International Journal of Computer Sciences and Engineering
Vol.13, Issue.2, pp.39-48, February 2025
ISSN: 2347-2693 (Online)
Available online at: www.ijcseonline.org

Research Article

Diseased Area Segmentation Using a Novel Gray-Scale Thresholding Algorithm and Classification Using a New Deep CNN Model for Apple Fruit Sorting

M. Henila[1*]

[1]Dept. of BCA, Ethiraj College for Women, Chennai, TamilNadu, India

*Corresponding Author:

Received: 30/Dec/2024; **Accepted**: 31/Jan/2025; **Published**: 28/Feb/2025. **DOI**: https://doi.org/10.26438/ijcse/v13i2.3948

Abstract: A novel Gray-Scale Thresholding Method [GSTM] for segmenting the region of interest and a Deep Convolutional Neural network model for the apple fruit sorting system has been proposed in this paper. First, the GSTM method converts the acquired colour image into a grayscale image and calculates the threshold value using the Gray pixel values. The acquired colour apple image was then segmented using the calculated threshold value to extract the diseased/defective part alone for further processing. Second, a Deep Convolutional Networking model was designed to classify the apple images as sound or diseased/defective apple images. The result obtained using the GSTM was compared with similar Grayscale thresholding methods like Otsu and Kapur. It was found that GSTM's execution time was less and the visual segmentation was good compared to the other two methods in extracting the diseased/defective area. The Deep Convolutional Network using GSTM segmented images gave a classification/sorting accuracy rate of 91.67%.

Keywords: Thresholding Algorithm, Region of Interest, Classification, Deep Neural Network, Apple Fruit Sorter

1. Introduction

The food industry is significantly growing and facing productivity problems. Increasing demand for the world's food resources has made the food supply complex and multifaceted with the worldwide supply chain bringing many benefits but, also presenting a major food safety headache. If the supply chain is long and complicated, there is a high risk of contamination and spoilage of food items like apple fruits. In the apple food industry, sorting is done by hand to remove the fruits which are contaminated or spoiled and are unsuitable to market or storage due to damage by mechanical injuries, insects, diseases, etc. This work is usually carried out manually and done before washing the apple fruits. Removing damaged/diseased apples from the healthy ones reduces losses by preventing secondary contamination. Apple sorting is done either at the farm level or in the pack houses. The sorting technology reduces the risk of contamination, driving up the safety of food on the production line, and also offers a smart investment for processors and manufacturers. Sorting and quality analysis systems should be effective in enhancing food safety. Also, the ability to employ a data management system when using machines ensures more real-time monitoring and controlling of apple production lines, creating further efficiencies in a safer, and higher quality apple samples going to consumers. Also, the sensor-based sorting technologies allow the computer processors increases the accuracy in identifying defects/disease in the apple fruit produce. Ensuring higher resolution and better contrast, the technology used will continue to improve the quality and food safety of apple fruits. Apple fruit sorters are coupled with intelligent software [embedded in micro-chips], with high-end cameras for recognizing each apple's colour, size, texture and shape; as well as the colour, size, and location of a defect on the apple fruit's surface. Few intelligent apple fruit sorters even allow the user to define a defective apple fruit based on the total defective surface area of any given apple image. The camera sensors capture an apple's response to the energy source and image processing techniques are used to manipulate the raw data acquired. The image processing extracts and categorizes information about specific features of the diseased/defective area. Then, the user defines the accept or reject criteria to determine which is good and a sound apple sample. The image processing art and science lies in developing algorithms or methods that maximize the effectiveness of the sorter while presenting a simple user interface to the operator. So, in this paper to maximize the use of a sorter, a novel image segmentation algorithm to identify the region of interest using a thresholding method and Deep Convolution neural network model for sorting the same has been proposed. Image segmentation has been referred to as a process to divide images into more useful regions to identify the region of interest before the image analysis procedure. Many research works are done for segmenting an image. Among them, using threshold values for segregating the foreground and background images are more advantageous in many ways. It has been the simplest among all methods for segmenting the region of interest from any RGB colour image. The thresholding method is based on the pixel intensity level values in an image. Thresholding method work with the assumption that the pixel values can be grouped in a certain range of pixel intensity values in one class and the remaining pixel intensity values in another class. The accuracy of classification process in any image analysis highly depends on how well the region of interest from an image is segmented.

