Düzce University Journal of Science & Technology, 13(2025) 857-867

Düzce University
Journal of Science & Technology

Research Article

Tomato Sorting System Based on Type Using Deep Learning

Eren Yiğit GÜLEM [a], Boran DURSUN [a], Hayrettin TOYLAN [a,*]

[a] *Department of Mechatronics, Faculty of Technology, Kırklareli University, Kırklareli, TÜRKİYE*
** Corresponding author's e-mail address: hayrettintoylan@klu.edu.tr*
DOI: 10.29130/dubited.1569117

ABSTRACT

The tomato is a vegetable that is cultivated globally and plays a significant role in the culinary traditions of numerous countries. This vegetable needs to be separated after collection to meet the requirements of obtaining different flavors outside the growing season. This study focuses on the automatic separation of Rio tomatoes, which are preferred for tomato paste and sauces, from Fujimaru tomatoes using artificial intelligence and image processing techniques. Convolutional neural network (CNN), R-CNN, and Fast-CNN models were used to classify two different tomato types, and their performances were compared. According to the experimental results, it was observed that the CNN model achieved 94.1% accuracy, 93.5% precision, 94.7% recall, and 94.1% F1 score in the classification of Rio type tomatoes, and 92.4% accuracy, 91.8% precision, 93% recall, and 92.4% F1 score in the classification of Fujimaru type tomatoes. The hardware and software components used in this study are low cost, flexible, and modular. Experimental results show that the proposed model and system have high accuracy, precision, and efficiency rates.

Keywords: *Deep learning, convolution neural network (CNN), tomatto sorting*

Derin Öğrenme Kullanarak Türüne Bağlı Domates Sınıflandırma Sistemi

ÖZ

Domates, dünya genelinde yetiştirilen ve ülkelerin yemek kültürlerinde önemli bir yer tutan sebzedir. Bu sebzenin yetiştirildiği mevsim dışında ve farklı lezzetler elde etme gereksinimlerini karşılamak için toplandıktan sonra ayrıştırılması gerekir. Bu çalışma, yapay zeka ve görüntü işleme tekniklerini kullanarak salçalık ve soslarda tercih edilen Rio cinsi domateslerin Fujimaru cinsi domateslerden otomatik olarak ayrılması üzerine odaklanmaktadır. İki farklı domates türünü ayırmak için konvolüsyon sinir ağı (CNN), R-CNN ve Fast-CNN modelleri kullanılmış ve performansları karşılaştırılmıştır. Deneysel sonuçlara göre, Rio cinsi domateslerin sınıflandırılmasında, CNN modelin %94.1 doğruluk, %93.5 hassasiyet, %94.7 geri çağırma ve %.94.1 F1 skoru; Fujimaru cinsi domateslerin sınıflandırılmasında, %92.4 doğruluk, %91.8 hassasiyet, %93 geri çağırma ve %.92.4 F1 skoru ile daha başarılı sonuçlar elde ettiği görülmüştür. Bu çalışmada kullanılan donanım ve yazılım bileşenleri düşük maliyetli, esnek ve modülerdir. Deneysel sonuçlar, önerilen modelin ve sistemin yüksek doğruluk, hassasiyet ve verimlilik oranlarına sahip olduğunu göstermektedir.

Anahtar Kelimeler: *Derin öğrenme, evrişimli sinir ağı, domates ayırma*

Received: 18/10/2024, Revised: 16/01/2025, Accepted: 13/02/2025

I. INTRODUCTION

Turkey is third in global tomato production, following China and India, with an output of approximately 13 million tonnes. The majority of tomato production (65.5%) was destined for the table tomato market, with the remainder (34.5%) allocated to tomato paste [1]. Traditionally, tomato sorting and grading are done manually, making the process time-consuming and susceptible to human error [2]. With the advancing technology, these systems can be controlled fully automatically using machine learning and artificial intelligence techniques.

