

# Block-based compressive sensing in deep learning using AlexNet for vegetable classification

Indrarini Dyah Irawati<sup>1</sup>, Gelar Budiman<sup>2</sup>, Sofia Saidah<sup>2</sup>, Suci Rahmadiani<sup>2</sup> and Rohaya Latip<sup>3</sup>

<sup>1</sup> School of Applied Science, Telkom University, Bandung, West Java, Indonesia

<sup>2</sup> School of Electrical Engineering, Telkom University, Bandung, West Java, Indonesia

<sup>3</sup> Department of Communication Technology and Network/Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Serdang, Selangor, Malaysia

## ABSTRACT

Vegetables can be distinguished according to differences in color, shape, and texture. The deep learning convolutional neural network (CNN) method is a technique that can be used to classify types of vegetables for various applications in agriculture. This study proposes a vegetable classification technique that uses the CNN AlexNet model and applies compressive sensing (CS) to reduce computing time and save storage space. In CS, discrete cosine transform (DCT) is applied for the sparsing process, Gaussian distribution for sampling, and orthogonal matching pursuit (OMP) for reconstruction. Simulation results on 600 images for four types of vegetables showed a maximum test accuracy of 98% for the AlexNet method, while the combined block-based CS using the AlexNet method produced a maximum accuracy of 96.66% with a compression ratio of 2×. Our results indicated that AlexNet CNN architecture and block-based CS in AlexNet can classify vegetable images better than previous methods.

**Subjects** Artificial Intelligence, Computer Vision, Neural Networks

**Keywords** AlexNet, Classification, Compressive sensing, Convolution neural network, Deep learning, Vegetable

Submitted 22 February 2023

Accepted 2 August 2023

Published 16 November 2023

Corresponding author

Indrarini Dyah Irawati,  
indrarini@telkomuniversity.ac.id

Academic editor

Giovanni Angiulli

Additional Information and  
Declarations can be found on  
page 12

DOI 10.7717/peerj-cs.1551

© Copyright

2023 Dyah Irawati et al.

Distributed under

Creative Commons CC-BY 4.0

**OPEN ACCESS**

## INTRODUCTION

Indonesia is one of the world's top producers of horticultural crop commodities, especially vegetables. Vegetables produced by Indonesian farmers are mostly used to meet daily household needs, but some are industrially processed for various processed products that are even exported abroad. Vegetables are foodstuffs that contain several vitamins and minerals and provide many benefits for the health of the human body. In addition, growing vegetables gives farmers a high chance of increasing their profits in Indonesia. Therefore, there is a need for a digital approach to classify vegetables more easily.

Technology is very influential in everyday life because individuals can carry out activities more effectively and efficiently (Zeng, 2017). Previous research (Chauhan, Ghanshala & Joshi, 2018) used CNN to detect and recognize objects in images by imitating the image recognition system in the human visual cortex. Another study (Hameed, Chai & Rassau, 2020) utilized CNN for object categorization with promising results. CNN

architecture was also studied for sound/music classification in research by [Doukhan & Carrive \(2017\)](#). GoogLeNet architecture has also been used to monitor freshness in bananas, and the results showed that the model could detect banana freshness with an accuracy of 98.92%, which was higher than the human detection rate. The fruits tested in this study were limited to bananas and strawberries ([Ni et al., 2020](#)). Sorting tomatoes conventionally can be costly, unproductive, and the reliability is uncertain. Therefore, this study aimed to design a powerful AI-based controller for the process of sorting tomatoes into three categories (namely raw, ripe, and defective (ripe and rotten)), as well as provide designs for a cost-effective tomato sorting machine. An information processing approach with reliable estimation capabilities that works based on the workings of the biological nervous system in human brain cells is known as an artificial neural network (ANN). In agriculture, ANN can be used to detect oil palm plants infected with *Ganoderma* through remote sensing based on the Unmanned Image Aerial Vehicle (UAV) ([Ahmadi et al., 2022](#)). Research by [Haggag et al. \(2019\)](#) developed a control algorithm based on CNN-ANN with 100% classification performance for all classes. Classification results reach 100% for raw and ripe data and 90% for mature and defective (ripe and rotten). [Pratondo, Elfahmi & Novianty \(2022\)](#) developed a model to classify the Zingiberaceae species from images that were directly photographed with a cellphone camera. The pre-trained visual geometri group (VGG-19) and Inception V3 with ImageNet weights using transfer learning simulations resulted in accuracies of 92.43% and 94.29%, respectively. Identifying weeds in plantation vegetable crops is quite challenging because there are many species of weeds; therefore, deep learning technology and image processing technology are used to identify them. CenterNet architecture is trained to detect vegetables with precision, recall, and F1-score values of 95.6%, 95%, and 95.3% respectively ([Jin, Che & Chen, 2021](#)).

The classification of vegetables and circular fruit, namely apples, lemons, oranges, pomegranates, tomatoes, and colored peppers, is carried out using the GoogLeNet architecture. The results of this study obtained a GoogLeNet training accuracy of 96.88% and test accuracy of 96% ([Yuesheng et al., 2021](#)). Classification of images of fruits and vegetables in supermarkets is a complex problem. CNN have shown promising results for object classification. Transfer learning and ensemble techniques can be used to overcome the scarcity of datasets. Accuracy results were obtained using the GoogLeNet architecture in 95% training, 93% testing, 95% ensembles, and 96% training accuracy results, 94% testing, and 98% ensembles obtained by MobileNet ([Hameed, Chai & Rassau, 2020](#)). Detection of organic and non-organic fruits and vegetables using machine learning methods was developed using a combination of spectroscopy. The best accuracy was 95.1%, which was obtained from the artificial neural network (ANN) architectural model ([Natarajan & Ponnusamy, 2021](#)).