2. Related Work

Few research works that had been done in the area of segmentation and classifications are discussed below; A survey on colour feature extraction methods after segmentation was conducted and presented as an article by Srikanth, Kumar, and Kumar[20]. Zhang[25], has done a survey on neural network classification methods and have presented an article. A review on the identification, classification, and grading of fruits using machine learning was done and presented by Behera et al.[4]. A colour image segmentation using an automatic seed region growing algorithm using the Watershed method has been proposed by Garcia-Lamont et al. [6]. Abd Elaziz et al. [1], have proposed a multi-objective and multi optimization algorithm for Gray-image segmentation and tested it against three other similar optimization methods. It was found that the proposed method showed good segmentation accuracy than the compared methods. Ishak [9], proposed a 2D multidimensional thresholding based on Renyi and Tsali's entropies and tested the accuracy rate using multimodal and noisy images. A clustering validation method was adapted by Baya´ etal. [3] to design a clustering stability, that automatically segmented the images and it was observed that the results in segmenting were good. Yang et al. [22], have introduced two steps as a classification process. In the first step, lactating sows were segmented using top-view images using a fully-connected convolution network. And in the second step, refinement of the image was done using the fully-connected neural network and Otsu threshold algorithm that gave a better segmentation result and 96.6% accuracy rate. Sha et al.[18], have proposed a robust 2D Otsu's thresholding method for segmenting colour images with a region post-processing method for dealing with the noise in the segmented images (Gaussian noise and salt & pepper noise). Their result showed a good accuracy rate. An image segmentation method based on Gray stretch and threshold algorithm combines discrete wavelet transform with the principles of the largest between class variance for segmenting color images was proposed by Liu et al.[13]. In the article by Manic et al.[14], an image multi- thresholding method based on Kapur and Tsallis entropy concepts, as well as the fire fly algorithm, was proposed and the segmentation of the proposed method was found to be good. A supervised learning task to classify images and an insight about an unsupervised classification was proposed by Lele [12].An innovative training criteria based on analysis on error back propagation was proposed by Xin and Wang[21], and their experimental results proved that M^3CE can enhance the performance of cross-entropy. Ireri et al.[8], proposed a tomato grading machine vision system. The proposed system detected calyx and stalk scar at an average accuracy of 95.15% for both defected and healthy tomatoes by histogram thresholding value based on the mean green-red pixel values of the regions of interest. Defective regions were classified by RBF-SVM classifier using the L*a*B*colour space pixel values. The model gave an accuracy rate of 98.9%. Richtsfeld et al. [17], introduced a framework for segmenting RGB images by processing the data hierarchically. After pre-clustering parametric surface patches were estimated. Different relations between patch-pairs are calculated, which was used to derive perceptual grouping principles, and support vector machine classification was employed to learn the perceptual grouping. Also, the proposed framework was able to segment objects, even if they are overlapped or jumbled in Video scenes. Gal Lego and Pardàs[5], have proposed a novel foreground segmentation system that performs a more complete Bayesian segmentation between foreground and background class. The results presented in this paper showed that the system was robust in segmenting the foreground object and the background. Gongal et al.[7], used a CCD camera and a time-of-flight light-based 3D camera for estimating the size of an apple in tree canopies. To measure the fruit size, the major axis value was estimated based on (i) the 3D coordinates of pixel values on the apple surfaces, and (ii) the size of the individual pixel values within apple surfaces in 2D space. Using 3D coordinates, and the distance between pixel pairs within apple surface regions were calculated.69.1% was the accuracy rate for estimating the major axis using 3D coordinates. The accuracy of size estimation was 84.8% when the pixel-based method was used. Pratondo et al. [16] have proposed a frame work that integrated machine learning with a region-based active contour model. Classification probability scores from machine learning algorithm, which are regularized using a non-linear function, is used to replace the pixel intensity values during energy minimization. The k-nearest neighbours and the support vector machine were integrated with the Chan-Vese method and the results obtained were compared with the normal methods of Chan-Vese and Li et al. The frame-work that was proposed gave a better accuracy and was less sensitive to parameter tuning. Yossy et al. [24] introduced a new method to sort mangoes as ripe and not ripe. Using the colour feature, this method sorted the mangoes with 94% accuracy rate. To utilize the CNN's potential power, an image classification method using multi-stage feature input a new method was proposed by Yim et al. [23] with an increased classification accuracy rate of 0.38%, 3.22%, and 0.13% respectively, when tested using CIFAR-10, NORB, and SVHN datasets. Sharma et al. [19] have presented an analysis on performances of various convolutional neural networks for identifying objects in real-time video scenes and it was found that GoogLe Net, ResNet50 was better than Alex Net in classifying the objects. Also, it was found that convolutional neural networks varied across different categories of classifications. An approach for spot segmentation of micro array using global segmentation method was proposed by Karthik and Manjunath [11] and it was found that the performance of the proposed method was accurate than Fuzzy C-means and Morphological segmentation. Nesar et al. [2] in their work have proposed a machine vision method to identify disease in apple fruits. It is found that the nearest neighbour classifier and the Euclidean distance using 80 training apple samples generated the best accuracy rates, at 100% for the stem and 97.5% for the calyx. Zhi Yong et al. [26] have proposed a apple sorting method. In this method, fuzzy and traditional PID algorithms are used to simulate the operation of brushless DC motor. By using the machine learning models for colour detection, the SVM model finally gave the classification of three types of apple samples with an accuracy rate of 96.7%.

Limitations of existing research works area a) Accuracy rate of segmentation using most of the methods are still very low. b) The number of samples used for testing is less. c) Segmentation speed is low d) Due to improper segmentation, classification or any other image analysis was inaccurate.

To bridge this gap, this paper proposes a) a novel thresholding method (algorithm) for segmenting the region of interest, b) a thresholding method that enhances the segmentation speed and c) a deep Convolutional neural network model that will sort the apple as diseased and sound with a good accuracy rate. Therefore, this paper has been organized under three sections. Section 2 gives the insight of the proposed methodology GSTM, section 3 discusses the results obtained using the proposed methodology GSTM and the designed Deep Convolutional neural network model for classification, section 4 gives the conclusion.

3. Materials and Methods

3.1 Image Acquisition and Dataset

Apple fruit images were acquired using an image-capturing system with a CCD camera, and two LED lights. The camera used for acquiring apple images was, SonyW810 compact CCD camera with 6x optical zoom with 20.1MP image-sensor, built-in autofocus, 6x optical zoom, and the number of effective pixels 20.4 MP. The sensor type was 1/2 3 type(7.76mm) Super HAD *CCD*^{TM} with focal length ranging from 4.6 to 27.6mm.10 to 15 images of each apple fruit were acquired and the proposed method was applied to sort them. Figure.1 shows the block diagram of the image-capturing system. Using the camera, 6 to 12 images of an apple fruit at 360 degrees was acquired when it passes through the conveyor belt. Acquired images were then sent to a computer system connected to the camera via a cable. The proposed methods can be coded, and embedded in microchips of an apple fruit sorter for the sorting process. The experimental images were of dimensions 235x214, 275x183, 224x224, and 226x223 pixels. These apple images were acquired from different types of apple samples with diseases like rot, scab, worm, cork spot, powdery mildew, etc., and sound apple samples. Table.1 shows details about the number of training and testing sample images of apples used for this experiment. Data augmentations like hearing, rotating, rescaling, width shifting, height shifting, flipping, filling, and zooming was applied to these 540 training apple images to get a total [540 x 8] of 4320 training/validation apple input samples.