The agriculture sector is undergoing a significant transformation with the integration of automation and artificial intelligence (AI) technologies. These developments are not only revolutionizing traditional farming practices but also paving the way for smart agriculture. This is essential to satisfy the increasing global demand for food. Automation and AI technologies offer significant advantages such as increasing farming productivity, ensuring quality control, reducing labor costs, and enabling precision farming [3]. Tomato grading is a critical step in production processes to improve quality control and efficiency [4]. AI-powered automated systems can significantly alleviate these issues by providing consistent and accurate grading results.

Wan et al. presented the backpropagation neural network (BPNN) method to determine the ripeness levels of Roma and Pear type tomatoes and categorized them as green, orange, or red according to their ripeness level. The accuracy rate obtained with this method was determined to be 99.31% [5]. Kaur et al. used a backpropagation neural network (BPNN) to rank and sort a total of 53 tomato images using camera setups developed by themselves. The proposed method achieved an accuracy rate of 92% [6].

The most widely used deep learning method for analysing image data is convolutional neural networks, known as CNN. Agarwal et al [7] used CNN models to identify tomato plant diseases and compared the performance of these models with traditional machine learning methods and obtained more successful results. Priyadharshini et al. applied CNN, R-CNN, Fast R-CNN and Faster R-CNN models to identify diseases in tomato leaves. In particular, it was stated that the model developed with Faster R-CNN performs better than the CNN, R-CNN, and Fast R-CNN models, with an accuracy rate of 98% [4]. The authors, Vini and Rathika (2024) proposed a fast and efficient CNN architecture is proposed for the classification of tomato leaf disease. The integration of stochastic gradient descent optimizers into the CNN network increased the success rate of the network. The proposed method achieved an accuracy rate of approximately 99.39% in disease classification [8]. Sun et al. focused on diagnosing tomato pests and diseases using Squeeze and SE Net (SSNet), a Convolutional Neural Network (CNN). In their research, they examined the effects of dataset balance and data volume on model performance. The results show that SSNet achieves accuracy rates of 98.80% and 98.39% for tomato pests and diseases, respectively [9]. In their study, Amune et al. (2024) focused on the post-harvest sorting procedure of tomatoes. In their research, they classified tomatoes according to size, color, and quality criteria. In quality classification, they preferred the CNN model due to its performance [10]. Research reveals that existing studies are more focused on the pre-harvest processes of tomatoes. However, there is limited research on the real-time post-harvest sorting of different tomato types. This study aims to address the problem of sorting tomato types, which is currently a labor-intensive process that requires a great deal of human observation.

The main objective of this study is to automatically separate different types of tomatoes using artificial intelligence and image processing techniques. Within the scope of this study, two types of tomatoes classified as table and tomato paste are separated by analyzing their visual characteristics through a Convolutional Neural Network (CNN)-based model. The developed low-cost prototype system has achieved successful results in separating tomatoes in real-time. In addition, the performance of the CNN model is compared with R-CNN and Fast-CNN models, and the results are presented.

The remaining part of this paper is organized as follows: Section 2 covers the experimental setup and the deep learning methods employed. Section 3 presents the dataset used in this study along with the details of the experimental results. Finally, Section 4 provides the conclusions.

II. MATERIALS AND METHODS

A. EXPERIMENTAL SETUP

The Tomato Sorting System (TSS) developed in this research consists of a camera, a conveyor belt, and an automatic sorting unit (Figure 1). A camera with a resolution of 1.3 megapixels (H × V = 1280 × 1024 pixels) and a frame rate of 25 frames per second (fps) was used to obtain detailed images of tomatoes. The automatic sorting unit consists of servo motors and a conveyor motor that directs the tomatoes into the appropriate bins according to their type.

