SXC JOURNAL
Volume 1, 2024

Data-driven Smart Farming to Grade and Classify Tomatoes using CNN and FFNN for Agricultural Innovation

Binisha Joshi, Bikal Konda, and Rajan Karmacharya*
Department of Computer Science, St. Xavier's College, Kathmandu, Nepal
*Email: rajankarmacharya@sxc.edu.np**

ABSTRACT

Identifying images poses a challenge in computer vision, but the use of deep learning methods has greatly enhanced the performance of image classification systems. In this research, Convolutional Neural Networks (CNN) and Feed Forward Neural Networks (FFNN) have been utilized for image classification. CNN is extremely effective in picture classification, which extracts relevant information from images using convolutional and pooling layers to minimize the dimensionality of the derived features, while FFNN algorithm is a classic neural network with fully linked layers. It can be used to further process the features extracted by CNN. The study makes use of CNN and FFNN models to train a huge dataset of tomato images to categorize them based on their type, ripeness, and damage status. CNN is found to be more effective in the case of tomato classification as compared to FFNN algorithm in all the use cases. The accuracy for classification of an image (tomato or not) using CNN is 95.83%, type classification using CNN is 81.52%, whereas using FFNN is 66.30%; ripeness grading for CNN is 92.86%, whereas for FFNN it is 57.14%; and damage status grading is 92.86% using CNN and 67.86% using FFNN. Therefore, it can be concluded that quality processing of tomatoes can be improved using CNN.

Keywords: Image classification, tomatoes, CNN, FFNN

Introduction

The agricultural sector is a significant contributor to Nepal's economy – as much as 36% to the GDP – and provides employment to approximately two-thirds of the population. Agriculture spans about 30% of the country's total land area. The government is actively engaged in modernizing, diversifying, commercializing, and promoting the agricultural industry. The adoption of "smart farming" concepts is being explored as a means to enhance productivity and bring about benefits in the agricultural sector (Agriculture old, n.d.).

In the agricultural production system, information and communication technology is integrated into machinery, tools, and sensors. The integration of technologies like the Internet of Things and cloud computing is further advancing this trend by incorporating additional robotics and artificial intelligence into farming practices. The goal of smart farming is to enhance the quality and quantity of agricultural products while reducing the need for human labor, ultimately striving for optimal results (The complete guide to smart farming & agriculture, n.d.).

The nightshade family includes fruits such as the tomato (*Solanum lycopersicum*), which is native to South America. Lycopene, an antioxidant linked to a variety of health benefits, including a lower risk of cancer and heart disease, is mostly found in tomatoes. Tomatoes can be of a variety of hues, including yellow, orange, green, and purple, in addition to its mature color of red. Furthermore, there are numerous tomato subspecies with varying shapes and flavors (Tomatoes: Nutrition facts and health benefits, 2019).

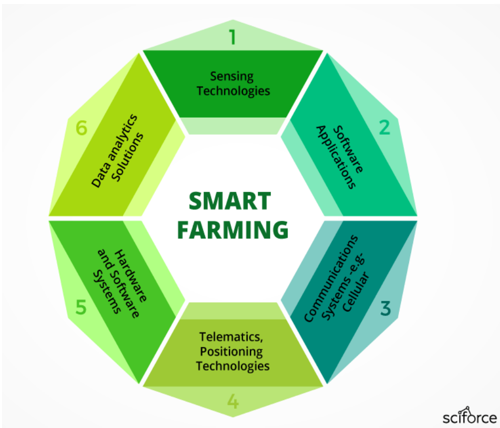


Figure 1. Smart farming concept *Note.* (Amos, n.d.)

There are more than 45+ types and varieties of tomatoes. But this research substantially classifies them into three introductory types of tomatoes (by shape and size) – Classic Tomatoes (Regular- Sized Tomatoes), Cherry Tomatoes (Mini Tomatoes), and Beefsteak Tomatoes (Large Tomatoes).

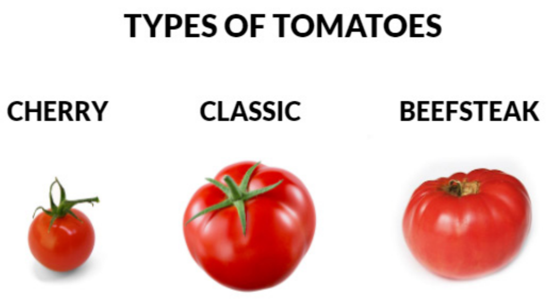


Figure 2. Classification of tomatoes

This study focuses mostly on grading tomato maturity (ripe/unripe and damaged/undamaged) and tomato type (cherry, classic, and beefsteak). To grade and classify the tomatoes, no physical intervention or human labor is required when employing the smart farming idea. The aim of this study is firstly image classification (whether the image is of tomato or not) and then further classification of tomato type and grading its ripeness level and damage status.

Literature review

Tomatoes are becoming more popular, and their quality is becoming more important to the consumers. Tomato ripeness has a big impact on their quality. Traditional tomato categorization is based on farmers' experience, however it is frequently erroneous and time-consuming. Image processing, computer vision, and machine learning methods, especially leveraging pre-trained CNN models like VGG16, VGG19, and ResNet101, have enabled the advancement of top-notch agriculture practices. In determining cherry tomato maturity, the VGG19 model has the highest precision (94.14%). However, only three types of tomatoes were investigated, making high accuracy easier to achieve. More tomato kinds will be added in the future, and the CNN models will be retrained for more precise categorization (Huynh et al., 2021).

Because of their complicated physical features and substantial nonlinearity, categorizing tomatoes and diagnosing the disorders are difficult. Because of the minor distinctions between ripe and unripe tomatoes, even the Human Visual System has difficulty distinguishing between them. In order to effectively classify tomato ripeness, digital images must overcome nonlinear obstacles. The research cited here proposes using a Support Vector Machine (SVM) classifier to differentiate ripe and unripe fruits, and a Multiclass Support Vector Machine (MSVM) classifier to detect faults. Because of its simple testing setup and dependability, the proposed method is ideal for incorporation into the tomato supply chain. Implementing this approach may result in enhanced early differentiation of tomatoes in the value chain, which will benefit the producers (Kumar et al., 2020).