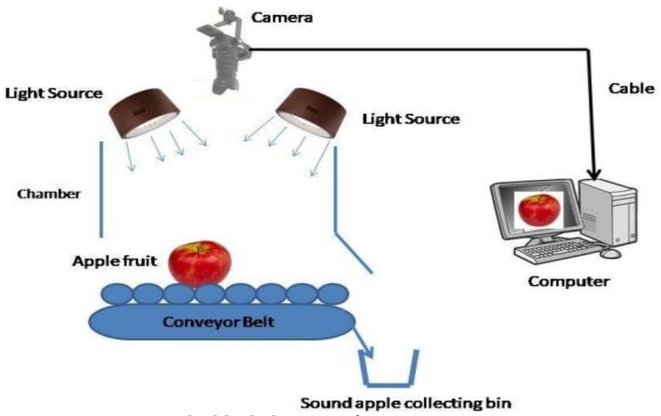


Figure 1: The block diagram of image acquisition system

3.2 Proposed Method

The proposed method involves two different phases. During the first phase, the given training and testing apple images were segmented to obtain the region of interest (defective/diseased area) using the proposed GSTM algorithm/method. In the second phase, all segmented apple images are trained and tested for classification/sorting using the proposed deep convolutional neural network model. The same is diagrammatically represented in Figure.2 Segmenting an image into a meaningful form of objects is an important job in the image processing process.

Thresholding algorithms were used for segmenting the required foreground region of interest from an image from its background. This proposed thresholding method/algorithm can be used for extracting the diseased/defective region of interest from an apple's image. Also, it is a stochastic thresholding approach which works on discrete pixel values. The threshold value was calculated using the grayscale values of the given image. Then, using the calculated threshold value the given RGB colour apple image was segmented for extracting the region of interest. The resultant image was a binary image

Table 1: Image dataset

Apple type	Diseased/Sound	No. of samples Training Validation [540]	Test set[120]
Golden Delicious	Diseased	128	29
	sound	46	22
Red Delicious	Diseased	100	10
	sound	22	11
Granny smith	Diseased	122	14
	sound	60	13
Fuji	Diseased	10	8
	sound	22	11

3.3 Thresholding Algorithm-GSTM:

Let "Iset" is the apple colour image data set such that Iset=[$Im_1,Im_2,Im_3\ldots Im_N$], where 'N' is the total number of colour images in the apple image dataset. For each *Imi*, where $i{=}i..N$, Algorithm1 is applied to obtain the threshold value. Algorithm1 depicts the steps involved in the proposed algorithm GSTM.

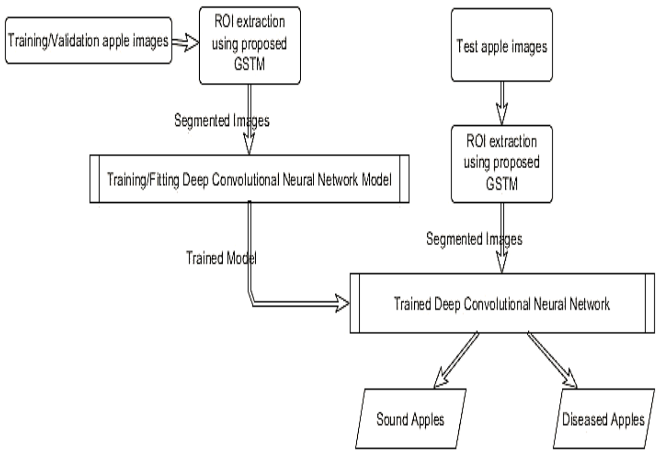


Figure 2: The proposed apple fruit sorting method

As the first step, the acquired apple colour image Im_i was converted to a grayscale image *Gim*. Then, the square root of each Gray pixel value was summed to calculate the threshold value. The square root of the individual values were calculated to transform each pixel value to generate an effective threshold value 'Thresh' that helps in segmenting the region of interest accurately. To transform the segmented image into a binary image, every pixel of the given apple image Im_i, i= i..N was taken and compared with the 'Thresh' value to decide the corresponding pixel value in the binary image Bim_i, where i=1..N using the rule mentioned below:

$$\text{Bim} = \begin{cases} 0 \; if \; Im(m,n) \; > \; T \\ 1 \; if \; Im(m,n) \; < \; T \end{cases}$$

Algorithm 1: Proposed GSTM

INPUT : Colour RGB image
OUTPUT : Segmented binary image Scan the input RGB colour image

ImGim ← Gray(*Im*)
Row ← No. of rows in *Gim*
Col←No. of columns in *Gim*
sum = 0
For *r = 1 to Row* **do**
  For *j=1 to Col* **do**
    $sum \leftarrow sum + \sqrt{Im_{i,j}}$
  end
end
Calculate threshold value 'Thresh' such that,

$$Thresh \leftarrow \frac{sum}{255}$$

Segment and generate a binary image using 'Thresh' value

$$\text{Bim} = \begin{cases} 0 \; if \; Im(m,n) \; > \; T \\ 1 \; if \; Im(m,n) \; < \; T \end{cases}$$

3.4 Proposed Deep Convolutional Neural Network model for classification:

Convolutional Neural Networks also called as Con vNet or CNN are a category of Neural Networks [**30**] that very effective in areas like image recognition and image classification. Convolutional neural networks have been successful in identifying diseases in apple fruits. Four major operations done by the Convolutional neural networks are i) Convolution ii) Non-Linearity iii) Pooling or Sub Sampling iv) Classification (Fully-Connected Layer). These operations act as the basic building blocks of every Convolutional Neural Network.
The overall training process of the proposed Deep Convolution Neural Network model has been summarized below:

Step 1: All filters and parameters/weights has to be initialized with some random values.

Step 2: The neural network takes a segmented training apple image as input, and goes through the forward propagation steps of convolution, ReLU/Sigmoid, and pooling layers along with forward propagation in the Fully-connected layer [as shown in Figure.3] and the output probabilities for each class are calculated.

Step 3: Calculate the total error using a loss function at the output layer such that; [**29**]

$$TotalError = \sum (targetprobability – outputprobability)^2$$

Step 4: Using Back propagation the gradients of the error was calculated for all the weights in the neural network

Step 5: The gradient descent was used to update all the filter values, weights, and parameter values to minimize the output error.

Step 6: The weights were adjusted according to their contribution to the total error and the same images were given as input again to check whether the output probabilities are closer to the target probabilities. .

Step 7: Repeat steps 2-7 with all other apple images in the training set

Thus, the model was optimized to classify the apple images from the training set. After the training process was complete, a test image set was given to the proposed Deep neural network model to check the classification accuracy rate.