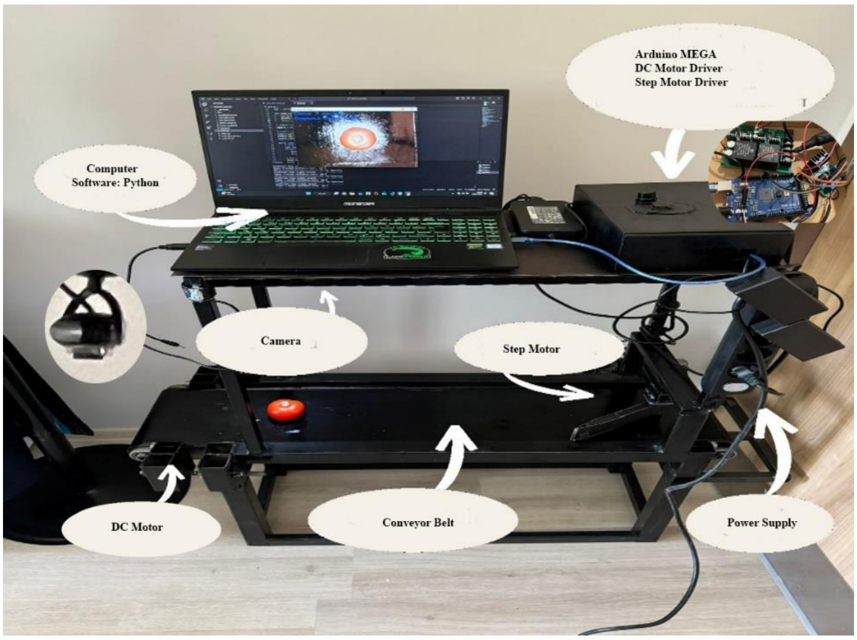


Figure 1. *Tomato sorting system*

TSS software components are developed using Python and its extensive libraries. Python is a flexible programming language widely preferred in machine learning and image processing [3]. TensorFlow and Keras libraries were used to create and train the CNN model. These libraries allow for the development of complex neural networks and effective model training. They are also very useful for processing large data sets and performing high-performance computations for deep learning tasks [11]. The Arduino IDE was preferred for programming the Arduino Mega microcontroller. This hardware enables cost-effective control of components. Table 1 presents the properties of the components used in the TSS.

Table 1. Electronic components of the proposed TSS

Unit	Parameters	Value
Brushed DC motor	Rated Voltage (V)	24
	Rated torque (Ncm)	14
	Rated speed (min)	3250
	Rated current (A)	2.84
	Ambient temperature	-30°C to +40°C
Arduino Board	Type	Mega 2560
	Microcontoller	Atmega 2560
	Operation Voltage	3.3-5 V
	Dijital Pins	54
	Analog Pins	16
Camera	Resulution	1280x1224
	Fps	25
	Color Space	RGB
	Image sensor	1.3 Megapixel 1/3" CMOS

B. DEEP LEARNING

Deep learning methods use artificial neural networks to mimic human thinking and learning processes. These methods can work effectively on text, audio, image, and video data and offer high accuracy rates in extracting their own features. However, these processes require large amounts of input and computers with high processing power. In this section, the deep learning methods used in this study for the classification of tomato types will be briefly reviewed.

B. 1. Convolutional Neural Network (CNN)

Tradiatonal image classification involves describing the entire image by either manually extracting features or using feature selection methods, followed by employing a classifier to identify the object category. The difference from traditional image processing in CNN models is that this model is an end-to-end learning process. The original image enters the network, the training and prediction processes are performed in the network, and the result is obtained [12]. Consequently, the extraction of features from the image is particularly important. CNN (Convolutional Neural Network) is one of the most widely used algorithms with the ability to process, classify, and segment images. CNNs are frequently preferred in tasks such as pattern recognition, classification, and segmentation, especially in large image data sets. The CNN model consists of multiple layers. While the first layers extract features from images, the last layers use these features for classification. The basic layers of the model are as follows:

The first layer is the convolution layer. This layer extracts important features by analysing the relationships between the pixels of the images. The pooling layer is used to eliminate unnecessary parameters when necessary. This layer ensures that important features are preserved while reducing the size of the input data. In the next stage, the feature map matrix is converted into vectors and then the Fully Connected (FC) layer comes into play. The vectorised features are fed into a neural network. The FC layer takes these vectorised features and integrates them to build a model. As a result, the softmax function is usually used in the final layer of a CNN and calculates the probabilities of 'n' different events [13].