Tomato quality is critical for consistent marketing, and customer impression is heavily influenced by maturity. Recognizing ripeness phases is critical for making high-quality items. Automation of ripeness evaluation can improve quality product manufacturing, benefiting the important worldwide tomato sector. An automated method to classify different stages of tomato ripeness using color characteristics was proposed by the authors of this article. This suggested method extracts and classifies features using Support Vector Machines (SVMs), Linear Discriminant Analysis (LDA), and Principal Components Analysis (PCA) (El-Bendary et al., 2015).

A research carried out in 2020 had the objectives to enhance the packing and market value of cherry fruits by developing an efficient grading system that utilized an upgraded CNN algorithm. Image analysis was used to recognize normal and irregularly shaped cherries in the study. To improve the generalization potential of the CNN, a novel hybrid pooling strategy combining max pooling and average pooling was used. K-Nearest Neighbors (KNN), Artificial Neural Networks (ANN), Fuzzy logic, and Ensemble Decision Trees (EDT) were compared to older methods such Histogram of Oriented Gradients (HOG) and Local Binary Patterns (LBP). The enhanced CNN approach surpassed prior methods in simulation, obtaining an accuracy of 99.4%. This implies that CNN and image processing techniques can efficiently replace traditional cherry grading methods, resulting in improved market control and cherry fruit export

(Momeny et al., 2020).

A research on Sun Bright tomatoes in 2015 investigated how ripeness affects the quality of tomatoes for both processing and consumption. The purpose was to learn how the optical properties of tomatoes, specifically their absorption and scattering properties, altered as they ripened. 281 "Sun Bright" tomatoes at various stages of ripeness were studied using hyperspectral imaging. The study's goal was to create classification models for tomato maturity based on optical absorption (μa) and scattering (μs') spectra. The study attempted to categorize tomatoes into six or three maturity groups by using Partial Least Squares-Discriminant Analysis (PLS-DA) models utilizing these optical characteristics, including solo and a combination (μa &μs', eff) data (Zhu et al., 2015).

Another research in the early 2000s discussed the use of color image processing to assess tomato quality maturity. RGB and Lab* color schemes were employed for the picture analysis. The findings were as follows: The radical regression curve of G(36) was judged to be 70% average correct, the pixels count of G(36) showed the highest correlation coefficient from tomato maturity, the level of a* also rises in accordance with maturation while the b* value did not change significantly, and the average value of a* for the upper surface can be used for the maturity index (Gejima et al., 2004).

Recent advances in computer vision have enabled new agricultural applications, most notably accurate yield estimation for improved harvesting, marketing, and logistics planning. A method for categorizing fresh market tomatoes (Roma and Pear varieties) based on their maturity levels (green, orange, and red) was studied. Color features were combined with a backpropagation neural network (BPNN) classification algorithm in this approach. To capture tomato images, a computer vision-based device was developed, and image processing techniques were employed to isolate tomato targets. The area for color feature extraction was determined as the largest inscribed circle on the tomato's surface, which was divided into five concentric circles for this purpose. The tomato's maturity level was represented by the average hue values from each sub-region. Subsequently, these color characteristics were utilized as inputs in the BPNN to ascertain the maturity levels of the tomato samples (Wan et al., 2018).

After studying many relevant research papers and articles, it was found that many algorithms, such as CNN, SVM, LDA, PCA, etc., were used in the field of smart farming and image classification for classification of fruits such as banana, pear, tomato, cherry, etc. Pre-trained CNN algorithms, such as VGG16, VGG19, ResNet101, were also used in various researches. Of all the algorithms, CNN was found to be a better choice for the research purpose

Research methodology

The datasets used for the training are collected via GitHub along with real time images; for testing purpose, real time images are used. The models were trained on around 10,000 tomato images of different types, ripeness, and damage status.

Figure 3. Training images

Artificial neural network (ANN)

An artificial neuron network (also known as a neural network) is a computer model of how nerve cells in the human brain work. Artificial neural networks (ANNs) use learning procedures to update their answers on their own or to learn when fresh data is presented to them. An artificial neural network has three or more linked layers. The top layer is made up of neurons that are used as input. These neurons send information to deeper layers, which send the final output information to the final output layer (Rouse, 2023). The numerical values that connect the neurons are referred to as weight. The weights between neurons determine the neural network's learning capacity. As artificial neural networks learn, the weights of the neurons change. Weights are assigned at random first. The "activation function" is used to standardize the output of neurons (Artificial neural network - Applications, algorithms and examples, n.d.).

Convolutional neural network

A CNN is a deep learning neural network designed for processing structured arrays of data, such as photos. The convolutional layer is a special type of layer that provides convolutional neural networks their strength. The design of a convolutional neural network is a multi-layered feed-forward neural network formed by progressively stacking many hidden layers on top of one another, allowing it to acquire hierarchical features due to its sequential development (LeCun & Benaissa, n.d.).

CNN has three main layers, namely, Convolutional layer, Pooling layer, and Fully-connected (FC) layer.

Convolutional layer

The convolutional layer is the central component of a CNN, and it is also where majority of the processing occurs. The only components required are input data, a filter, and a feature map. Assume that the input is a color image made up of a 3D pixel matrix. As a result, the input will have three dimensions: height, width, and depth, which correspond to the RGB values in a picture. In addition, we have a feature detector, also known as a kernel or filter, which will traverse the image's receptive fields and assess whether or not the feature is there (Ratan, 2020).

Pooling layer

In convolutional neural networks, the feature map formed by a preceding convolutional layer and a non-linear activation function is often utilized as the basis for pooling. The essential phases of the pooling process are quite similar to those of the convolution procedure. You select a filter and place it over the output feature map of the previous convolutional layer. Based on the type of pooling operation you select, the pooling filter determines the output on the receptive field (the region of the feature map beneath the filter). The most commonly used strategies are max-pooling and average pooling (What is pooling in a convolutional neural network (CNN): Pooling layers explained, 2021).