3.4.1 Architecture: The proposed Deep CNN [Figure.3] sequential model had, three convolution blocks with a max pool layer in each of them and a fully connected deep neural network. Input layer to the convolution layer was made up of input data from the segmented apple images of size 150x150x2, where 150x150 specifies the width and height of the images and 2 represents the dimensions which represented the input features. Convolutional layer one was followed by a max-pooling layer with a 3x3 window and 32 filters. The second convolutional layer was followed by a max-pooling layer with 64 filters and a 3x3 window. And, the third convolutional layer was followed by a max-pooling layer with 128 filters and a 3x3 window.

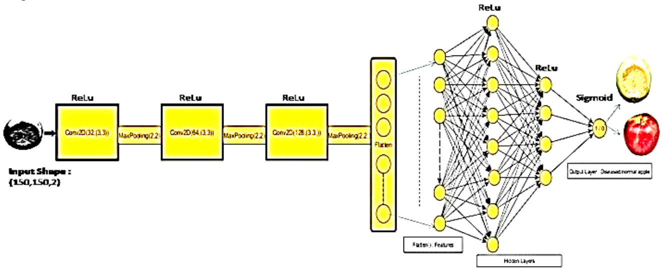


Figure 3: Proposed Deep Convolutional Neural Network architecture

In the fully connected layer of the convolution neural network, features obtained from the third convolutional max-pool layer were flattened and was given as input to the fully connected network designed for classification. The flatten layer was followed by two hidden layers one with 8 nodes and the other with 4 nodes. Both the hidden layers use the ReLU as their activation function. The output layer with one node was used the sigmoid activation function. The proposed model will output class probabilities for classifying the apple samples into two classes namely" Diseased/defective" and" Sound".

3.4.2 Parameters:

Activation functions:

The mathematical equations that determine the output of a neural network are called Activation function. This function is attached to every neuron in the neural network, and determines whether it should be activated or not, and this is done based on whether an input to each neuron is relevant for the model's prediction. Also, an activation function normalizes the output of each neuron to a range between 1 and 0 or between the range -1 and 1 [**27**].

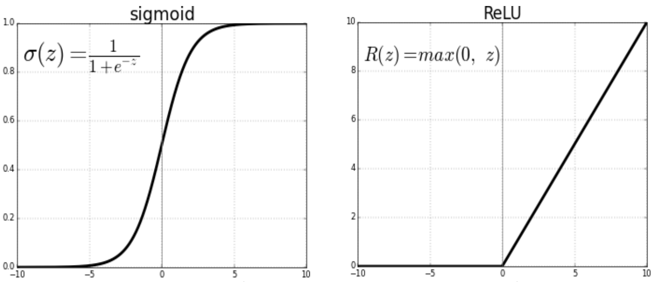


Figure 4: Graph of ReLu an Sigmoidal activation functions

Stochastic gradient descent with back propagation of errors is used to train deep neural networks so, an activation function is needed which will act like a linear function. The activation function must also provide a good sensitivity to the activation input sum and to avoid easy saturation.

1. **The rectified linear activation function or ReLU:** It is a linear function that will output the input directly if it is +ve, otherwise, it will output 0. A model that uses ReLu is easier to train and gives a better performance. ReLu activation function is f(x)=max(0,x).
2. **Sigmoid [28]:** The main advantage of using the sigmoid activation function is that it gives a smooth gradient, prevents jumping in output values, and the output values bound between 0 and 1. It normalizes the output of each neuron giving a clear prediction. The main reason for using the sigmoid function is that it exists between 0 and 1. It is especially used for the models that need to predict the probability as an output.

Figure.4 shows the graph of both ReLu and Sigmoid activation functions. Proposed DCNN was trained with the following parameters;
Number Of Learnable/trainable parameters:389233
Batch-size: 10
epochs:100
regularization:0.012

3.4.3 Segmentation assessment measures and methods:

Statistical measure: Wilcoxon signed-rank test has been a non-parametric test used to find whether two dependent samples are taken from populations having the same distribution. It assumes that the information is there in the signs of the differences and magnitudes between all paired observations. The Wilcoxon signed-rank test is similar to the t-test, when the data of a population does not follow a normal distribution. It can be used to compare any two groups. It is also known as the Rank-Sum test or the Signed-Rank test. This test helps to determine whether two or more pairs of sets are different from one another in a statistically significant manner.

Otsu's thresholding [15]: In Otsu's method goodness criteria for searching the threshold value is used. It uses within-class variance and between-class variance to select the threshold value that gives the optimum value. It uses the σ^2_{B} / σ^2_{T}, where σ^2_{B} is the between-class variance and σ^2_{T} is the total variance. Otsu algorithm is as follows;

1. Acquire the RGB color image, Convert it to grayscale.
2. Compute the histogram as well as probabilities of each intensity level
3. Initialize $\omega_i(0)$ and $\mu_i(0)$.
4. Repeat steps 4, 5 for all possible threshold's maximum intensity
5. Update ω_i and μ_i
6. Find $\sigma^2_{B}(T)$.
7. The final threshold value corresponds to the maximum $\sigma^2_{B}(T)$.

Kapur's thresholding [10]: Kapur's entropy function segments the grayscale image by maximizing the entropy value of the histogram. The entropy concept given by was used by Kapur but from a different point of view. Two probability distributions of the entire image is used in this method. One probability distribution is for the foreground object and the other is for the background object. The sum of the individual entropy is then maximized. Thus, resulting in an equal probability of Gray-levels in each region. Maximizing the sum of homogeneities in Gray-levels within the object and background will make the Gray-levels equally distributed in both the regions.