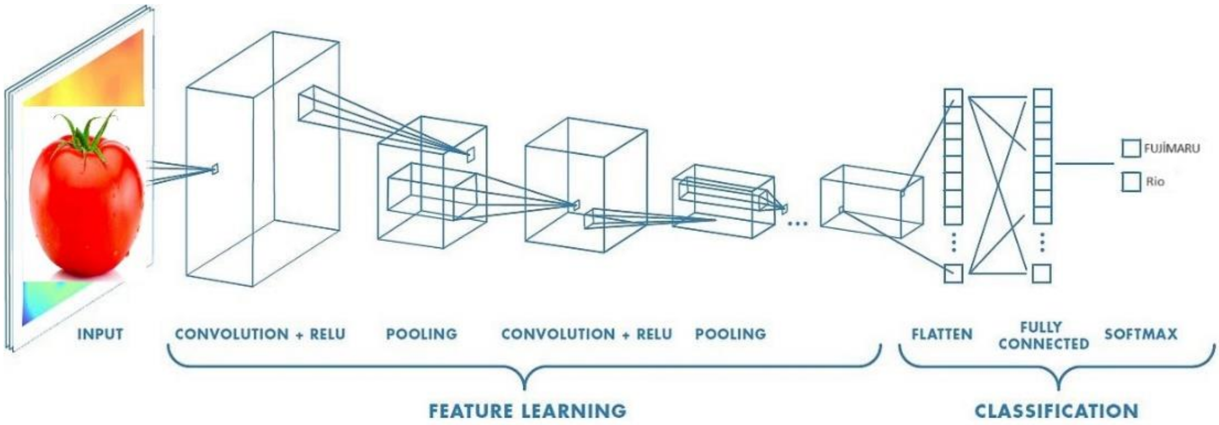


Figure 2. *Convolutional Neural Network (CNN) structure used in the TSS*

The CNN used for tomato sorting consists of an input layer, convolutional layers with ReLU activation, pooling layers, fully connected layers, and a softmax output layer (Figure 2). The parameters of the CNN,R-CNN and Faster R-CNN model applied in this study are presented in Table 2.

Table 2. *CNN,R-CNN and Faster R-CNN model parameters*

Parameters	**CNN/R-CNN /Faster R-CNN**
Number of convolution layer	4
Number max pooling layer	4
Activation function	Relu/Softmax
Learning rate	0.0001
Number of epoch	20
Batch size	32
Optimizer	SGD (Stochastic Gradient Descent)

The flowchart of the CNN model with four convolutional layers designed to separate tomatoes according to their types is presented in Figure 3.

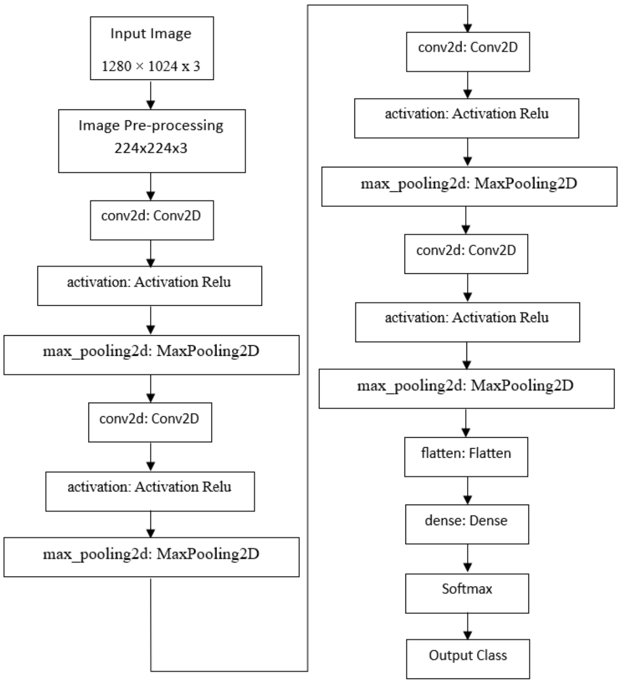


Figure 3. *Flowchart of CNN model for the TSS*

B. 2. Region-based Convolutional Neural Networks (R-CNN)

R-CNN extracts approximately 2,000 possible object regions to identify and classify the positions of objects in an image. In the first stage, a selective search algorithm is used to identify possible object regions. This algorithm generates many potential object regions on the image, which are then processed by a CNN (Convolutional Neural Network). Each object region is processed through a CNN model and converted into feature vectors. The resulting feature vectors are evaluated by a separate SVM (Support Vector Machine) classifier for each object region[14].