Research by [Li, Fu & Yu \(2017\)](#) conducted a comparison of the Deep-CNN (DCNN), ANN, and TMPL methods with high-resolution remote sensing imagery to identify oil palm trees. The results of this study showed that CNN obtained the highest accuracy value, 92–97% in four regions, then ANN with 83–85%, and then TMPL with 70–77%. An intra-class vegetable recognition system used the CNN method with 24 categories of vegetables. The accuracy of the vegetable recognition rate obtained in this study was 95.50%

(Sudharshan Duth & Jayasimha, 2020). Another study developed by Litvak, Divekar & Rabaev (2022) provided the Urban Planter dataset for classifying plant species consisting of 1,500 images with 15 types of houseplants. This study also tested DCNN different configurations and pre-trained models and web applications to evaluate the architectural model. Deep learning-based real-time vegetable recognition system has been used for agricultural robots. Comparative experiments showed that the recognition system could effectively identify seven different types of vegetables (the average AP can reach 87.89%, and the detection speed is up to 38FPS) (Zheng et al., 2019). DCNN-based vegetable image classification was tested by Ahmed, Mahmud Mamun & Zaman Asif (2021). Comparison tests of several CNN architectures such as VGG16, MobileNet, InceptionV3, and ResNet were carried out using transfer learning scenarios to get the optimum architecture built upon the accuracy and effectiveness of the new image data set. Fruit and vegetable classification using an android application with a deep learning method is easy to access. In Agrawal et al. (2021), the research was split into two parts: the first part explored the model through its training and development, and the second part was developed to obtain a simple model through trimming techniques so that the process was lighter but still maintained accuracy. Previous research also used AlexNet architecture in classifying the types of vegetables, which consisted of five classes of vegetables. The result of the accuracy test in this study was 92.1% (Zhu et al., 2018). The same study was also conducted using three classes of vegetables and the You Only Look Once algorithm and succeeded in predicting with an accuracy of 61.6% (Sachin et al., 2019).

The need for storage space and image data transmission is often an obstacle in the classification process because storage requires sufficient space. In contrast, the image data transmission process requires considerable time and bandwidth. The application of CS on image data can minimize the delivery time of image data and requires less storage space. The research combining CS procedure and deep learning is rare. The novelty of our method is the CS contribution to the system with the deep learning procedure. In this study, the development of research was carried out by applying CS to the CNN, thus the main contributions of this article consist of (1) efficient storage usage, (2) delivery time reduction, (3) The accuracy of testing is tremendous although the data quality is reduced and (4) data bandwidth is reduced before training and testing process obtaining an efficient data transmission.

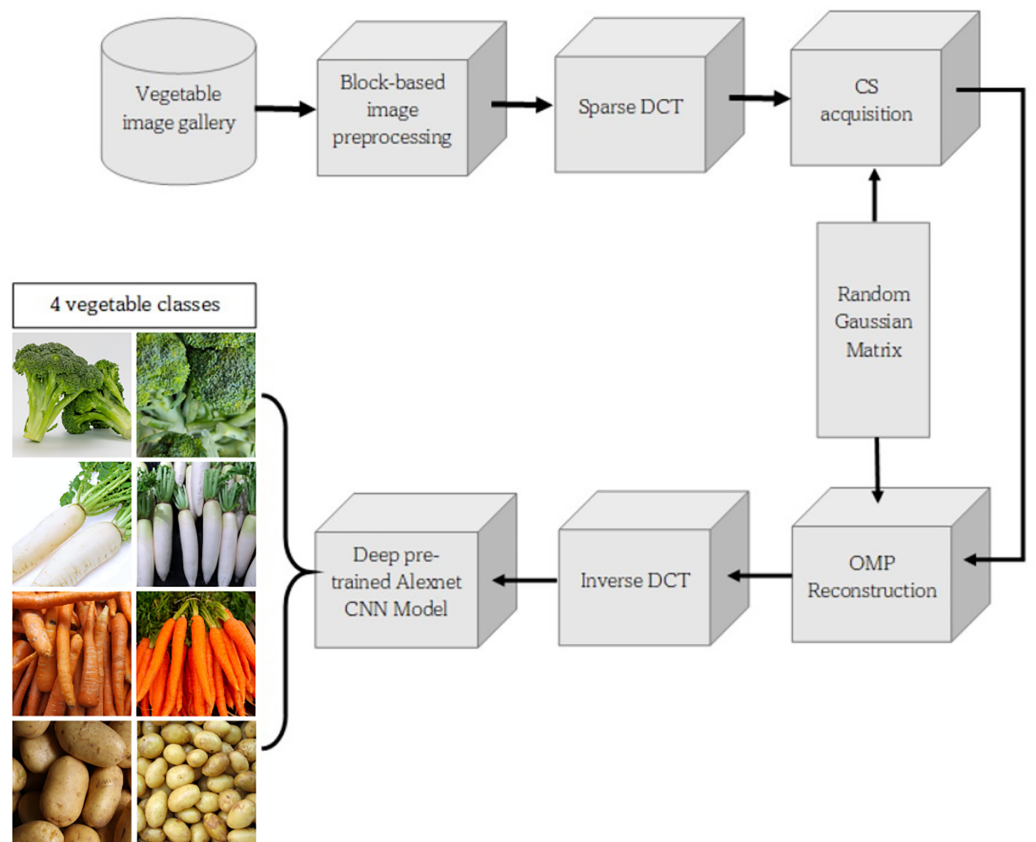
## MATERIALS AND METHODS

The dataset used in this study is a secondary dataset obtained from Kaggle in JPG format as shown in Fig. 1 (Ahmed, Mahmud Mamun & Zaman Asif, 2021). The data consisted of four classes (image data of broccoli, carrot, potato, and radish), and each class was made up of RGB images. The dataset varied and was taken from different perspectives so that the CNN algorithm could recognize objects and learn more about the object under study.