Fully connected layer

Neural networks are made up of a collection of interdependent non-linear functions. Each function is carried out by a single neuron (or perceptron). In fully connected layers, the neuron changes the input vector linearly using a weights matrix. The result is then transformed nonlinearly using a nonlinear activation function. Each input into the input vector influences every output into the output vector. However, not all weights have an effect on all outputs (Unzueta, 2022).

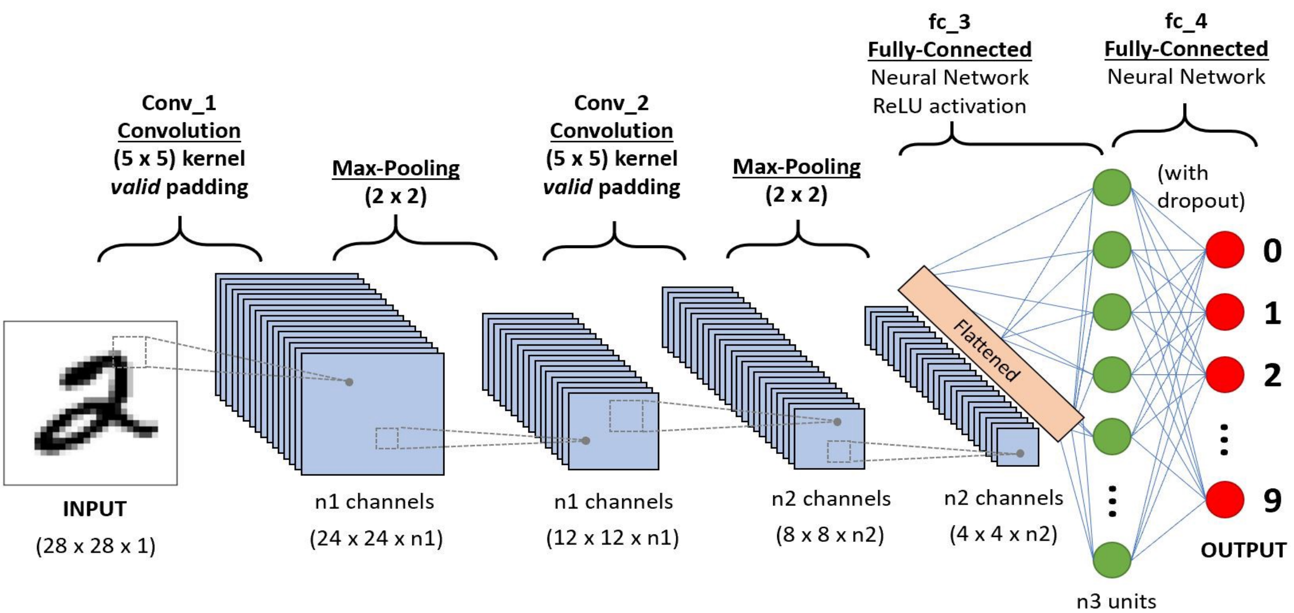


Figure 4. Convolutional neural network
Note. (Ratan, 2020)

Feed forward neural network

Feed-forward networks are artificial neural networks that do not have looping nodes. Because all input is simply transmitted forward, this type of neural network is also known as a multi-layer neural network. The data received at the input nodes transfer across covert layers, and the output at the nodes comprises data flow. There are no network links that can be altered to send data back from the output node.

The following is how a feed forward neural network approximate functions:

- Classifiers are determined by an algorithm employing the equation y = f*(x).
- Therefore, category y is given to input x.
- The feed forward model states that y = f (x; θ). The function's closest approximation is determined by this value (Understanding feed forward neural networks in deep learning, n.d.).

Activation function:

By including an activation function, an artificial neural network may learn complex patterns in data. The output of a neural network is determined by mathematical equations known as activation functions (Activation functions — All you need to know! | By Sukanya Bag | Analytics vidhya, n.d.).

ReLU activation function:

If the input is positive, the rectified linear activation function, or ReLU, which is a non-linear or piecewise linear function, will output the input directly; if the input is negative, it will output zero. It is the activation function that is most frequently employed in neural networks, particularly in convolutional neural networks (CNNs) and multilayer perceptrons. It is written as f(x) = max(0, x) in mathematics (Praharsha, n.d.).

Softmax activation function

The neural network's unprocessed outputs are converted into a vector of probabilities—basically, a probability distribution across the input classes—by the softmax activation function. Think about an N-class multiclass classification issue. The result of the softmax activation is an output vector with N elements, the entry at index i representing the likelihood that a certain input belongs to class i. Mathematically, it is expressed as (Softmax activation function: Everything you need to know, n.d.).

Result Analysis

A total of 70 tomato and 26 non-tomato images were used for testing purpose; a total of 92 images for type classification, which included 27 cherry, 55 classic, and 2 beefsteak tomatoes; 56 images for damage status grading (48 undamaged and 8 damaged tomatoes); and 38 ripe and 18 unripe tomatoes, an accumulation of 56 images for ripeness grading testing. The testing results and comparisons are displayed and discussed below.

