Emerging Information Science and Technology
Vol. 5, No. 1, (2024), pp. 46-51

# Discrete Curvelet Transform Feature Extraction For Mangosteen Fruit Surface Damage Detection

Nafi Ananda Utama*, Wahyu Indah Triyani, Slamet Riyadi, Cahya Damarjati

*Universitas Muhammadiyah Yogyakarta, Jl. Brawijaya, Tamantirto, Kasihan, Bantul, Yogyakarta 55183, Indonesia*
**Corresponding Author:* nafi@umy.ac.id

***Abstract***

*Mangosteen (Garcinia mangostana L) is one of the commodities of Indonesian fruit and is used as an export primadona that became the basis of Indonesia to increase the currency of the country. The quality of the fruit can be seen from the surface, whether there is damage or not. The sorting that the farmers have been doing all this time is still using the conventional way, that is, with the sense of sight. This conventional method seems to be less effective because it takes a lot of energy, takes a long time, and there are different perceptions between farmers. To solve this problem, a method of surface quality extraction of mango fruit will be developed based on image processing. The initial stage of image processing is with the image size equation then the image is converted to grayscale mode, then a discrete curvelet transformation is performed. The next stage is the extraction of mean, energy, entropy, standard deviation, variance, sum, correlation, contrast, and homogeneity. The result of the subsequent feature extraction is used to enter a value at the classification stage. From some of these extractions it will be known which extraction has the highest accuracy value. The method of classification used is Linear Discriminant Analysis (LDA) with the method of K-Fold Cross Validation which in this study is divided into 4-fold cross validation. After testing on 120 images, the highest value of accuracy is with extraction of standard characteristics deviation of 91.7% and variance of 88.4%.*

***Keywords**: Discrete curvelet transformation; Feature extraction; Image processing*

## 1. Introduction

Mangosteen (*Garcinia mangostana* L) is a commodity of Indonesian fruit and is used as an export primadona so it becomes an Indonesians trust to increase the currency of the country. The importance of mangrove from the export side can be seen from the constantly increasing volume of mango exports. Therefore, it requires high quality mango fruit to be able to export and compete with other countries. This research focuses only on extraction of the surface quality characteristics of mango fruit. Inspection of the surface quality of the mango fruit is still using conventional methods, i.e. by observing manually with the sense of vision. This traditional method feels less effective because it takes a lot of energy, takes a long time, the subjectivity factor and the fatigue factor.

Matlab is a programming language developed by MathWorks and dedicated to numerical computing, visualization, and programming. [1]. Using matlab allows users to analyze data, develop algorithms, and create models and applications. Matlab provides a wide range of tools that will facilitate program writing so you can focus more on research outcomes and innovations. A curvelet transformation is a directed multi-scale transformation that allows the almost optimal nonadaptive sparse appearance of an object that has many edges. [2]. A multiscale method that cuts data into sets of different frequency levels, so that each level can be analyzed and classified. Analysis of curvelet transformation results can be done using Linear Discriminant Analysis (LDA). LDA is a method used in statistics to classify one or more objects of a group based on a particular characteristic.

Character extraction is a phase of comparing one image with the other. The purpose of the extraction is to obtain the pattern of each characteristic of the fruit so that it can be found the similarity of characteristics between one fruit and the other fruit. Character extraction is generally done based on colour, shape, and texture [3]. In this study used character extraction based on texture in the fruit.

## 2. Related Works

Image processing research using curvelet transformation methods has been done a lot before with a variety of classifications used, one of which has been carried out by Khoje, Bodhe, & Adsul, (2013) in his research “Automated Skin Defect Identification System for Fruit Grading Based on Discrete Curvelet Transform” aimed to develop a methodology for evaluating the quality of fruit objectively using texture analysis based on the curve transform.

As a multi-resolution approach, the curvelet has the ability to examine the fruit surface with low and high resolution to extract global and local details of the fruit's surface. The study used four characteristic extractions: energy, entropy, mean, and standard deviation used to characterize the fruit's surface texture. The features used to clarify the surface of the fruit are the Support Vector Machine (SVM) and Probabilistic Neural Networks (PNN) and the classification performance is tested for two categories of grading the fruit: healthy and defective. The results show that the best SVM classification is obtained with an accuracy of 96%. The study concluded that with curvelet transformation gives promising insights to estimate damage to the skin of the fruit.

## 3. Method

This research method includes manual classification, design and preparation program and testing.

### 3.1 Manual Classification

Manual classification is done by grouping images based on the level of damage on the surface of mango fruit. With manual classification this can be used as a measure in calculating the degree of accuracy of programs through comparison of manual classifications and using programs. The final result that will be used is a manggis surface image that has been grouped according to the corruption and non-resolution of the manggis image.