Figure 2 shows the block diagram of the proposed method. The input was an RGB vegetable image during the preprocess stage. The original image was resized to  $256 \times 256$  pixels. The next step in the process was block splitting, which made the image into blocks with a size of  $8 \times 8$ .



**Figure 1** Vegetable dataset examples. (A) Broccoli. (B) Carrots. (C) Potatoes. (D) Radishes (Kaggle, Ahmed, Mahmud Mamun & Zaman Asif, 2021). Full-size DOI: 10.7717/peerj-cs.1551/fig-1



**Figure 2** Block diagram of proposed method (Kaggle, Ahmed, Mahmud Mamun & Zaman Asif, 2021). Full-size DOI: 10.7717/peerj-cs.1551/fig-2



In the CS stage, the sparse features of signals in a transform domain were used to perform non-uniform sampling of signals at a sample rate significantly lower than Nyquist's sampling frequency, allowing the sampling and compression processes to be merged (Ouyang, Hu & Hou, 2019). The sparsing process was carried out on the  $8 \times 8$  block. The next step was to convert the two-dimensional or  $8 \times 8$  signal into a  $64 \times 1$  column vector.

CS consists of two stages: compression and reconstruction. For compression, the signal is acquired with a Gaussian measurement matrix ( $A$ ) measuring  $M \times 64$  so that a matrix of  $M \times 1$  size is obtained. While the reconstruction process aims to restore the original signal, OMP is one of the greedy and iterative reconstruction algorithms that performs calculations repeatedly to get optimal results (Usman, 2017). The next stage is transforming the signal from the frequency domain to the spatial domain using IDCT to get the original signal (Irawati et al., 2022).

CNN is a neural network commonly used in image data, and consists of neurons with weight, bias, and activation functions (Doukhan & Carrive, 2017). It is an extension of the multilayer perceptron (MLP). The CNN architecture consists of two main parts, feature learning and classification. In the first stage of feature learning, the convolution stage uses a kernel of a certain size. The number of kernels used depends on the number of features generated from the calculation. After the activation function using a rectifier linear unit (ReLU) and the activation function process is exited, the pooling process is carried out. This process is repeated several times until it is obtained. This feature map is sufficient to continue into the classification section.

In the LSVRC-2010 ImageNet contest, AlexNet was proposed to conduct classification on 1.2 million high-resolution photos. Non-saturating neurons and a GPU implementation of convolution operations made training go faster (Sustika et al., 2018). The suggested neural network system came in first place among participants at the ILSVRC2012. A test error rate of 13.3% was achieved by the networks. AlexNet has been utilized for a variety of tasks since its initial publication, including object identification (Drayer & Brox, 2014), picture segmentation (Long, Shelhamer & Darrell, 2015), and video categorization (Ng et al., 2015).

The block-based CS algorithm in this study is described in two steps: the training and testing process. The training algorithm is described in detail according to the following steps:

1. Read an RGB image and resize to a certain resolution ( $N \times N$ ) obtaining  $x(n, m)$ .
2. Generate random normal distributed  $A$  ( $M \times L$ ) matrix for CS acquisition where  $L = B^2$ ,  $B$  is the size of segmented image and  $M \ll L$ .
3. Apply segmentation of the image into  $B \times B$  pixels obtaining  $x_i$  ( $B \times B$ ) where  $i$  is a block number of the block-based image.
4. Transform  $x_i$  into DCT domain by 2-D DCT obtaining  $X_i$  ( $B \times B$ ).
5. Reshape 2-D matrix  $X_i$  into 1-D vector obtaining  $v_i$  with size  $B^2 \times 1$ .
6. Apply CS acquisition  $Av_i$  obtaining  $Y_i$  with size  $M \times 1$ .

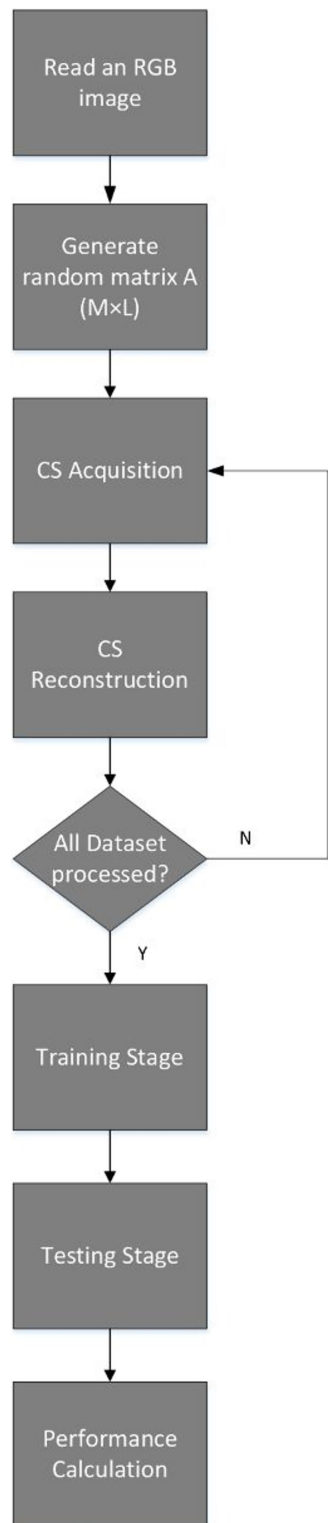
7. Reconstruct  $Y_i$  by OMP using  $A$  obtaining  $\hat{v}_i$  with size  $B^2 \times 1$ .
8. Reshape 1-D vector  $\hat{v}_i$  into 2-D matrix  $\hat{X}_i$  with size  $B \times B$ .
9. Apply 2-D IDCT to  $\hat{X}_i$  obtaining  $\hat{x}_i$  with size  $B \times B$ .
10. Repeat steps 4–9 until all blocks are processed.
11. Combine all blocks from step 10 obtaining the reconstructed image  $\hat{x}(n, m)$  with size  $N \times N$ .
12. Repeat steps 1–11 until we have 400 reconstructed images for training. These images are ready to go to the deep learning training stage.
13. Train the dataset consisting of 400 images using the AlexNet technique in CNN model to obtain the trained network. Save the trained network.