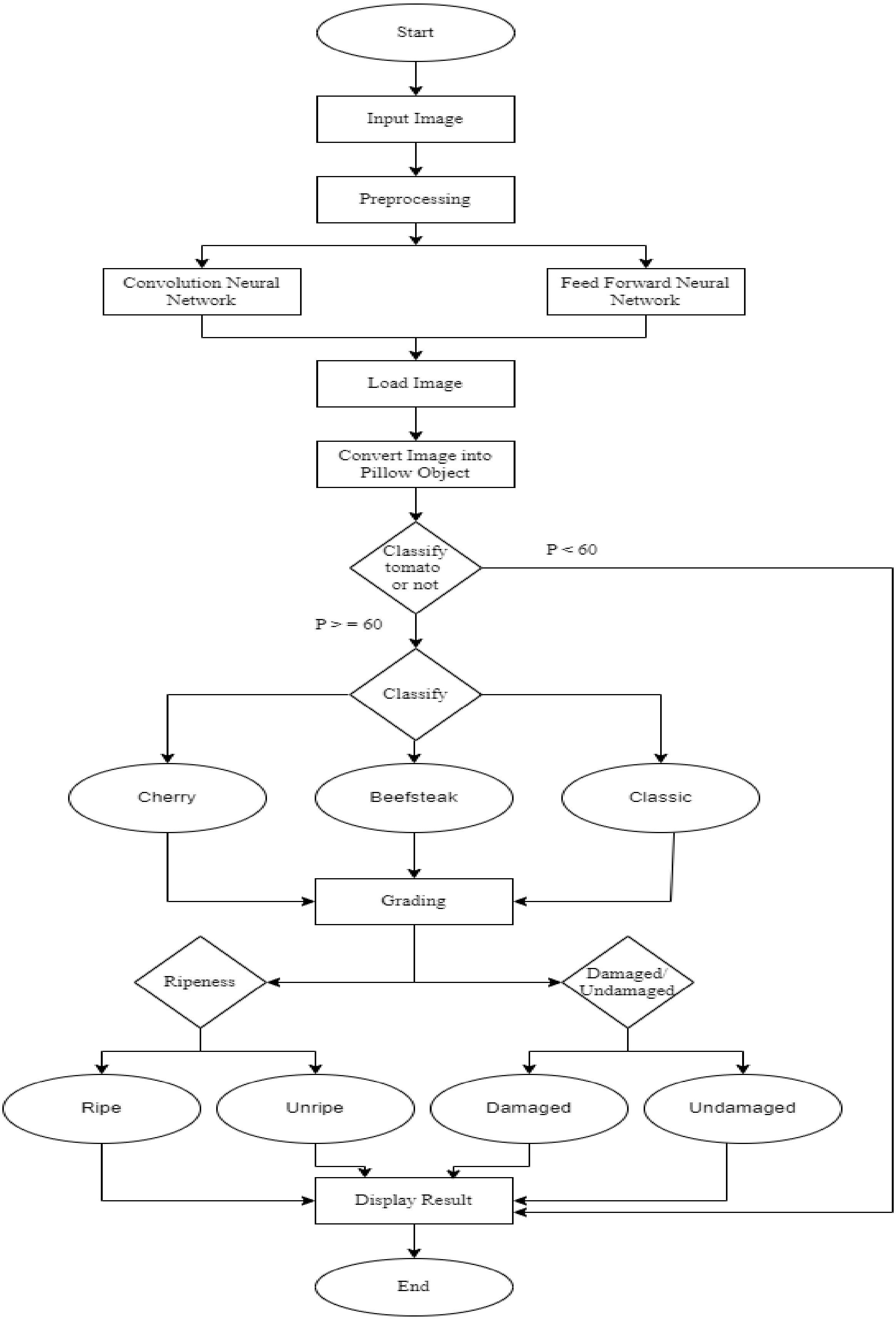


Figure 5. Workflow of the algorithm

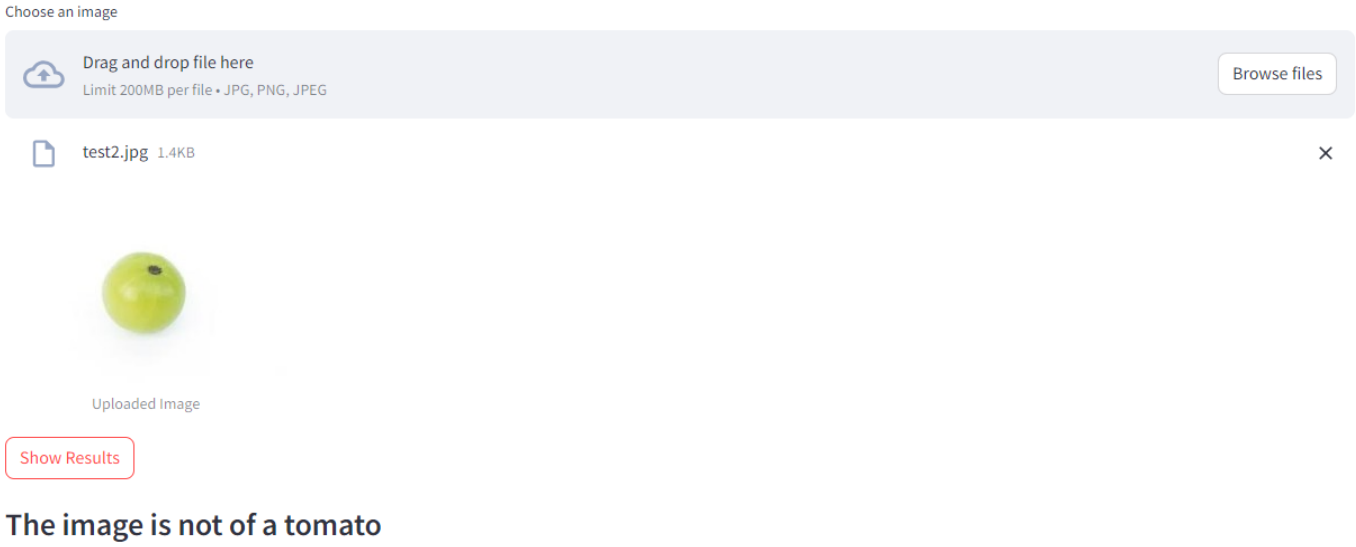
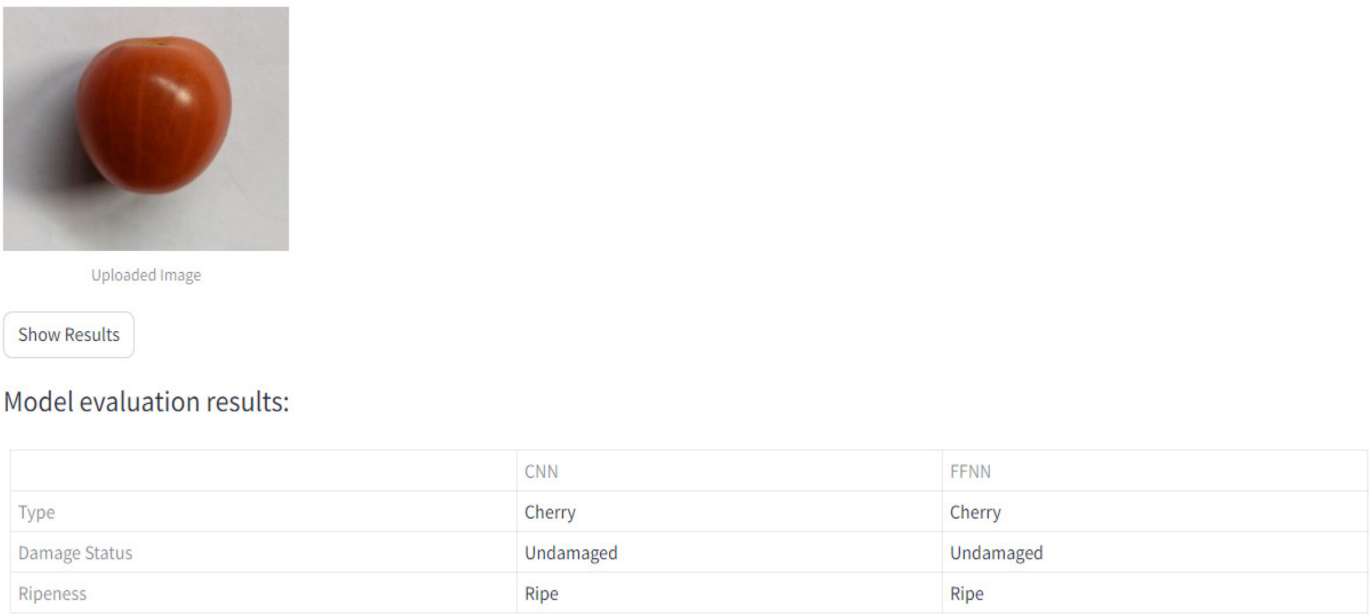


