a color-based algorithm. Shinta A's research using the KNN algorithm shows that the tomato skin color feature with Euclidean distance k=3 gives the best results with 92% accuracy. Meanwhile, research by Seyed M N, Amir T, Abdolabbas J using a fuzzy algorithm to classify ripe tomatoes based on color, size and hardness of the fruit succeeded in achieving an accuracy level of between 87,4% to 93,3%, depending on the configuration of the fuzzy algorithm used. Thus, the color-based classification method proved effective in identifying the level of ripeness of tomato fruit.

2. RESEARCH METHOD

2.1. Data Collection

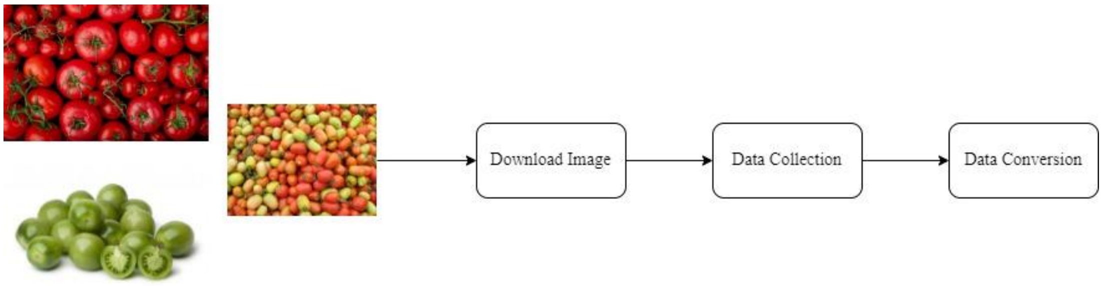

Figure 1. Data Collection Stages

Figure 1 explains that when classifying tomatoes, it is necessary to prepare the required tomato image dataset. Tomato photo data was taken randomly from internet, there were 70 images consisting of 36 red tomatoes and 34 green tomatoes. Images are taken in JPG format and Each tomato color has its own folder. The data obtained is converted into tomato image data with different sizes. These tomato images were used during training and testing.

2.2. Preprocessing and Selection of Tomato Fruit Images

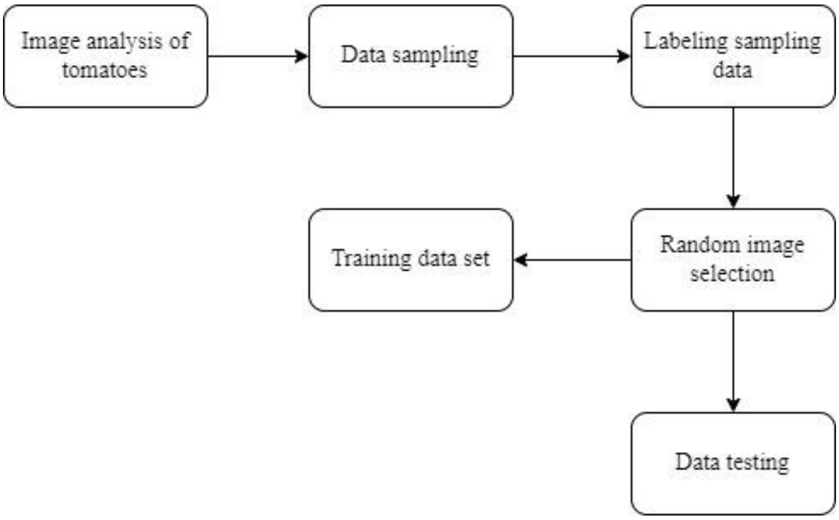

Figure 2. Preprocessing Stages and Selection of Tomato Fruit Images

Preprocessing and sorting of tomato fruit images is shown in Figure 2. The first stage is to analyze one by one the images obtained from internet to ensure that the images used are images of red and green tomatoes. Images of tomatoes are put into each folder with separate folders for red and green tomatoes. Normalize the resulting data to tomato image data by adjusting it to the same size, namely 8x8 pixels. Refine images and determine the ratio of the number of images used as training data and test data. The training dataset used contained 56 images out of a total of 70 images, and the test data used contained 14 images. The image data used as a training data set has a number attached to each image. Examples of labeling are shown in Figure 3 and Figure 4.