While the steps for the testing phase are as follows:

1. Read an RGB image and resize to the certain resolution ( $N \times N$ ) obtaining  $x(n, m)$ .
2. Generate random normal distributed  $A(M \times L)$  matrix for CS acquisition where  $L = B^2$ ,  $B$  is the size of segmented image and  $M \ll L$ .
3. Apply segmentation of the image into  $B \times B$  pixels obtaining  $x_i (B \times B)$  where  $i$  is a block number of the block-based image.
4. Transform  $x_i$  into DCT domain using 2-D DCT to obtain  $X_i (B \times B)$ .
5. Reshape 2-D matrix  $X_i$  into 1-D vector obtaining  $v_i$  with size  $B^2 \times 1$ .
6. Apply CS acquisition  $Av_i$  obtaining  $Y_i$  with size  $M \times 1$ .
7. Reconstruct  $Y_i$  by OMP using  $A$  obtaining  $\hat{v}_i$  with size  $B^2 \times 1$ .
8. Reshape 1-D vector  $\hat{v}_i$  into 2-D matrix  $\hat{X}_i$  with size  $B \times B$ .
9. Apply 2-D IDCT to  $\hat{X}_i$  obtaining  $\hat{x}_i$  with size  $B \times B$ .
10. Repeat steps 4–9 until all blocks are processed.
11. Combine all blocks from step 10 to obtain the reconstructed image  $\hat{x}(n, m)$  with size  $N \times N$ .
12. Load the trained network and apply testing to the reconstructed image using AlexNet technique and the trained network. Save the testing result.
13. Repeat steps 1–12 for the next testing of the reconstructed image. Calculate the deep learning performance of all testing images according to accuracy, precision, recall and F1-score.

The training and testing process is also illustrated in [Fig. 3](#).

The confusion matrix is one of the parameters used to evaluate the classification method ([Sokolova & Lapalme, 2009](#)). At the testing stage, the stored training data will be processed. The test will be presented in the form of a confusion matrix to obtain precision, recall, F1-scores, and accuracy as shown in [Eqs. \(1\)–\(4\)](#). The true negative (TN) value is the amount of negative data that is correctly classified by the system. True positive (TP) is the amount of positive data that is correctly classified by the system. False positive (FP) is the



**Figure 3** Training and testing flowchart.

Full-size  DOI: [10.7717/peerj-cs.1551/fig-3](https://doi.org/10.7717/peerj-cs.1551/fig-3)

number of positive data that is incorrectly classified by the system. False negatives (FN) are the number of negative data that the system incorrectly classifies.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{F1-score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

The compression performance measurement is expressed in a compression ratio (CR), which is the ratio between the original signal and the compressed signal, as shown in Eq. (5) (Irawati, Suksmono & Edward, 2018).

$$\text{CR} = \frac{\text{uncompressed size}}{\text{compressed size}} \quad (5)$$

## RESULTS

In this study, the system model designed for dataset testing uses the AlexNet architecture with predetermined input parameters as shown in Table 1. The dataset was divided into three out of 600 images: 70% training data with 420 images, 10% validation data with 60 images, and 20% test data with 120 randomly selected images. The dataset used for the CNN process had been reconstructed during the CS process. The programming platform used in this research was Google Collaboratory, starting from the pre-processing, compressive sensing, and classification stages.

The accuracy and loss performance of the training and validation process using the AlexNet method are shown in Fig. 4, while the combination method of block-based CS on AlexNet with a CR of 2× is shown in Fig. 5. Figure 4 shows that the result of the training accuracy was 98% while the validation accuracy was 96%. The result of the training loss was 14% and validation loss was 15%. In Fig. 5, the results showed that the training accuracy was 97% while the validation accuracy was 86%. Loss training performance was 11% and validation loss was 36%.