The most significant advantage of R-CNN is its high accuracy in object detection. However, the model has a high computational cost and long training times. Additionally, because the R-CNN model runs a CNN for each region separately, the processing time can be considerably longer, particularly when working with large images.

B. 3. Faster R-CNN

Faster R-CNN is an enhanced version of the R-CNN model and offers faster and more efficient performance in object detection tasks. Fast R-CNN inputs the entire image into the CNN instead of processing the 2000 region proposals generated by selective search one at time. It determines region proposals from the feature map obtained from the image and resizes them to a fixed size using a RoI (Region of Interest) pooling layer. Finally, A softmax layer is used to evaluate the results from the model and produce probabilistic results[15].

C. PERFORMANCE EVALUATION METRICS

The performance of the compared models is presented based on the following evaluation metrics described by Sokolova and Lapalme [16].

Accuracy: The number of correctly classified positive examples is denoted "True Positive" (TP), while the number of correctly classified negative examples is denoted "True Negative" (TN). The accuracy of a classification is calculated by dividing the number of correctly classified examples by the sum of the number of correctly and incorrectly classified (False Positive (FP), False Negative (FN)) examples. Accuracy is calculated using eq. (1).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

F1 Score: This measurement considers both precision and recall simultaneously, representing a harmonic average of the two variables. F1 Score is calculated using eq. (2).

$$F1\,Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{2}$$

Precision: It indicates the proportion of samples classified as positive that are genuinely positive. Precision is calculated using eq. (3).

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

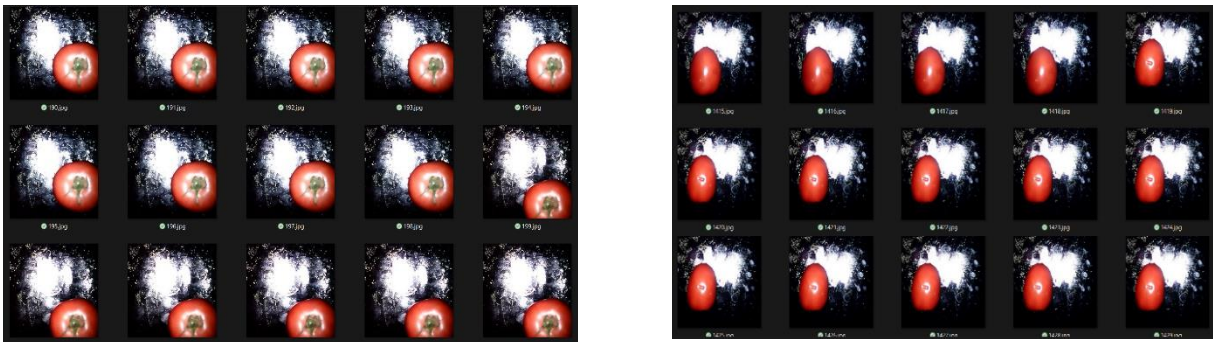
Recall: It indicates the proportion of samples classified as positive to the total positive in the data. Equality is given 4.

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

III. RESULTS AND DISCUSSION

A. DATASET

The dataset was created with images obtained from two different tomato types in a homogeneous background and constant lighting conditions. This dataset contains a total of 1,785 images of Fujimaru and Rio species taken from different angles. The collected image data was divided into 80% training, 10% validation, and 10% test data. The training process was carried out without using transfer learning techniques.



(a) ***(b)***

Figure 4. *a)FUJİMARU tomato data, b)Rio tomato data.*

The images captured by the camera are quite large and may contain unwanted noise. To improve computational efficiency and enhance performance, a three-stage image preprocessing approach was implemented. First, the images were resized to dimensions of 224x224x3. Next, the Red, Green, and Blue (R, G, B) color components were normalized to a range of [0, 1]. Finally, random noise in the images was reduced using the Gaussian blur technique.