Figure 3. Red Tomato Image Labeling

Figure 4. Green Tomato Image Labeling

Data that has been foldered for each image can be input into the library. Before creating a model, it is necessary to change to a one hot encoding scheme. Due to changing category features to a format that functions better in classification. Then combine the images of 2 types of tomatoes randomly as in Figure 5. Next, the data is processed into a CNN model.

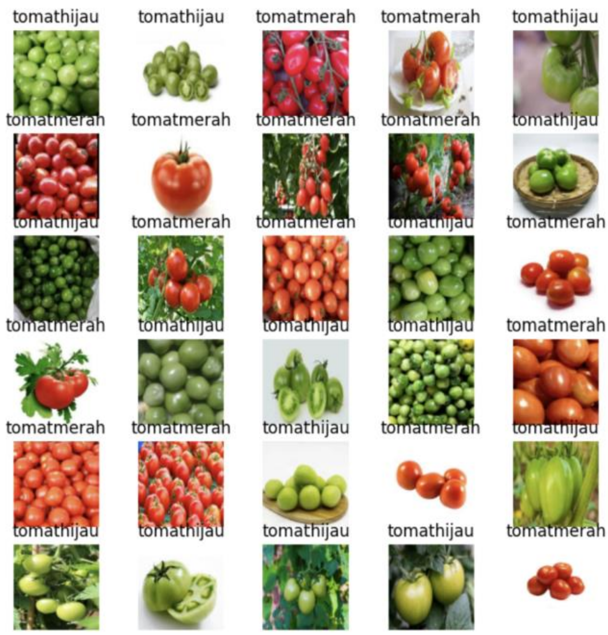

Figure 5. Training Images Randomly

2.3. System Planning

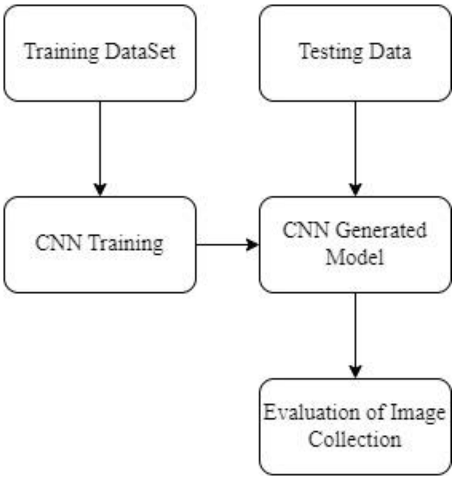

Figure 6. System Planning

2.3.1. Training Dataset

The training dataset was taken by downloading tomato images via a browser in jpg image format. The total images used were 70 images of tomatoes from 2 classes which were used as training data. The image size will be changed based on the CNN algorithm format to produce a better image. The dataset is labeled according to the file section.

2.3.2. CNN Training

The image that has been given a label will be processed with the CNN model used, which is trained with 7 hidden layers (convolutional layers). The workings of the 7 layers are as follows: 6 network layers, with the same architecture for each network layer followed by non-linear ReLU elements and a 2 x 2 MaxPooling. The criteria for selecting numbers in the network layer are related to the convergence on the error rate during the process learning. In this case study, it took 5 or 6 iterations (to increase the network layer number) to find the calculation [16]. The CNN training architecture used in this research can be seen in Figure 7. Stochastic gradient descent was used for CNN training using a small and the same random data set for each iterative learning phase.