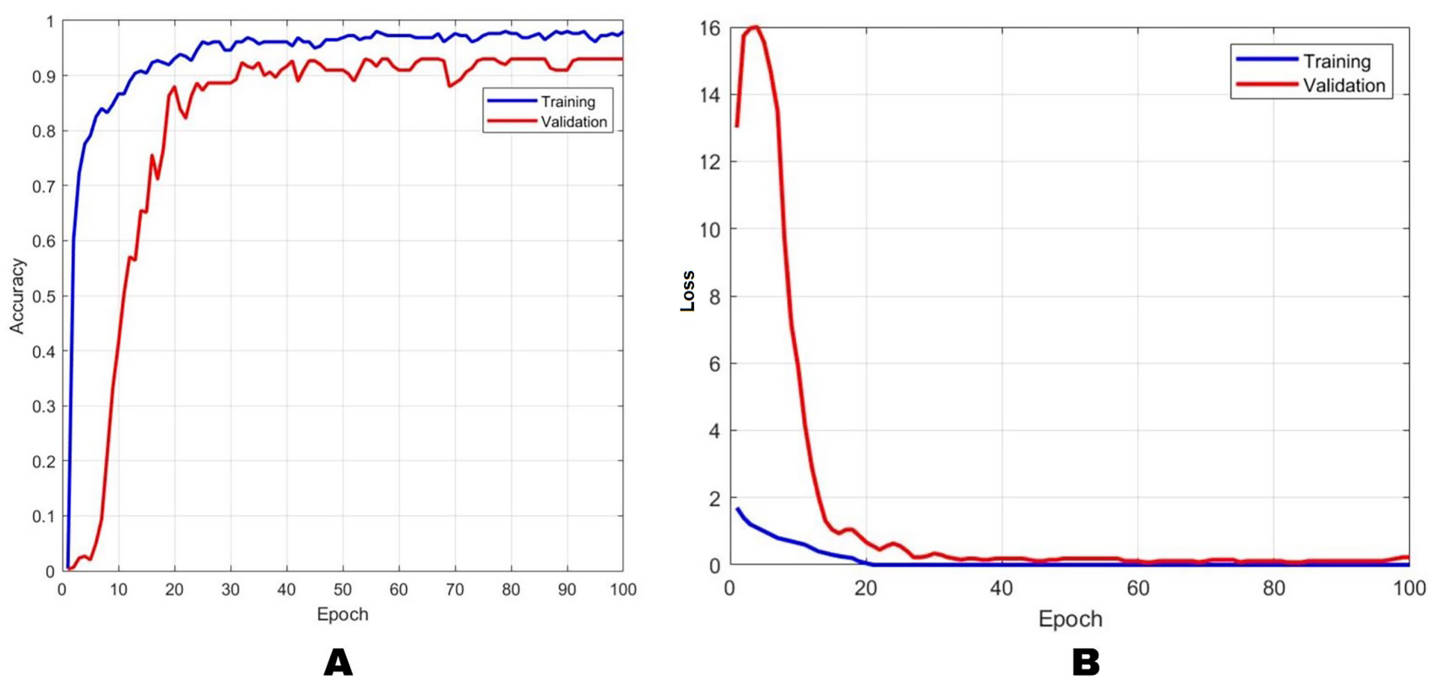
In the testing phase, 20% of the images from the dataset were used. Figure 6 shows that 112 images were correctly predicted according to their class. This condition shows that the application of block-based CS on the CNN model using the AlexNet architecture is promising for classifying vegetables such as carrots, broccoli, radishes, and potatoes. The description of the confusion matrix on the model's ability to predict different types of vegetables is shown in Table 2.

In this article, we tested the effect of several CRs on the system performance when using AlexNet only vs the combination of blocked-based CS and AlexNet. There were five CRs: 32, 16, 8, 4, and 2. Table 3 shows the performance parameters based on the F1-score, recall, precision, testing accuracy, and testing loss. Based on these results, we determined that the



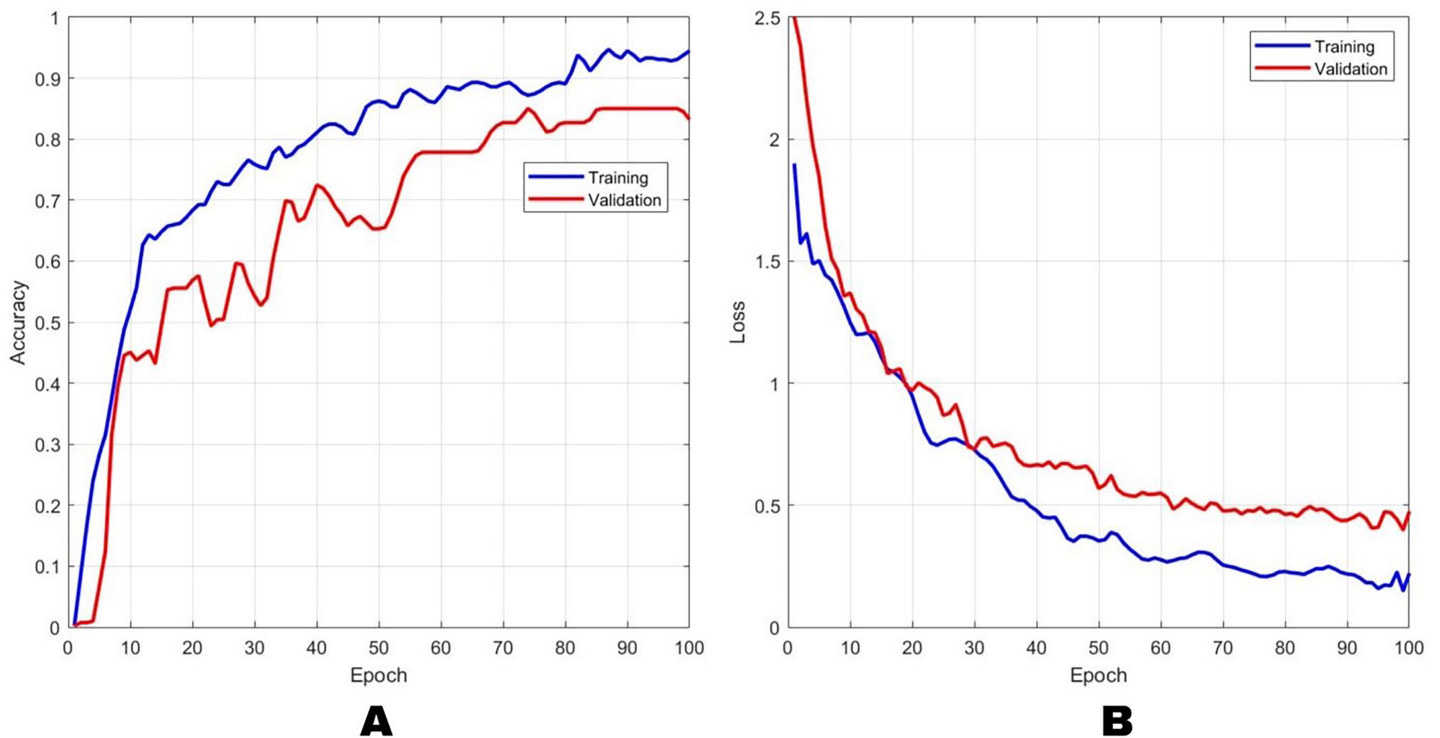
**Table 1** Predetermined input parameter.