B. PERFORMANCE OF CNN MODELS

The compared CNN models were trained and tested with the above mentioned dataset. Accuracy, F1 score, precision and recall were calculated using Equation (1), Equation (2), Equation (3) and Equation (4) respectively. The performance curve of the Fast R-CNN model is shown in Figure 5.

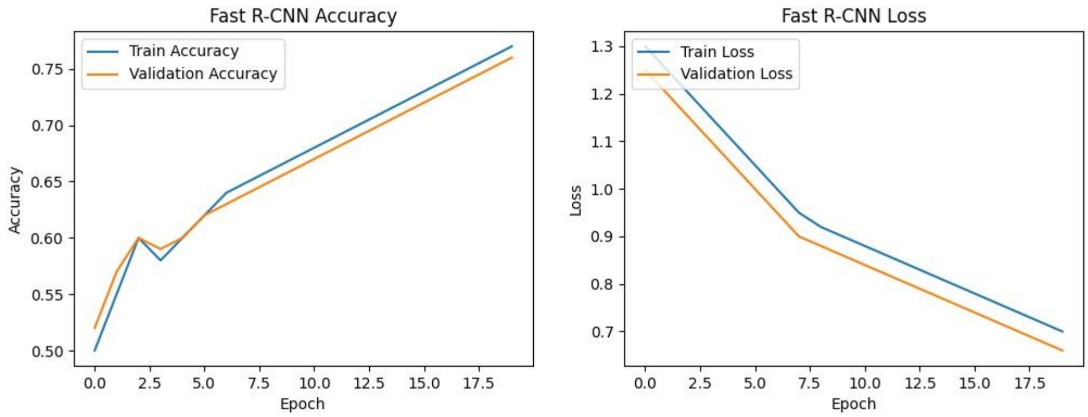


Figure 5. *Fast R-CNN model performance curves*

The training and validation success of the Fast R-CNN model demonstrate a continuous increase across both graphs. The training success is approximately 75%, which is very close to the validation success. Furthermore, the training and validation losses also decrease over time, reaching a value of approximately 0.7. This demonstrates that the overall performance of the model is satisfactory and that there is no overfitting in the validation process. The models were evaluated with 20 epochs, a batch size of 32, and a learning rate of 0.0001. The performance curve of the R-CNN model is shown in Figure 6.

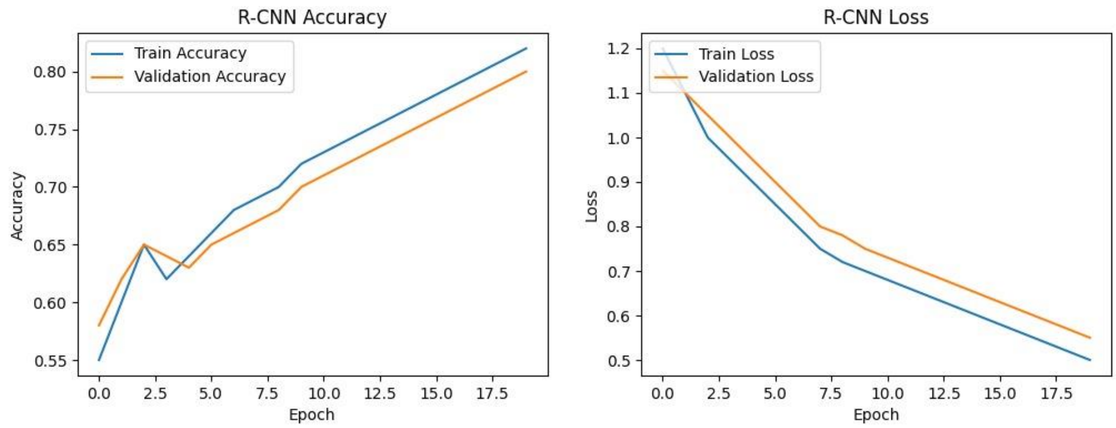


***Figure 6**. R-CNN model performance curves*

While the R-CNN model shows a similar accuracy increase to Fast R-CNN, it shows a superior performance with accuracy rates reaching approximately 80%. The training and validation losses decreased to approximately 0.6. These findings show that the R-CNN model provides an effective performance with higher accuracy and lower loss. The models were evaluated with 20 epochs, 32 Batch Size and 0.0001 learning rate. The performance curve of the simple CNN model is shown in Figure 7.