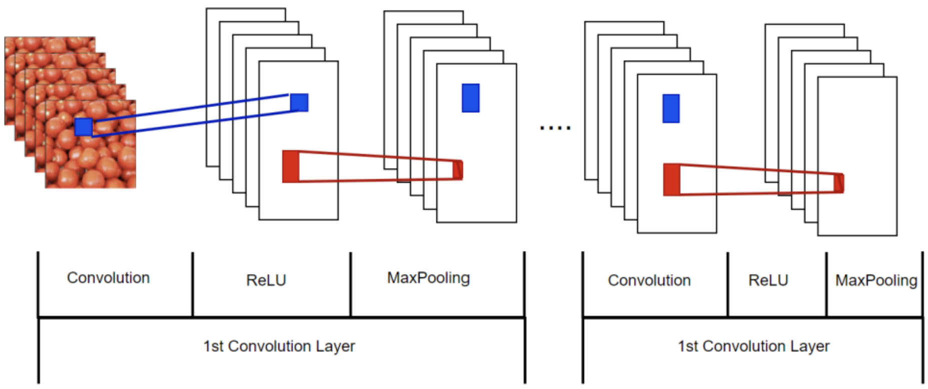

Figure 7. CNN Architecture

2.3.3. CNN Generated Model

The model generated by the CNN is saved to contain the data that has been considered. The CNN model can then be used at any time with experimental data to evaluate the accuracy of the model. The model that will be used is a simple Multi Layer Perceptron (MLP) model, therefore the Sequential model used. The CNN model has several layers including 2D convolution, pooling dropout. It can be seen in Figure 8.