Parameters	Value
Resize	256 × 256
CS ratio	6.25%
Optimizer	SGD
Epoch	100
Batch size	64
Dropout	0.4
Learning rate	0.001

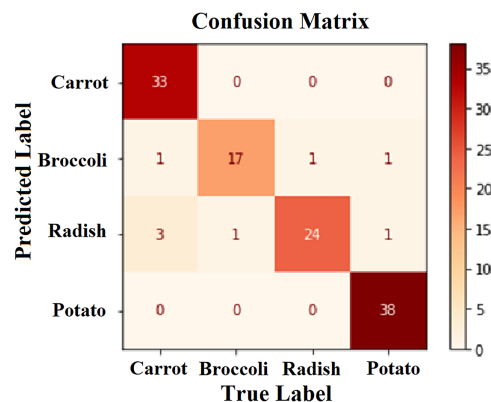


**Figure 4** The results of training and validation on the AlexNet method for (A) accuracy and (B) loss. The blue line indicates the training results, and the orange line represents the testing results. [Full-size !\[\]\(9d188a796ceef961be962a3cd4b57b68\_img.jpg\) DOI: 10.7717/peerj-cs.1551/fig-4](https://doi.org/10.7717/peerj-cs.1551/fig-4)

smaller the CR, the higher the precision, recall, F1, and test accuracy values, and the value of testing loss was directly proportional to CR. There is a trade-off between accuracy and CR. When the CR is getting better or higher, the accuracy will be lower. This is in accordance with the quality of the image which gets worse when the CR is higher, so that the deep learning system will find it more difficult to get training accuracy. However, the lower the accuracy, the higher the CR, which has a good impact on the efficiency of the process. In this case we choose a moderate value between accuracy and CR so that both provide benefits or a win-win solution depending on the performance focus, whether to choose CR or accuracy.



**Figure 5** The results of training and validation on the Block-based CS+AlexNet method for (A) accuracy and (B) loss. The blue line indicates the training results, and the orange line indicates the testing results. [Full-size !\[\]\(1663bb69f307a960345edb0e712f8c02\_img.jpg\) DOI: 10.7717/peerj-cs.1551/fig-5](https://doi.org/10.7717/peerj-cs.1551/fig-5)



**Figure 6** Confusion matrix.

[Full-size !\[\]\(339a16584d5da0f0a3ca4e9ec17bf6a1\_img.jpg\) DOI: 10.7717/peerj-cs.1551/fig-6](https://doi.org/10.7717/peerj-cs.1551/fig-6)

**Figure 7** shows a comparison of the computational time of each CR. Looking at the figure, it can be determined that the larger the CR, the greater the computation time, and the computation time without CS was the longest compared to that of the proposed combination technique.

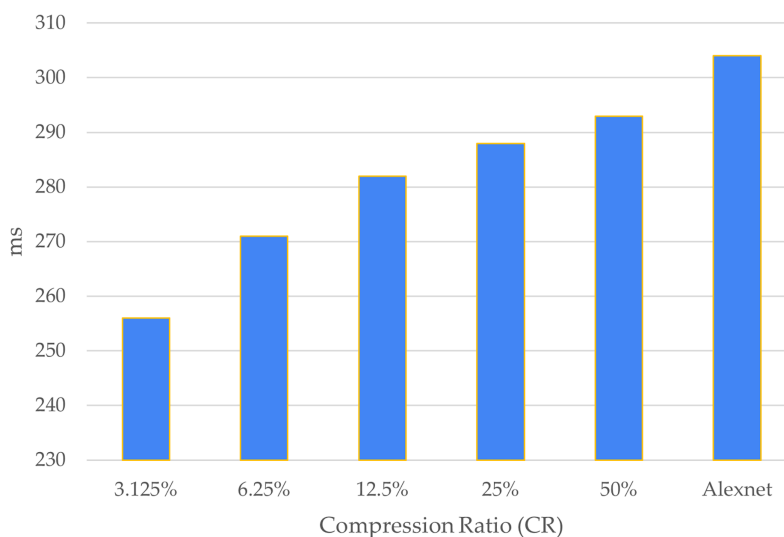
The performance analysis of the proposed method compared with a previous study is shown in [Table 4](#). Both proposed methods have higher accuracy performance than other

**Table 2** Precision, recall and F1-score values of the model to predict vegetables: carrot, broccoli, radish, potato.

Class	Precision	Recall	F1-score
Carrot	0.96	0.88	0.92
Broccoli	1.00	0.93	0.96
Radish	1.00	0.94	0.97
Potato	0.82	0.97	0.89

**Table 3** Performance parameters of the proposed method.

Method	CR	Precision (%)	Recall (%)	F1-score (%)	Testing accuracy (%)	Testing loss (%)
Block-based CS in AlexNet	32	81.25	83.50	82.35	81.66	52.02
	16	88.00	88.00	88.00	88.33	28.01
	8	92.25	93.00	92.62	92.50	17.35
	4	94.75	95.00	94.87	95.00	16.73
	2	96.25	96.25	96.25	96.66	14.90
AlexNet	-	97.45	98.00	97.55	98.00	15.00