Figure 6. Image not a tomato



	CNN	FFNN
Type	Cherry	Cherry
Damage Status	Undamaged	Undamaged
Ripeness	Ripe	Ripe

Figure 7. Cherry, ripe and undamaged tomato

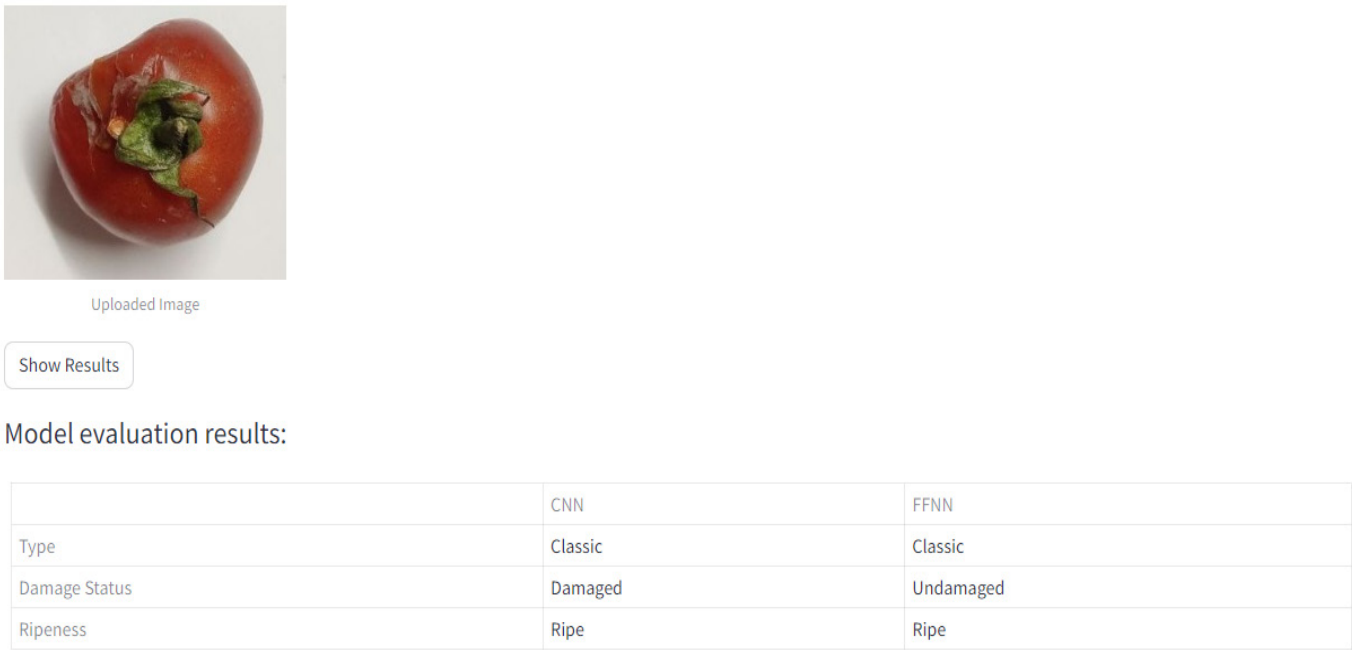


Figure 8. Classic, ripe and damaged tomato

Table 1. Accuracy results

Accuracy	CNN	FFNN
Classification of image (tomato or not)	95.83%	-
Type classification	81.52%	66.30%
Grading ripeness	92.86%	57.14%
Grading damage status	92.86%	67.86%