4. Results and Discussions

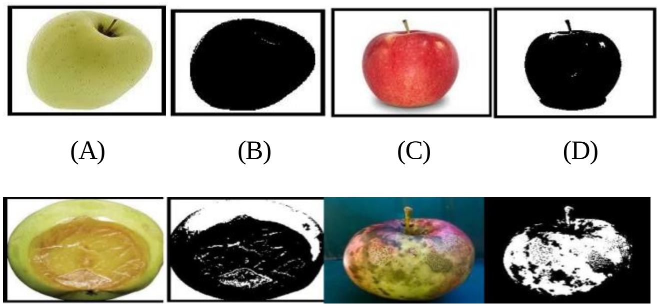
4.1 Performance of GSTM:

After executing Algorithm 1, a binary image with the region of interest [foreground image] was obtained as output. Figure. 5 shows the result of the GSTM algorithm with the original image, and the GSTM segmented output for four different sample test images. If visually observed, it can be noted that the diseased/defective part in the apple image has been segmented well. This accurate segmentation of the region of interest by GSTM, helped in a good classification process. As the proposed GSTM method is a grayscale thresholding method, the performance was compared with similar grayscale thresholding algorithms like Otsu's and Kapur's thresholding algorithms. The performance of the GSTM was compared with them, in terms of visual similarity and execution time. Figure.6 shows the visual segmentation accuracy of the region of interest (defective/diseased area) from the apple image, using the proposed GSTM method. Compared to Otsu's, and Kapur's segmented images. Also, the execution time of the proposed GSTM algorithm was less when compared to Otsu's and Kapur's algorithms. Table 2 shows the execution time of all the three algorithms in seconds. Execution time was calculated under the system environment with a configuration of Intel(R) Core(TM) 7200U CPU - 2.50 GHz to 2.70 GHz. From Table 2 it is evident that the execution time of the proposed GSTM is less when compared to Otsu and Kapur algorithms. The same has been statistically proved using Wilcoxon signed-rank test. To test the significance between GSTM and Otsu as well as GSTM and Kapur the following hypothesis was framed.

H0: There is no difference between the execution times [GSTM Vs Otsu and GSTM Vs Kapur]

H1: There is a difference between the execution times [GSTM Vs Otsu and GSTM Vs Kapur]

Figure.7 shows the output obtained using the Wilcoxon signed-rank test at a 5% level of significance using SPSS software. Both Figure. 7a and Figure. 7b hypothesis summary states that the median values between GSTM and Otsu as well as between GSTM and Kapur are equal to zero at 0.05 significance level, indicating a value that rejects the null hypothesis H0 for both the analysis.



(A) (B) (C) (D)

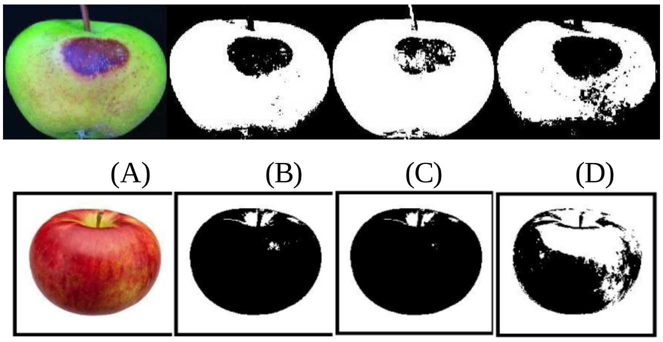
(E) (F) (G) (H)

Figure 5: A,C - Sound input images ; B,D - Proposed GSTM segmented sound apple images ; E,G - Diseased apple input images ; F,H - Proposed GSTM segmented diseased apple images

Thus, it was proved that there was a significant difference between the execution times of GSTM and Otsu as well as GSTM and Kapur. The same can be observed by manually inspecting Table 2.

4.2. Performance of Proposed Deep Convolutional Neural Network model:

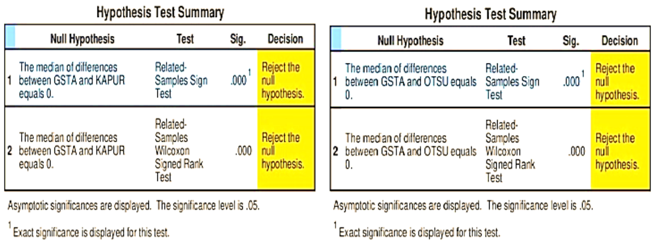
The proposed Deep CNN sequential model was trained using an apple dataset with 540 training/validation sample images. After applying data augmentations like shearing, rotating, rescaling, width shifting, height shifting, flipping, filling, and zooming a total 540 x 8 = 4320 training/validation apple dataset was obtained. These images were then segmented using the proposed GSTM algorithm/method and were given as input to the Deep Convolutional neural network. Training, and validation split was taken as 80% and 20%.



(A) (B) (C) (D)

(E) (F) (G) (H)

Figure 6: A,E- Original images from the dataset; B,F - Proposed GSTM segmented images ; C,G - Otsu's segmented images; D,H-Kapur's segmented images



Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of differences between GSTA and KAPUR equals 0.	Related-Samples Sign Test	.000[1]	Reject the null hypothesis.
2	The median of differences between GSTA and KAPUR equals 0.	Related-Samples Wilcoxon Signed Rank Test	.000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

[1] Exact significance is displayed for this test.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of differences between GSTA and OTSU equals 0.	Related-Samples Sign Test	.000[1]	Reject the null hypothesis.
2	The median of differences between GSTA and OTSU equals 0.	Related-Samples Wilcoxon Signed Rank Test	.000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

[1] Exact significance is displayed for this test.

(A) (B)

Figure 7: Wilcoxon signed-rank for execution times in seconds: A - GSTM Vs Kapur B - GSTM Vs Otsu

Table 2: Execution time in seconds – Proposed GSTM, Otsu, and Kapur

Apple Sample	Proposed GSTM	Otsu	Kapur
Sample1	0.02187	0.07077	0.02615
Sample2	0.02332	0.03814	0.02275
Sample3	0.01154	0.01657	0.01731
Sample4	0.01573	0.02317	0.02224
Sample5	0.01046	0.01697	0.01634
Sample6	0.01340	0.01690	0.01989
Sample7	0.01408	0.02724	0.01939
Sample8	0.02081	0.03157	0.02148
Sample9	0.00896	0.01558	0.01418
Sample10	0.01642	0.02544	0.01685

The result of the proposed Deep CNN architecture was designed and coded using python with the following input, output shapes and trainable/learnable parameter values. The input layer had no earnable parameters as it had only the input-segmented apple image data. The Convolution Layer 1 had inputs of sizes