**Figure 7** Computational time. The x-axis shows the compression ratio (CR) and the y-axis shows the times.

Full-size DOI: 10.7717/peerj-cs.1551/fig-7

**Table 4** Performance comparison with previous study.

Data train	Data test	Accuracy	Method
19,800	2,400	92.10%	AlexNet ( <i>Zhu et al., 2018</i> )
2,520	1,680	91.17%	AlexNet without augmentation ( <i>Turaev, Almisreb &amp; Saleh, 2020</i> )
2,520	1,680	86.00%	AlexNet with augmentation ( <i>Turaev, Almisreb &amp; Saleh, 2020</i> )
480	120	98.00%	AlexNet (proposed)
480	120	96.66%	Block-based CS in AlexNet (proposed)

previous studies. *Zhu et al. (2018)* used more than 23,200 images and *Turaev, Almisreb & Saleh (2020)* used 4,200 images, while only 600 images were used in this study. However, even with less data, our testing accuracy was higher. For our next research, we plan to optimize the CNN architecture after CS processing by several procedures, such as early stopping and over-fitting reduction as researcher suggestion from *Naushad, Kaur & Ghaderpour (2021)*.

## CONCLUSION

In this study, the CNN architecture AlexNet was used for vegetable image classification. Comparison tests were carried out limited on the AlexNet system only and AlexNet combined with block-based CS. In the CS system, DCT sparsity technique, Gaussian normal measurement matrix, and OMP reconstruction algorithm were used. Both proposed methods were tested on 600 datasets for four types of vegetables. The maximum testing accuracy results reached 98% for the AlexNet method, while the block-based CS in AlexNet method achieved a maximum accuracy of 96.66% with 2× compression rate. This proved that the two proposed systems provided better accuracy than previous studies. Future studies need to explore the addition of plant species and implementation when using sensors.

## ACKNOWLEDGEMENTS

We thank A. V. Senthil Kumar from Hindusthan College of Arts and Science, India for the comments on the manuscript.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

This research was funded by Telkom University through the 2022 International research scheme (No: KWR4.072/PNLT3/PPM-LIT/2022) in collaboration with University of Putra Malaysia. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Grant Disclosures

The following grant information was disclosed by the authors:

Telkom University through the 2022 International Research Scheme: KWR4.072/PNLT3/PPM-LIT/2022.

University of Putra Malaysia.

### Competing Interests

The authors declare that they have no competing interests.

### Author Contributions

- Indrarini Dyah Irawati conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.



- Gelar Budiman conceived and designed the experiments, performed the computation work, authored or reviewed drafts of the article, and approved the final draft.
- Sofia Saidah performed the computation work, authored or reviewed drafts of the article, and approved the final draft.
- Suci Rahmadiani performed the experiments, prepared figures and/or tables, and approved the final draft.
- Rohaya Latip analyzed the data, authored or reviewed drafts of the article, and approved the final draft.

### Data Availability

The following information was supplied regarding data availability:

The code is available at GitHub and Zenodo:

- [https://github.com/gelar1978/CS\\_vegetables\\_Alexnet/tree/main/CS\\_vegetables\\_Alexnet](https://github.com/gelar1978/CS_vegetables_Alexnet/tree/main/CS_vegetables_Alexnet).

- Irawati, Indrarini Dyah, Budiman, Gelar, Saidah, Sofia, Rahmadiani, Suci, & Latip, Rohaya. (2023). Block-based compressive sensing in deep learning using AlexNet for vegetable classification (Final). Zenodo. <https://doi.org/10.5281/zenodo.8358537>.

The data are available at: M. I. Ahmed, S. Mahmud Mamun and A. U. Zaman Asif, "DCNN-Based Vegetable Image Classification Using Transfer Learning: A Comparative Study," 2021 5th International Conference on Computer, Communication and Signal Processing (ICCCSP), Chennai, India, 2021, pp. 235-243, DOI [10.1109/ICCCSP52374.2021.9465499](https://doi.org/10.1109/ICCCSP52374.2021.9465499).

### REFERENCES

- Agrawal Y, Jain B, Mishra SK, Indu S. 2021.** Android application for vegetable and fruit classification. In: *2021 International Conference on Intelligent Technologies (CONIT)*. Piscataway: IEEE, 1–7.
- Ahmadi P, Mansor S, Farjad B, Ghaderpour E. 2022.** Unmanned aerial vehicle (UAV)-based remote sensing for early-stage detection of ganoderma. *Remote Sensing* **14**(5):1239 DOI [10.3390/rs14051239](https://doi.org/10.3390/rs14051239).
- Ahmed MI, Mahmud Mamun S, Zaman Asif AU. 2021.** DCNN-based vegetable image classification using transfer learning: a comparative study. In: *2021 5th International Conference on Computer, Communication and Signal Processing (ICCCSP)*. Piscataway: IEEE, 235–243.
- Chauhan R, Ghanshala KK, Joshi RC. 2018.** Convolutional neural network (CNN) for image detection and recognition. In: *1st International Conference on Secure Cyber Computing and Communication (ICSCCC)*. Piscataway: IEEE, 278–282.
- Doukhan D, Carrive J. 2017.** Investigating the use of semi-supervised convolutional neural network models for speech/music classification and segmentation. In: *The 9th International Conferences on Advances in Multimedia (MMEDIA)*.
- Drayer B, Brox T. 2014.** Training deformable object models for human detection based on alignment and clustering. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **8693(PART 5)**:406–420 DOI [10.1007/978-3-319-10602-1\\_27](https://doi.org/10.1007/978-3-319-10602-1_27).