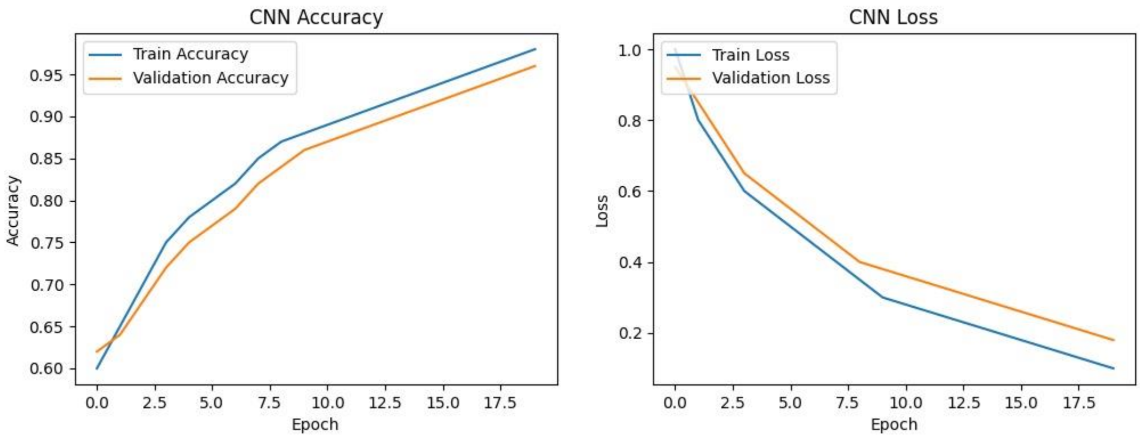


Figure 7. *CNN model performance curves*

The simple CNN model achieved the highest accuracy and lowest loss values compared to the other two models. While the training accuracy increased to approximately 95%, a similar increase was observed in the validation accuracy. Training and validation losses decreased continuously to approximately 0.2. These results show that the CNN model exhibits superior performance in both training and validation phases and that the model learns effectively. The models were evaluated with 20 epochs, 32 Batch Size and learning rate of 0.0001. Table 3 presents CNN, R-CNN and Fast R-CNN models performance

Table 3. *CNN, R-CNN and Fast R-CNN models performance*

Model	Dataset	Accuracy (%)	Precision	Recall	F1 Score
CNN	FUJİMARU Tomato	92.4	91.8	93.0	92.4
	Rio Tomato	94.1	93.5	94.7	94.1
R-CNN	FUJİMARU Tomato	76.3	80.2	82.5	81.3
	Rio Tomato	81.2	79.2	83.6	81.3
Fast R-CNN	FUJİMARU Tomato	75.3	78.6	85.6	81,9
	Rio Tomato	72.6	74.5	75.9	75,1

In this study, the recognition process is carried out on images taken with the help of a camera. Therefore, there is not more than one different pattern in the visual field. For this reason, it can be said that the simple CNN model gives a successful result compared to other models. While simple CNN networks are successful in identifying the class of the object, they are not sufficient to find its location. In this study, since there is no need to find the location at the same time, models that would cause additional computational costs were avoided. Simple CNN has a lower computational cost and runs faster than R-CNN and Faster R-CNN. This is an important advantage when working with large data sets.

IV. CONCLUSION

In this study, we developed a Tomato Sorting System (TSS) capable of classifying tomatoes by type using a deep learning approach. To identify the most effective deep learning algorithm for the TSS, we evaluated the performance of three models: Simple CNN, R-CNN, and Fast R-CNN. Our results showed that the Simple CNN model achieved the highest accuracy, with 92.4% for Fujimaru type tomatoes and 94.1% for Rio type tomatoes.

The most important advantages of the developed TSS system are the use of low-cost hardware and the ability to easily adapt to different agricultural crops. However, various improvements can be made to further increase the speed of this system, which can classify about 100 tomatoes per hour. Among these improvements, the use of industrial cameras and powerful processors stand out. In conclusion, this study demonstrates the potential of artificial intelligence and image processing techniques to improve agricultural applications and emphasizes the need for continuous research and development in this field.