Table 3: Classification Analysis [Proposed DCNN and Machine learning models]

Thresholding method	Classification Model	TP	FP	FN	TN	Recall	Precision	Accuracy	F-score
Proposed GSTM	**Proposed DCNN**	77	2	8	33	0.91	0.94	91.67	0.925
	Naïve Bayesian	71	4	18	27	0.80	0.87	81.67	0.833
	SVM	70	7	8	35	0.90	0.83	87.50	0.864
	KNN	54	24	6	36	0.90	0.60	75.00	0.720
	Logistic Regression	76	12	22	10	0.78	0.45	71.67	0.573
	Linear discriminant analysis	64	12	34	10	0.65	0.45	61.67	0.536
	Decision tree	52	10	44	14	0.54	0.58	55.00	0.562
Otsu	**Proposed DCNN**	54	4	28	34	0.66	0.89	73.33	0.759
	Naïve Bayesian	54	6	36	24	0.60	0.80	65.00	0.686
	SVM	72	6	36	6	0.67	0.50	65.00	0.571
	KNN	72	12	30	6	0.71	0.33	65.00	0.453
	Logistic Regression	36	10	64	10	0.36	0.50	38.33	0.419
	Linear discriminant analysis	32	12	64	12	0.33	0.50	36.67	0.400
	Decision tree	44	32	40	4	0.52	0.11	40.00	0.183
Kapur	**Proposed DCNN**	58	6	8	48	0.88	0.89	88.33	0.884
	SVM	60	2	44	14	0.58	0.88	61.67	0.695
	KNN	60	6	10	44	0.86	0.88	86.67	0.868
	Logistic Regression	32	14	60	14	0.35	0.50	38.33	0.410
	Linear discriminant analysis	30	14	60	16	0.33	0.53	38.33	0.410
	Decision tree	18	4	68	30	0.21	0.88	40.00	0.338

****DCNN : Deep convolutional Neural Network, SVM : Support Vector Machine, KNN:K-Nearest Neighbour, TP: True Positive, FP: False Positive, FN: False Negative, TN : True Negative**

150x150x2, with output shapes of [148,148,32] and 896 trainable parameters. Its max-pooling 2D layer had an output of shapes [74,74,32]. At Convolution Layer 2, the output shape was [72,72,64] with 18496 trainable parameters. Its max-pooling 2D layer had an output of shapes [36,36,64]. And, convolution layer 3 had an output shape [34,34,128] with 73856 trainable parameters. Its max pooling 2D layer had an output of shapes [17,17,128]. At the fully connected neural network, the flatten/input layer had the shape 36992. Hidden layer 1 had 8 neurons/nodes with activation function ReLu, regularization 0.012 and 295944 trainable parameters. Hidden layer 2 has 4 neurons/nodes with activation function ReLu, regularization 0.012 and 36 parameters. Finally, the Output layer with Sigmoid activation function had 5 trainable parameters. Thus, the total number of trainable/learnable parameters were 389233. And the total number of non-trainable parameters was zero. The model was trained and fitted with training batch size 10 and validation batch size as 3. The learning rate was observed and adjusted using the metrics like accuracy and loss for 100 epochs. Figure. 8a shows the learning accuracy rate graph and it can be observed that the accuracy rate gradually increases and moves upwards indicating a good accuracy rate. And Figure. 8b shows the loss rate graph indicating a steady low loss rate for 100 epochs. RMSPROP Optimization method was used for this model. And, the loss function used was binary-cross entropy. RMSPROP Optimization method was used for this model. And, the loss function used was binary-cross entropy. To evaluate/test the designed Deep convolutional neural network, 120 test images were used. The confusion matrix metrics like TP (Diseased/Defective apple images classified as diseased/defective correctly), TN (Sound apple images classified as sound apples), FP (diseased/defective apple images classifies as sound apple images), FN (Sound apple images classified as diseased/defective apple images) values were obtained. Recall/Sensitivity, Precision/Specificity, F-score, and testing. The accuracy rate was also calculated for GSTM segmented images, Otsu segmented images, and Kapur segmented images for comparison with the proposed DCNN and other classification models. The result of the same has been tabulated in Table.3. If observed accuracy in classifying the 120 test apple images using GSTM segmented images using the proposed DCNN was 91.67%, using Otsu segmented images was 73.33%, and Kapur segmented images was 88.33%.

Table 4: Classification Analysis [Proposed DCNN and Existing CNN models]

Thresholding method	Classification Model	TP	FP	FN	TN	Recall	Precision	Accuracy	F-score
Proposed GSTM	**Proposed DCNN**	77	2	8	33	0.91	0.94	91.67	0.925
	VGG (Visual Geometry Group)	71	4	18	27	0.80	0.87	81.67	0.833
	Dense Net	70	7	8	35	0.90	0.83	87.50	0.864

*****DCNN: Deep convolutional Neural Network, TP : True Positive, FP : False Positive, FN : False Negative, TN : True Negative***

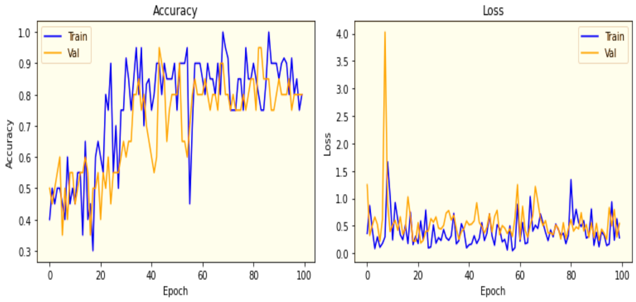
The GSTM method has got higher accuracy rate, F-score, Recall/sensitivity, and Precision/Specificity percentages than other classification models. Classification accuracy, Sensitivity, Precision, and F-score using test apple images were calculated using Eqn. (1), Eqn. (2), Eqn. (3), Eqn. (4), respectively.

$$\text{Accuracy} = \frac{TN+TP}{FN+FP+TP+TN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

$$\text{F-Score} = \frac{2 \times Precision \times Recall}{Precision+Recall} \quad (4)$$



(A) (B)