- Haggag M, Abdelhay S, Mecheter A, Gowid S, Musharavati F, Ghani S. 2019.** An intelligent hybrid experimental-based deep learning algorithm for tomato-sorting controllers. *IEEE Access* 7:106890–106898 DOI 10.1109/ACCESS.2019.2932730.
- Hameed K, Chai D, Rassau A. 2020.** A Progressive weighted average weight optimisation ensemble technique for fruit and vegetable classification. In: *16th IEEE International Conference on Control, Automation, Robotics and Vision (ICARCV)*. Piscataway: IEEE, 303–308.
- Irawati ID, Hadiyoso S, Budiman G, Fahmi A, Latip R. 2022.** A novel texture extraction based compressive sensing for lung cancer classification. *Journal of Medical Signals & Sensors* 12(4):278–284 DOI 10.4103/jmss.jmss\_127\_21.
- Irawati ID, Suksmono AB, Edward IJM. 2018.** Measurement matrix for sparse internet data based compressive sampling. In: *22nd International Computer Science and Engineering Conference (ICSEC)*. Piscataway: IEEE, 1–6.
- Jin X, Che J, Chen Y. 2021.** Weed identification using deep learning and image processing in vegetable plantation. *IEEE Access* 9:10940–10950 DOI 10.1109/ACCESS.2021.3050296.
- Li W, Fu H, Yu L. 2017.** Deep convolutional neural network based large-scale oil palm tree detection for high-resolution remote sensing images. In: *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. Piscataway: IEEE, 846–849.
- Litvak M, Divekar S, Rabaev I. 2022.** Urban plants classification using deep-learning methodology: a case study on a new dataset. *Signals* 3(3):524–534 DOI 10.3390/signals3030031.
- Long J, Shelhamer E, Darrell T. 2015.** Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Piscataway: IEEE, 3431–3440.
- Natarajan S, Ponnusamy V. 2021.** A review on the organic and non-organic fruits and vegetable detection methods. In: *International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*. Piscataway: IEEE, 1–5.
- Naushad R, Kaur T, Ghaderpour E. 2021.** Deep transfer learning for land use and land cover classification: a comparative study. *Sensors* 21(23):8083 DOI 10.3390/s21238083.
- Ng JY-H, Hausknecht M, Vijayanarasimhan S, Vinyals O, Monga R, Toderici G. 2015.** Beyond short snippets: deep networks for video classification. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Piscataway: IEEE, 4694–4702.
- Ni J, Gao J, Deng L, Han Z. 2020.** Monitoring the change process of banana freshness by GoogLeNet. *IEEE Access* 8:228369–228376 DOI 10.1109/ACCESS.2020.3045394.
- Ouyang H, Hu X, Hou H. 2019.** Reconstruction to power quality disturbance signal based on iterative-reweighed-least-squares algorithm. In: *Proceedings—2019 Chinese Automation Congress (CAC)*. Piscataway: IEEE, 3837–3840.
- Pratondo A, Elfahmi E, Novianty A. 2022.** Classification of *Curcuma longa* and *Curcuma zanthorrhiza* using transfer learning. *PeerJ Computer Science* 8(1):e1168 DOI 10.7717/peerj-cs.1168.
- Sachin C, Manasa N, Sharma V, Kumar NAA. 2019.** Vegetable classification using you only look once algorithm. In: *International Conference on Cutting-edge Technologies in Engineering (ICon-CuTE)*. Piscataway: IEEE, 101–107.
- Sokolova M, Lapalme G. 2009.** A systematic analysis of performance measures for classification tasks. *Information Processing & Management* 45(4):427–437 DOI 10.1016/j.ipm.2009.03.002.
- Sudharshan Duth P, Jayasimha K. 2020.** Intra class vegetable recognition system using deep learning. In: *International Conference on Intelligent Computing and Control Systems (ICICCS)*. Piscataway: IEEE, 602–606.

- Sustika R, Subekti A, Pardede HF, Suryawati E, Mahendra O, Yuwana S. 2018.** Evaluation of deep convolutional neural network architectures for strawberry quality inspection. *International Journal of Engineering & Technology* 7(4):75–80.
- Turaev S, Almisreb AA, Saleh MA. 2020.** Application of transfer learning for fruits and vegetable quality assessment. In: *14th International Conference on Innovations in Information Technology (IIT)*. Piscataway: IEEE, 7–12.
- Usman K. 2017.** Introduction to orthogonal matching pursuit. Telkom University, 1–13. Available at <https://korediantousman.staff.telkomuniversity.ac.id/files/2017/08/main-1.pdf>.
- Yuesheng F, Jian S, Fuxiang X, Yang B, Xiang Z, Peng G, Zhengtao W, Shengqiao X. 2021.** Circular fruit and vegetable classification based on optimized GoogLeNet. *IEEE Access* 9:113599–113611 DOI 10.1109/ACCESS.2021.3105112.
- Zeng G. 2017.** Fruit and vegetables classification system using image saliency and convolutional neural network. In: *IEEE 3rd Information Technology and Mechatronics Engineering Conference (ITOEC)*. Piscataway: IEEE, 613–617.
- Zheng YY, Kong JL, Jin XB, Su TL, Nie MJ, Bai YT. 2019.** Real-time vegetables recognition system based on deep learning network for agricultural robots. In: *Proceedings 2018 Chinese Automation Congress, CAC 2018*. Piscataway: IEEE, 2223–2228.
- Zhu L, Li Z, Li C, Wu J, Yue J. 2018.** High performance vegetable classification from images based on AlexNet deep learning model. *International Journal of Agricultural and Biological Engineering* 11(4):190–196 DOI 10.25165/j.ijabe.20181104.2690.