ACKNOWLEDGEMENTS: This work is supported by TUBİTAK 2209 A

V. REFERENCES

[1] S. Eğilmez, "Ürün raporu Domates 2022" (2025, 15 January). *Tarımsal Ekonomi ve Politika Geliştirme Enstitüsü*, Erişim: https://arastirma.tarimorman.gov.tr/tepge/Belgeler/PDF%20%C3%9Cr%C3%BCn%20Raporlar%C4%B1/2022%20%C3%9Cr%C3%BCn%20Raporlar%C4%B1/Domates%20%C3%9Cr%C3%BCn%20Raporu%202022-364%20TEPGE.pdf

[2] P. Wan, A. Toudeshki, H. Tan, H., and R. Ehsani, "A methodology for fresh tomato maturity detection using computer vision," *Computers and electronics in agriculture*, vol. *146*, pp. 43-50, 2018.

[3] S. Kaur, A. Girdhar, and J. Gill, "Computer vision-based tomato grading and sorting," In *Advances in Data and Information Sciences: Proceedings of ICDIS-2017*, Singapore, 2018, pp. 75-84.

[4] M. Agarwal, S. K. Gupta, and K. K. Biswas, "Development of Efficient CNN model for Tomato crop disease identification," *Sustainable Computing: Informatics and Systems*, vol. 28, p.100407, 2020.

[5] M. M. Mijwil, K. Aggarwal, R. Doshi, K. K. Hiran, and M. Gök, "The Distinction between R-CNN and Fast RCNN in Image Analysis: A Performance Comparison," *Asian Journal of Applied Sciences*, vol. 10(5), pp. 429-437, 2022.

[6] V. Shankar, V. Kumar, U. Devagade, V. Karanth, and K. Rohitaksha, "Heart disease prediction using CNN algorithm," *SN Computer Science*, vol. 1(3), p. 170, 2020.

[7] M. Maity, S. Banerjee, and S. S. Chaudhuri, "Faster r-cnn and yolo based vehicle detection: A survey," In *2021 5th international conference on computing methodologies and communication (ICCMC)*, India, 2021, pp. 1442-1447.

[8] S. L. Vini, and P. Rathika, "TrioConvTomatoNet: A robust CNN architecture for fast and accurate tomato leaf disease classification for real time application," *Scientia Horticulturae*, vol. 330, p.113079, 2024.

[9] L. Sun, K. Liang, Y. Wang, W. Zeng, X. Niu, and L. Jin, "Diagnosis of tomato pests and diseases based on lightweight CNN model," *Soft Computing*, vol. 28(4), pp. 3393-3413, 2024.

[10] A. Amune, S. Shinde, A. Samargade, S. Suryawanshi, and Y. Shirsat, "Tomato TriSort: Smart Sorting System for Size, Color, and Pest Detection," In *2024 4th International Conference on Sustainable Expert Systems (ICSES),* Nepal, 2024, pp. 539-545.

[11] M. Sokolova, and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Information processing & management*, vol. 45(4), pp. 427-437, 2009.

[12] L. Chen, S. Li, Q. Bai, J. Yang, S. Jiang, and Y. Miao, "Review of image classification algorithms based on convolutional neural networks," *Remote Sensing*, vol. 13(22), p.4712, 2021.

[13] Y. Gao, W. Liu, and F. Lombardi, "Design and implementation of an approximate softmax layer for deep neural networks," In *2020 IEEE international symposium on circuits and systems (ISCAS)*, Spain, 2020, pp. 1-5.

[14] G. Priyadharshini, and D. R. Judie Dolly, "Comparative Investigations on Tomato Leaf Disease Detection and Classification Using CNN, R-CNN, Fast R-CNN and Faster R-CNN," *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS),* India, 2023, pp. 1540-1545.

[15] M. Maity, S. Banerjee, and S. S. Chaudhuri, "Faster r-cnn and yolo based vehicle detection: A survey," In *2021 5th international conference on computing methodologies and communication (ICCMC)*, India, 2021, pp. 1442-1447.

[16] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, pp. 1137-1149, 2017.