Figure 8: Training metric graphs for 100 epochs A) Accuracy rate B) loss rate

The GSTM method has got higher accuracy rate, F-score, Recall/sensitivity, and Precision/Specificity percentages than other classification models

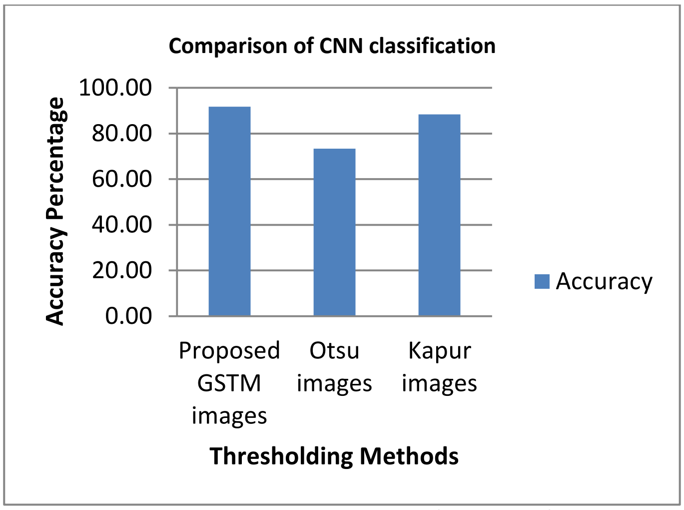


Figure 9: Bar chart comparing the accuracy rate of CNN classification using the proposed GSTM, Otsu and

Classification accuracy rate of the Deep convolutional neural network using GSTM, Otsu, and Kapur segmented images are represented using a bar chart in Figure. 9. It can be observed from the chart that the GSTM method with the Deep CNN model has given the maximum accuracy rate when compared to Otsu and Kapur methods. Also, Table. 4 shows the accuracy in classifying 120 test apple images using GSTM segmented images with respect to the proposed DCNN was 91.67%, VGG was 81.67%, and Dense Net was 87.50% respectively.

5. Conclusion and Future Scope

Image segmentation has been the primary step for any image analysis or classification process.

The reason for segmenting an image is to represent an image into some other form of image that helps in accurate analysis of any image. The proposed GSTM thresholding algorithm helped to extract the region of interest accurately. This accurate segmentation resulted in a good classification accuracy rate of sorting the acquired apple images as sound or defective/diseased ones using the proposed deep convolution neural network model. To conclude;

a) A new Gray-scale thresholding method GSTM has been proposed to segment the region of interest (defective/diseased parts) from an apple image.

b) The segmentation accuracy concerning visual accuracy and execution time was compared with similar Gray-scale thresholding algorithms proposed by Otsu and Kapur, and it was found that the visual, as well as segmentation accuracy of GSTM was good.

c) The execution speed of the proposed GSTM in segmenting the region of interest was higher than Otsu's and Kapur's methods.

d) A Deep convolutional neural network model was designed for the feature extraction and classification/sorting process.

e) The proposed GSTM with Deep CNN when compared to other existing Neural network gave the highest classification accuracy rate of 91.67%.

Thus, the proposed GSTM segmentation method and the Deep Convolution Neural network combination outperformed well in classifying/sorting apple fruit images and can be embedded in microchips of sorting machines with the required enhancements. In future a colour Gray-scale image segmentation algorithm can be developed and used. Also, a Deep CNN can be designed for obtaining more than 91.67% of accuracy.

Data Availability

This statement should describe how readers can access the data supporting the conclusions of the study and clearly outline the reasons why unavailable data cannot be released. Data was acquired using digital camera and from internet. So, it is not available in any repository.

Study Limitations

None.

Conflict of Interest

There is no conflict of interest [Single Author].

Funding Source

None. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors

Authors' Contributions

Author-1 researched literature and conceived the study, involved in protocol development, gaining ethical approval, patient recruitment, data analysis, and wrote the first draft of the manuscript. The authors also reviewed and edited the manuscript and approved the final version of the manuscript.

Acknowledgements

None

References

[1] Mohamed Abd Elaziz, Diego Oliva, Ahmed A Ewees, and Shengwu Xiong. "Multi-level thresholding-based grey scale image segmentation using multi objective multi-verse optimizer". *Expert Systems with Applications*, Vol.**125,** pp**.112– 129, 2019**.

[2] Nesar Mohammadi Baneh, Hossein Navid, Jalal Kafashan, Hatef Fouladi, and Ursula Gonzales-Barron. "Development and evaluation of a small-scale apple sorting machine equipped with a smart vision system". *Agri Engineering*, Vol.**5,** Issue**.1,** pp**.473–487, 2023**.

[3] Ariel E Baya, Monica G Larese, and Rafael Nam'ıas. *"*Clustering stability for automated color image segmentation". *Expert Systems with Applications*, Vol.**86,** pp**.258–273, 2017**.

[4] Santi Kumari Behera, Amiya Kumar Rath, Abhijeet Mahapatra, and Prabira Kumar Sethy. "Identification, classification & grading of fruits using machine learning & computer intelligence: a review." *Journal of Ambient Intelligence and Humanized Computing*, pp.**1–11, 2020.**

[5] Jaime Gallego and Montse Pardas. "Region based foreground segmentation ` combining color and depth sensors via logarithmic opinion pool decision". *Journal of visual communication and image representation*, Vol.**25,** Issue**.1,** pp**.184–194, 2014**.

[6] Farid Garcia-Lamont, Jair Cervantes, Asdrubal L ´ opez, and Lisbeth Rodriguez. "Segmentation of images by color features: A survey". *Neurocomputin*g, Vol.**292,** pp**.1–27, 2018.**

[7] A Gongal, M Karkee, and S Amatya. "Apple fruit size estimation using a 3d machine vision system." *Information Processing in Agriculture*, Vol.**5,** Issue**.4,** pp**.498– 503, 2018**.

[8] David Ireri, Eisa Belal, Cedric Okinda, Nelson Makange, and Changying Ji. "A computer vision system for defect discrimination and grading in tomatoes using machine learning and image processing. *Artificial Intelligence in Agriculture*, Vol.**2,** pp**.28–37, 2019**.