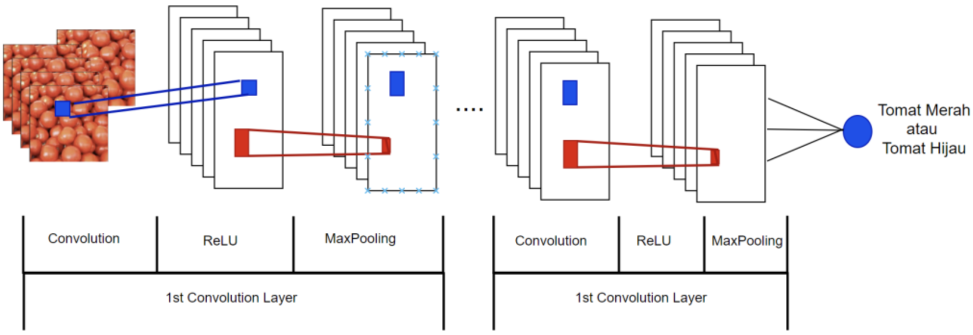

Figure 8. CNN Architecture for Tomato Fruit Classification

```
Model: "sequential_22"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 rescaling_36 (Rescaling)    (None, 180, 180, 3)       0

 conv2d_66 (Conv2D)          (None, 180, 180, 16)      448

 max_pooling2d_68 (MaxPooli  (None, 90, 90, 16)        0
 ng2D)

 conv2d_67 (Conv2D)          (None, 90, 90, 32)        4640

 max_pooling2d_69 (MaxPooli  (None, 45, 45, 32)        0
 ng2D)

 conv2d_68 (Conv2D)          (None, 45, 45, 64)        18496

 max_pooling2d_70 (MaxPooli  (None, 22, 22, 64)        0
 ng2D)

 flatten_22 (Flatten)        (None, 30976)             0

 dense_44 (Dense)            (None, 128)               3965056

 dense_45 (Dense)            (None, 2)                 258

=================================================================
Total params: 3988898 (15.22 MB)
Trainable params: 3988898 (15.22 MB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Figure 9. Model Obtained

Flatten_22 produces an output of 30976 which is obtained from the multiplication calculation of the previous dimensions, namely 22x22x64 = 30,976 and dense_44 outputs 128 which is a number that shows the number of neurons used. The parameter of 3,965,056 is obtained from the multiplication result of 30,976x128 = 3,965,056. Dense_45 shows the number of image categories used, so the parameters are 2x128+2= 258. The total parameters of the model are 3,988,898. then train the data on the pepper image into a model with a fit model. To fit the model, use epoch = 10, bacth_size = 32 and validation_split = 0.2. Epoch is the number of times the network will see the entire data set, batch_size is the number of training examples in one forward/backward pass. And it can be concluded that the higher the batch_size value, the more memory is needed. So the complete CNN Architecture Model for pepper classification adds one full connection network layer to the training architecture which can be seen in Figure 8. As the final classifier, the Softmax function is used.

2.3.4. Testing Data

In testing using 14 different tomato image data. This research uses the same training setting parameters, namely the maximum epoch value is 10. The convolution2dLeyer layer parameter settings use image size filter length and width 2 and filter number 10 and maxPooling2dLayer with pool size 2 and Stride 2. The experimental data is a set of tomato images. This data is input for the model obtained from CNN to predict the level of accuracy in images which are likely to be similar images of red tomatoes and green tomatoes.

2.3.5. Evaluation of Image Classified

This stage is the stage that evaluates the images that have been classified according to the results of the level of accuracy and suitability of the tomato image. To compare the value of accuracy, it can be done by concentrating on increasing the number of layers in the CNN architecture. And this addition will not change the values of previously determined parameters such as epoch values and learning rate. The convolutional layer used is layer seven, the accuracy value is obtained after the seventh layer which is the average of each layer's accuracy value. In this research, experimental data is input for the resulting model and can identify images of red tomatoes and green tomatoes which are the output results. The application used in this research uses the TensorFlow framework. TensorBoard is a histogram from TensorFlow, a graphic visualization tool that makes it easier to understand model parameters and their variations over time.

3. RESEARCH RESULTS AND DISCUSSION

3.1. Data Collection and Training Results