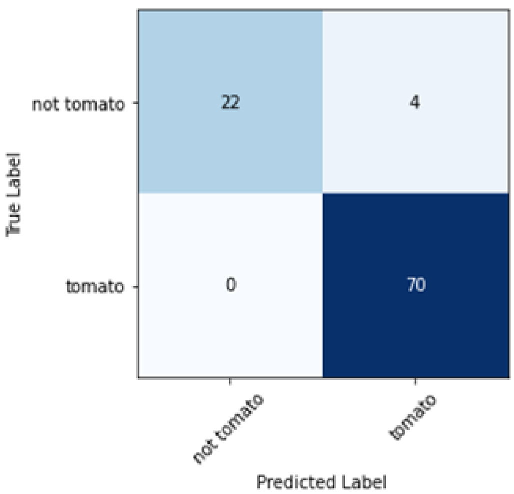


Figure 9. Confusion matrix for image classification (tomato or not) using CNN

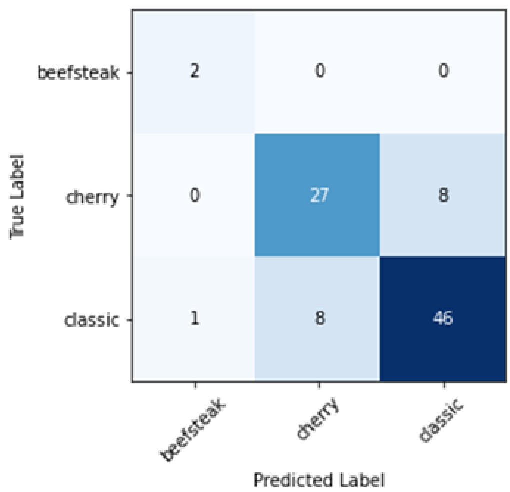


Figure 10. Confusion matrix for type classification using CNN

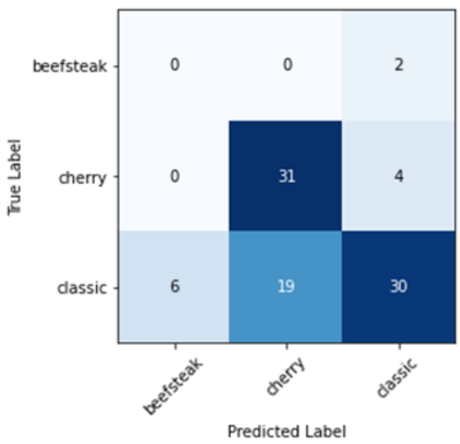


Figure 11. Confusion matrix for type classification using FFNN

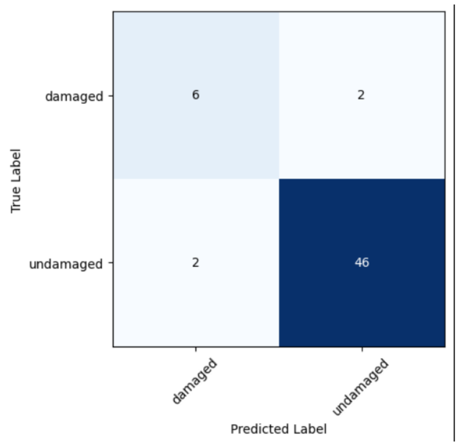


Figure 12. Confusion matrix for grading damage status using CNN

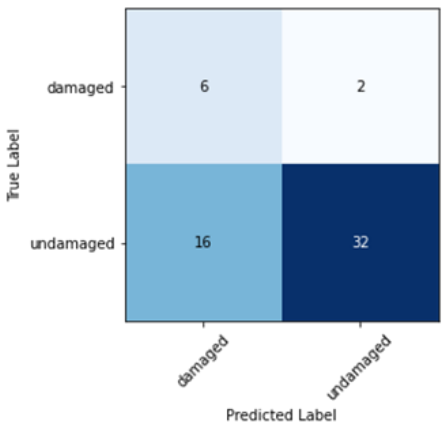


Figure 13. Confusion matrix for grading damage status using FFNN

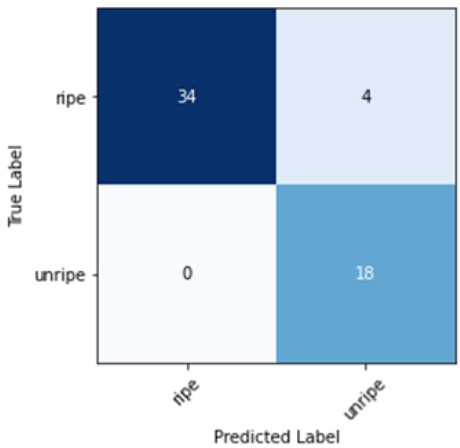


Figure 14. Confusion matrix for grading ripeness using CNN

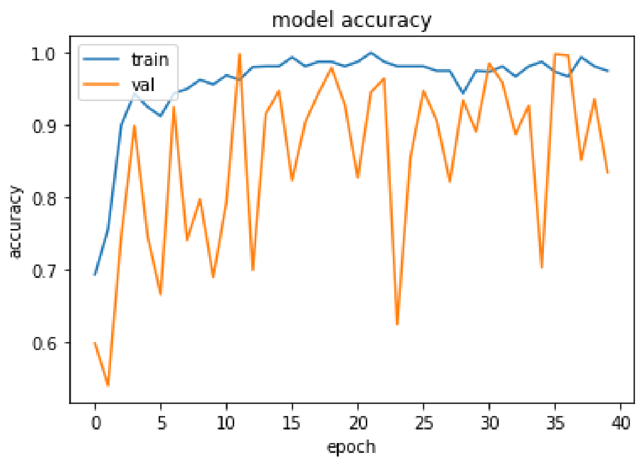


Figure 15. Accuracy curve for image classification (tomato or not) using CNN

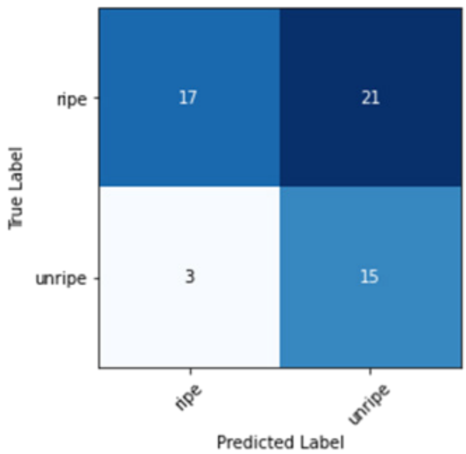


Figure 16. Confusion matrix for grading ripeness using FFNN

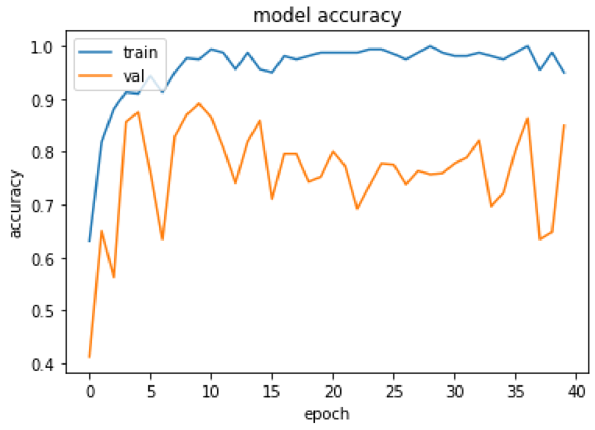


Figure 17. Accuracy curve for type classification using CNN

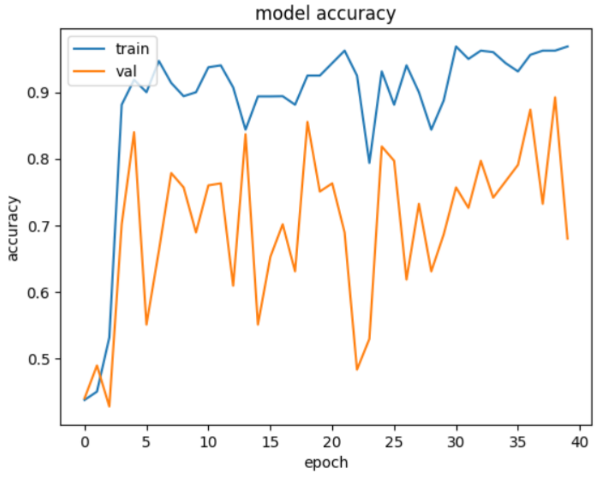


Figure 18. Accuracy curve for type classification using FFNN

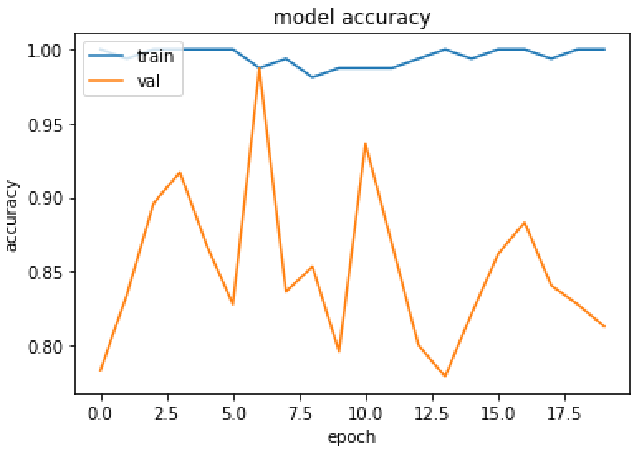