[9] Anis Ben Ishak. "A two-dimensional multilevel Thresholding method for image segmentation." *Applied SoftComputing*, Vol.**52,** pp**.306–322, 2017**.

[10] Jagat Narain Kapur, Prasanna K Sahoo, and Andrew KC Wong. *"*A new method for gray-level picture thresholding using the entropy of the histogram." *Computer vision, graphics, and image processing*, Vol.**29,** Issue**.3,** pp**.273–285, 1985**.

[11] SA Karthik and SS Manjunath. "An enhanced approach for spot segmentation of microarray images". *Procedia computer science*, Vol.**132,** pp**.226–235, 2018**.

[12] NS Lele. "Image classification using convolutional neural network". *International Journal of Scientific Research in Computer Science and Engineering*, Vol.**6,** Issue**.3,** pp**.22–26, 2018**.

[13] Leiming Liu, Ning Yang, Jinhui Lan, and Juanjuan Li. "Image segmentation based on gray stretch and threshold algorithm.*" Optik*, Vol.**126,** Issue.**6,** pp**.626–629, 2015**.

[14] K Suresh Manic, R Krishna Priya, and V Rajinikanth. "Image multithresholding based on kapur/tsallis entropy and firefly algorithm". *Indian Journal of Science and Technology,* Vol.**9,** Issue,**12,** pp.**89949, 2016**.

[15] Nobuyuki Otsu. "A threshold selection method from gray-level histograms". *IEEE transactions on systems, man, and cybernetics*, Vol.**9,** Issue.**1,** pp**.62–66, 1979**.

[16] Agus Pratondo, Chee-Kong Chui, and Sim Heng Ong. *"*Integrating machine learning with region-based active contour models in medical image segmentation.*" Journal of Visual Communication and Image Representation*, Vol.**43,** pp**.1–9, 2017.**

[17] Andreas Richtsfeld, Thomas Morwald, Johann Prankl Michael Zillich, and ¨ Markus Vincze. "Learning of perceptual grouping for object segmentation on rgb-d data". *Journal of visual communication and image representation*, Vol.**25,** Issue**.1,** pp**.64–73, 2014**.

[18] Chunshi Sha, Jian Hou, and Hongxia Cui. *"*A robust 2d otsu's thresholding method in image segmentation". *Journal of Visual Communication and Image Representation*, Vol.**41,** pp**.339–351, 2016**.

[19] Neha Sharma, Vibhor Jain, and Anju Mishra. *"*An analysis of convolutional neural networks for image classification". *Procedia computer science*, Vol.**132,** pp**. 377–384, 2018**.

[20] Thodeti Srikanth, P Pradeep Kumar, and Ashwin Kumar. *"*Color image segmentation using watershed algorithm". (*IJCSIT) International Journal of Computer Science and Information Technologies*, Vol.**2,** Issue**.5,** pp**.2332–2334, 2011.**

[21] Mingyuan Xin and Yong Wang. "Research on image classification model based on deep convolution neural network." *EURASIP Journal on Image and Video Processing*, pp**.40, 2019.**

[22] Aqing Yang, Huasheng Huang, Chan Zheng, Xunmu Zhu, Xiaofan Yang, Pengfei Chen, and Yueju Xue.*"*High-accuracy image segmentation for lactating sows using a fully convolutional network". *Biosystems Engineering,* Vol.**176,** pp**.36–47, 2018**.

[23] Junho Yim, Jeongwoo Ju, Heechul Jung, and Junmo Kim. *"*Image classification using convolutional neural networks with multi-stage feature". *In Robot Intelligence Technology and Application*s-3, pp.**587–594**, **2015**.

[24] Emny Harna Yossy, Jhonny Pranata, Tommy Wijaya, Heri Hermawan, and Widodo Budiharto."Mango fruit sortation system using neural network and computer vision". *Procedia computer science*, Vol.**116,** pp**.596–603, 2017**.

[25] Guoqiang Peter Zhang. Neural networks for classification: a survey." *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews),* Vol.**30,** Issue**.4,** pp**.451–462, 2000**.

[26] ZhiYong Zou, Tao Long, Qi Wang, Li Wang, Jie Chen, Bing Zou, and Lijia Xu. *"*Implementation of apple's automatic sorting system based on machine learning". *Food Science and Technology*, **42:e24922, 2022**.

[27] Poornima Devi. M, M. Sornam, "Contour based Character Segmentation and Nguyen-Widrow Weight Generation for Classification of Tamil Palm Leaf Script Characters - Machine Learning Approach*", International Journal of Computer Sciences and Engineering*, Vol.**7**, Issue.**5**, pp.**118-123, 2019.**

[28] PL.Chithra, M.Henila, "Fruits Classification Using Image Processing Techniques*", International Journal of Computer Sciences and Engineering,* Vol.**7**, Issue.**5**, pp**.131-135, 2019**.

[29] David Ademola Oyemade, Diseimokumor Favour Seregbe, "A Machine Learning Model for the Classification of Human Emotions," *International Journal of Computer Sciences and Engineering*, Vol.**12**, Issue.**4**, pp.**17-23, 2024**.

[30] Domo Omatsogunwa Ereku, Vincent I.E. Anireh, Onate Egerton Taylor, "Model For Email Spam Classification Using Hybrid Machine Learning Technique", *International Journal of Computer Sciences and Engineering*, Vol.**13**, Issue.**1**, pp.**24-32, 2025**.

AUTHOR PROFILE

Dr. M. Henila earned her MCA., M.Phil., B.Ed., P.G.D.C.S., and Ph.D. in Computer Science from Colleges affiliated to University of Madras. She is currently working as Associate Professor and Head, in the Department of BCA, Ethiraj College for Women, Chennai. She is a Reviewer for journal papers for Springer Nature Journals. She has published 5 research papers [with 85 citations and h-index 5] in reputed international journals including Soft Computing (Springer Nature), IET Image processing, IJCSE, BPAS, IJARCS (UGC referred, Scopus, SCI & Web of Science) and conference. Her main research work focuses on Image Processing, Neural Network, and Machine Learning. She has 24 years of teaching experience and 9 years of research experience.