The graphs in Figure 10 and Figure 11 are the training results obtained from the CNN model. Figure 10 where the x-axis is the Epoch while the y-axis is the accuracy value. Y-axis is where the accuracy values for validation data tend to be stable, so this model is good and optimal. Figure 10 shows the difference in accuracy between training and validation results. Meanwhile, the train accuracy value is 0,99 and starting from epoch 0 tends to be stable and continues to increase. Figure 11 is a loss graph from the learning process where the x-axis is Epoch while the y-axis is Loss with accuracy values. The error value for the training data starts at around 2,5 at epoch 4 while the validation data tends to decrease significantly.

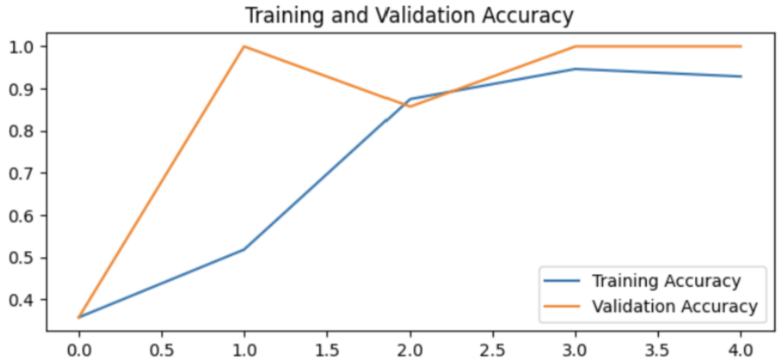

Figure 10. Training and Validation Accuracy Charts

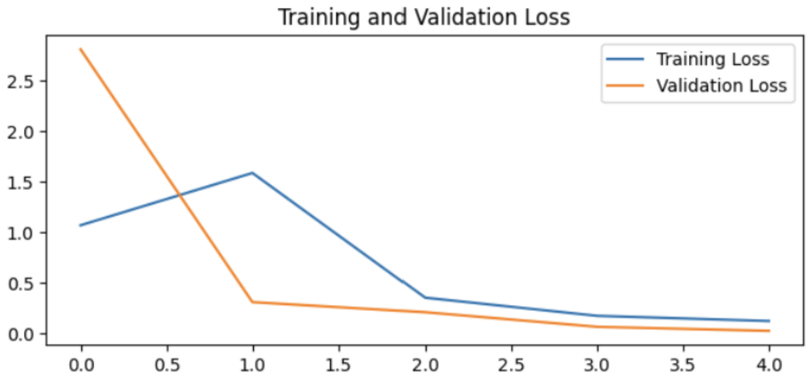

Figure 11. Training and Validation Loss Graph

3.2. Testing Process Results

Results of the process is carried out, namely the training and testing process using the CNN method using validation of the training and testing process. The test results can be seen in Table 1 for red tomatoes and Table 2 for green tomatoes.

Table 1. Red Tomato Test Results

File Name	**Image**	**Size**	**Accuracy**
tomatmerah (1).jpg		300x168	99,90 %
tomatmerah (7).jpg		246x205	91,30%
tomatmerah (14).jpg		300x168	98,22%
tomatmerah (27).jpg		226x223	89,66%
tomatmerah (32).jpg		290x174	77,51%

Table 2. Green Tomato Test Results

File Name	Image	Size	Accuracy
tomathijau (3).JPG		300x168	91,45%
tomathijau (14).JPG		225x225	90,80%
tomathijau (21).JPG		271x186	74,93%
tomathijau (29).JPG		225x225	93,00%
tomathijau (34).JPG		452x318	91,78%

From the results of observing the 2 data that have been taken, it can be concluded that the accuracy that is 91,318% (average accuracy from Table 1) is for red tomatoes, the size of the red tomato in index 1 (tomatmerah (1).jpg) has a size of 300x168 pixels which is the same as the green tomato. The accuracy fot green tomato is 88,392% (average accuracy from Table 2). The accuracy values below 80% are red tomatoes at index 32 (tomatmerah (32).jpg) and green tomatoes at index 21 (tomathijau (21).jpg). This proves that the system created is able to detect the accuracy level of red tomatoes and green tomatoes that come from the browser and are taken randomly. The accuracy values of the training and testing process using the CNN method from previously determined images can be seen in Table 3. The data in Table 3 explains that the accuracy value obtained from training and testing was 99% accuracy value from training 56 tomato images and the testing data had an accuracy value of 97% with 14 tomato images.

Table 3. Training and Testing Classification Accuracy Results

Data	Amount of data	Accuracy (%)
Training	56	99
Testing	14	97

4. CONCLUSION

Classification of red tomatoes and green tomatoes using the CNN method was successfully carried out. The training dataset used was 56 image data of red tomatoes and green tomatoes. The test used 70 images of red tomatoes and green tomatoes where all images were taken from the browser randomly with a total of 10 epochs. The classification accuracy value of red tomatoes and green tomatoes was influenced by the image size and image format, namely JPG. By using 7 convolutional layers, the resulting accuracy values range between 80% -100%. The average accuracy value resulting from testing data is 0,999. The use of convolutional layers is able to detect the strength of the shape of an image. Using a classification module with CNN, the results obtained in this research are better. This proves that the system is successful in detecting red tomatoes and green tomatoes.