Figure 19. Accuracy curve for grading ripeness using CNN

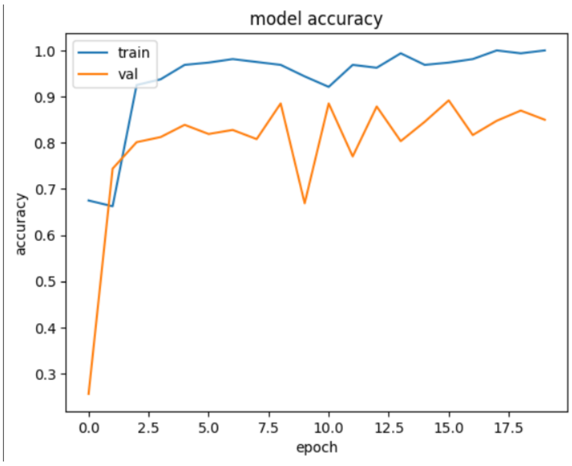


Figure 20. Accuracy curve for grading ripeness using FFNN

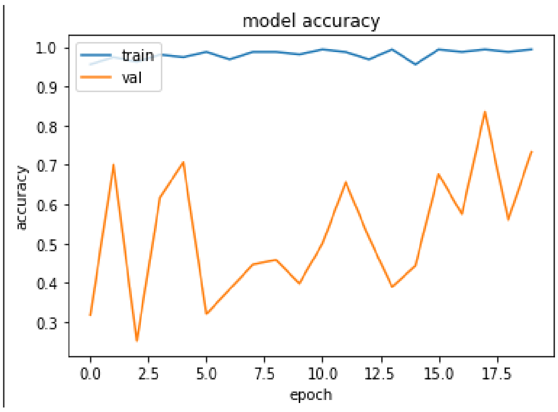


Figure 21. Accuracy curve for grading damage status using CNN

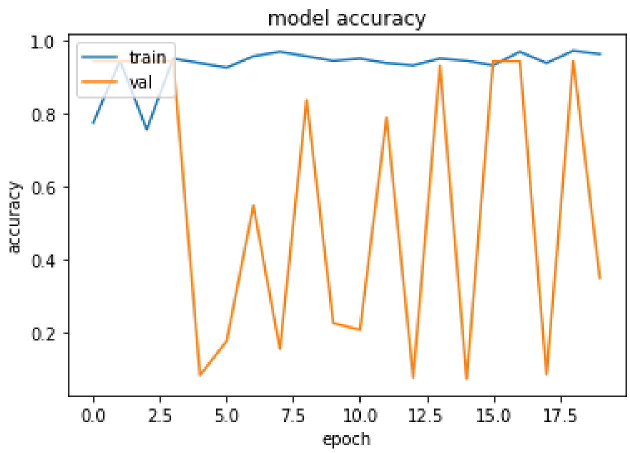


Figure 22. Accuracy curve for grading damage status using FFNN

Conclusion

This paper proposes deep learning neural network algorithms, namely, Convolutional Neural Networks (CNN) and Feed Forward Neural Networks (FFNN), for grading and classification of tomatoes. The tomato classification program is a vital tool for businesses such as food processing, distribution, and restaurants, allowing for effective image-based tomato categorization. CNN is found to be more effective in the case of tomato classification as compared to FFNN algorithm in all the use cases. The accuracies for classification of image (tomato or not) using CNN is 95.83%, type classification using CNN is 81.52% whereas with FFNN it is 66.30%, ripeness grading for CNN is 92.86% whereas for FFNN it is 57.14%, and damage status grading is 92.86% using CNN and 67.86% using FFNN.