5. SUGGESTED

This research can be developed by implementing real-time detection, counting the number of tomatoes and determining which tomatoes have good quality so that it can be implemented in the tomato sorting industry that is ready to be marketed.

6. REFERENCES

[1] M. N. Isnaini, Agustono and U. Barokah, "Identification of Agriculture, Forestry, and Fisheries Sectors in Mitigating Income Inequality in the Former Residencies of Madiun," Agriecobis (Journal of Agricultural Socioeconomics and Business), pp. 83-96, 2022.

[2] A. distan, "Pengertian dan Konsep Pertanian Berkelanjutan," Pemerintah Kabupaten Buleleng Dinas Pertanian, 28 October 2019. [Online]. Available: https://distan.bulelengkab.go.id/informasi/detail/artikel/pengertian-dan-konsep-pertanian-berkelanjutan-22. [Accessed 25 January 2024].

[3] R. S. L, A. N, D. V and K. T, "Fruit Classification System Using Computer Vision: A Review," International Journal of Trend in Research and Development, vol. 5, no. 1, pp. 22-26, 2018.

[4] M. S. Nasution and N. Fadillah, "Deteksi Kematangan Buah Tomat Berdasarkan Warna Buah dengan Menggunakan Metode YCbCr," InfoTekJar : Jurnal Nasional Informatika dan Teknologi Jaringan, vol. 3, no. 2, pp. 251-254, 2019.

[5] bdullah, A. Harjoko and O. Mahmod, "Classification of Fruits Based on Shape and Colorusing Combined Nearest Mean Classifiers," JURNAL RESTI(Rekayasa Sistem dan Teknologi Informasi), vol. 7, no. 1, pp. 51-57, 2023.

[6] X. Yang, Y. Ye, X. Li, R. Y. K. Lau, X. Zhang and X. Huang, "Hyperspectral Image Classification With Deep Learning Models," IEEE Transactions on Geoscience and Remote Sensing, vol. 56, no. 9, pp. 5408-5423, 2018.

[7] F. Milletari, "Hough-CNN: Deep learning for segmentation of deep brain regions in MRI and ultrasound," Computer Vision and Image Understanding, vol. 164, no. 1, pp. 92-102, 2017.

[8] M. T. Tombeng and Z. Ardian, "Prediksi Penjualan Supermarket Menggunakan Pendekatan Deep Learning," CogITo Smart Journal, vol. 7, no. 1, p. 160–169, 2021.

[9] Z.-Q. Zhao, P. Zheng, S.-T. Xu and X. Wu, "Object Detection With Deep Learning: A Review," IEEE Transactions on Neural Networks and Learning Systems, vol. PP, pp. 1-21, 2019.

[10] F. Chen and M. R. Jahanshahi, "NB-CNN: Deep Learning-Based Crack Detection Using Convolutional Neural Network and Naïve Bayes Data Fusion," IEEE Transactions on Industrial Electronics, vol. 65, no. 5, pp. 4392-4400, 2018.

[11] M. Duan, "A Hybrid Deep Learning CNN–ELM for Age and Gender Classification," Neurocomputing, vol. 275, no. 1, pp. 448-461, 2018.

[12] D. Wang, M. Zhang, Z. Li, J. Li, M. Fu, Y. Cui and X. Chen, "Modulation Format Recognition and OSNR Estimation Using CNN-Based Deep Learning," IEEE Photonics Technology Letters, vol. 29, no. 19, pp. 1667-1670, 2017.

[13] G. A. Sandag and J. Waworundeng, "Identifikasi Foto Fashion Dengan Menggunakan Convolutional Neural Network (CNN)," CogITo Smart Journal, vol. 7, no. 2, p. 305–314, 2021.

[14] S. Aprilisa and Sukemi, "Klasifikasi Tingkat Kematangan Buah Tomat Berdasarkan Fitur Warna Menggunakan K-Nearest Neighbor," Computer Science and ICT, vol. 5, no. 1, pp. 170-173, 2019.

[15] S. M. Nassiri, A. Tahavoor and A. Jafari, "Fuzzy logic classification of mature tomatoes based on physical properties fusion," Information Processing in Agriculture, vol. 9, no. 4, pp. 547-555, 2022.

[16] L. Craig and R. Awati, "convolutional neural network (CNN)," techtarget, January 2024. [Online]. Available: https://www.techtarget.com/searchenterpriseai/definition/convolutional-neural-network#:~:text=CNNs%20use%20a%20series%20of,layers%20to%20recognize%20detailed%20patterns.. [Accessed 29 January 2024].