CNN's prowess in pattern recognition and shape recognition makes it the best choice for complicated image classification, such as tomato types. It requires minimal preprocessing, has hierarchical feature learning, and is praised for its user-friendliness. CNN employs the use of convolutional layers which extracts low level features like edges and textures to recognize patterns and structure in the images. These features of CNN resulted in better accuracy while grading and classifying tomatoes as compared to FFNN.

Limitations and Recommendations

Potential enhancements in the future may encompass precise classification of a wide range of tomato varieties, integration of emerging characteristics, enlargement of datasets, improvement in adaptability to varying light conditions, and incorporation of machine learning frameworks for increased efficiency and accuracy. CNN algorithm was tested with 4 and 6 hidden layers, but 5 hidden layers gave better results and accuracy. Similarly, the FFNN algorithm uses a total of 4 hidden layers, with 3 and 5 hidden layers also tested for accuracy. Different numbers of hidden layers could also be tested for better accuracy for tomato classification. Adam Optimizer is used for model training whose learning rate is set to 0.001, but different optimizers and learning rates can also be tested for better accuracy. Ultimately, there is a prospect to explore real-time classification as well through the utilization of edge computing.

References

Activation functions — All you need to know! | by Sukanya Bag | Analytics Vidhya. (n.d.). Medium. Retrieved August 30, 2023, from https://medium.com/analytics-vidhya/activation-functions-all-you-need-to-know-355a850d025e

Agriculture Old. (n.d.). Federation of nepalese chambers of commerce & industry (FNCCI). https://www.fncci.org/agriculture-148.html

Amos, Z. (n.d.). *Smart farming: The future of agriculture*. IoT for all. Retrieved August 30, 2023, from https://www.iotforall.com/smart-farming-future-of-agriculture

Artificial neural network - Applications, algorithms and examples. (n.d.). TechVidvan. Retrieved August 30, 2023, from https://techvidvan.com/tutorials/artificial-neural-network/

The complete guide to smart farming & agriculture. (n.d.). Smarter Technologies. Retrieved August 30, 2023, from https://smartertechnologies.com/guides/the-complete-guide-to-smart-agriculture-farming/

El-Bendary, N., Hariri, E. E., Hassanien, A. E., & Badr, A. (2015, March). Using machine learning techniques for evaluating tomato ripeness. *Expert Systems With Applications*, *42*(4). 10.1016/j.eswa.2014.09.057

Gejima, Y., Zhang, H., & Nagata, M. (2004, March 22). Judgment on level of maturity for tomato quality using L*a*b* color image processing. *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*. 10.1109/aim.2003.1225540

Huynh, D. P., Van Vo, M., Van Dang, N., & Truong, T. Q. (2021, March). Classifying maturity of cherry tomatoes using Deep Transfer Learning techniques. *IOP conference series*, *1109*(1). 10.1088/1757-899x/1109/1/012058

Kumar, S., Esakkirajan, S., Bama, S. S.,

& Keerthiveena, B. (2020, July). A microcontroller based machine vision approach for tomato grading and sorting using SVM classifier. *Microprocessors and Microsystems*, *76*. 10.1016/j.micpro.2020.103090

LeCun, Y., & Benaissa, A. (n.d.). *Convolutional neural network definition*. DeepAI. Retrieved August 30, 2023, from https://deepai.org/machine-learning-glossary-and-terms/convolutional-neural-network

Momeny, M., Jahanbakhshi, A., Jafarnezhad, K., & Zhang, Y. (2020, August). Accurate classification of cherry fruit using deep CNN based on hybrid pooling approach. *Postharvest Biology and Technology*, *166*. 10.1016/j.postharvbio.2020.111204

Praharsha, V. (n.d.). *ReLU (Rectified Linear Unit) activation function*. OpenGenus IQ. Retrieved August 30, 2023, from https://iq.opengenus.org/relu-activation/

Ratan, P. (2020, October 28). *Convolutional neural network architecture | CNN Architecture*. Analytics Vidhya. Retrieved August 30, 2023, from https://www.analyticsvidhya.com/blog/2020/10/what-is-the-convolutional-neural-network-architecture/

Rouse, M. (2023, June 17). *What is an artificial neural network (ANN)? - Definition from Techopedia*. Techopedia. Retrieved August 30, 2023, from https://www.techopedia.com/definition/5967/artificial-neural-network-ann

Softmax activation function: Everything you need to know. (n.d.). Pinecone. Retrieved August 30, 2023, from https://www.pinecone.io/learn/softmax-activation/

Tomatoes: Nutrition facts and health benifits. (2019, December 17). Allure Vanguard. Retrieved August 30, 2023, from https://allure.vanguardngr.com/2019/12/tomatoes-nutrition-facts-and-health-benefits/

Understanding feed forward neural networks in deep learning. (n.d.). Turing. Retrieved August 30, 2023, from https://www.turing.com/kb/mathematical-formulation-of-feed-forward-neural-network

Unzueta, D. (2022, October 18). *Fully connected layer vs convolutional layer: Explained*. Built In. Retrieved August 30, 2023, from https://builtin.com/machine-learning/fully-connected-layer

Wan, P., Toudeshki, A., Tan, H., & Ehsani, R. (2018, March). A methodology for fresh tomato maturity detection using computer vision. *Computers and Electronics in Agriculture*, *146*. 10.1016/j.compag.2018.01.011

What is pooling in a convolutional neural network (CNN): Pooling layers explained. (2021, December 5). Programmathically. Retrieved August 30, 2023, from https://programmathically.com/what-is-pooling-in-a-convolutional-neural-network-cnn-pooling-layers-explained/

Zhu, Q., He, C., Lu, R., Mendoza, F., & Cen, H. (2015, May). Ripeness evaluation of 'Sun Bright' tomato using optical absorption and scattering properties. *Postharvest Biology and Technology*, *103*. 10.1016/j.postharvbio.2015.02.007