### 3.2 Design and Preparation Program

At this stage, a program has been designed to detect damage to the surface of mango fruit using the method of extraction characteristics curvelet and using the LDA classification method, this study uses some characteristics extraction results as LDA input i.e. mean, energy, entropy, standard deviation, variance, sum, correlation, contrast, homogeneity. From several extractions of such attributes will be determined which attribute extraction has the accuracy with the best value.

a. Resize Image

The image resize is intended to harmonize the input image before processing using the curvelet transformation method. At the stage of imaging the image data the surface of the mango fruit has different sizes, then the resize function is used to harmonise the image size to 512x512 pixels.

b. Konversi Citra RGB ke grayscale

Conversion of RGB images to grayscale images is aimed at simplifying input images thus reducing processing time. The first process of image conversion is reading the RGB image data folder, then calculating the number of image inputs to be converted to gray scale mode, followed by the process of converting images, after which grayscaled image data is saved.

c. Curvelet Transformation

At the transformation stage of this curvelet it starts from the reading of the image data resulting from the grayscale image conversion then the input data that has been stored in the variable is inserted into the transforming curve function for the further process of transformation, after transformation process is completed, the values of the curve coefficient as well as curve values will be obtained.

d. Properties Extraction

The characteristic extraction is calculated from the curvelet coefficient obtained from the process of transforming the discrete curvelets and will produce the required values, namely the values of mean, energy, entropy, standard deviation, variance, sum, correlation, contrast, homogeneity. The result of the extraction of the characteristics will be used to determine the surface of the fruit that is damaged and not damaged. This selected characteristic extraction result is used to give an assessment of the quality of the fruits of mango. The value of the extraction of the features will be analyzed to be entered as an input for the classification process using the LDA method.

e. Classification

The method of classification used in this study is the Linear Discriminant Analysis (LDA) method using two stages, the training image and the test image. The classification is divided into two classes, the broken class and the unbroken class. This method of classification takes input from the values of the output of characteristic extraction.

### 3.3 Testing

The purpose of the surface damage detection program is to determine the degree of accuracy of the program based on the data from the extraction results. The method used in this test phase is to compare the program detection results with the results of the detection manually.

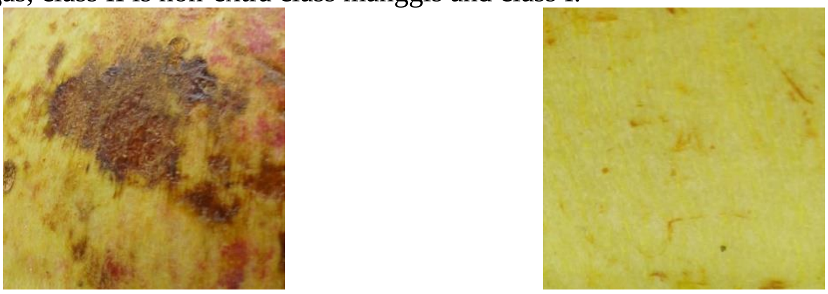
## 4. Result and Discussion

### 4.1 System Operating Principles

The operating principle of the system begins with the reading of corrupt and non- corrupt images that have been selected and collected in a particular folder. The set of images is then processed by pre-processing the image and extracting the characteristics that will result in a certain value. The image set is divided into training images and test images with the data set specified from 4-Fold Cross Validation. Linear Discriminant Analysis is done to do the training process. At the test stage, the new data is mapped onto the training area model so that the results will be located in a specific area and can be classified.

### 4.2 Data Collection Results

The total image data is 120 which is a combination of the training image and the test image, of the 120 images consisting of 60 corrupt images and 60 intact images. This image has varied qualities: Extra class, Class I, and Class II. Extra class is super quality manggis, class I is good quality manggas, class II is non-extra class manggis and class I.



(a) Damaged Mangosteen Image

(b) Undamaged Mangosteen Image

Figure **1.** Examples of Damaged and Undamaged Mangosteen Images

### 4.3 Program Design Result

#### 4.3.1 Resize Image

The data was taken using a digital camera with the same distance of about 50 cm and the same resolution. In data capture, one photo contains four manggis in the same class. This is done to shorten data capturing time. Then resize each mango with the image size changed to 512x512 pixels.

(a) Mangosteen image before croping

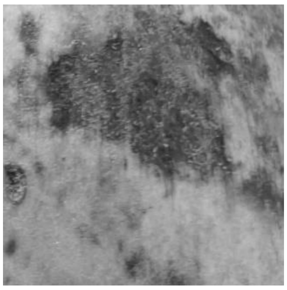
(b) (b) Mangosteen image after cropping and resizing

Figure **2**. Image Resize Process

### 4.3.2 Convert RGB to Grayscale

After the image is resized, the image color mode is changed from RGB (Red Green Blue) to grayscale color mode with the aim of simplifying the input image so as to reduce processing time.

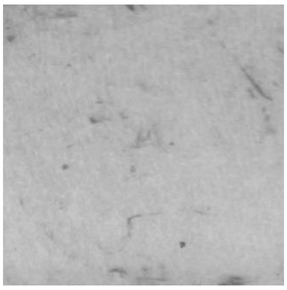
(a) RGB Image



(b) Grayscale Image

Figure **3**. Conversion Result of Damaged Mangosteen Image to Grayscale Image

(c) RGB Image

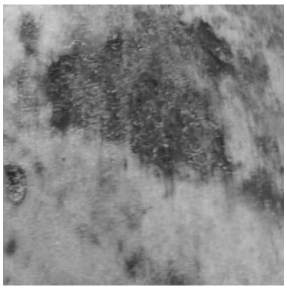


(d) Grayscale Image

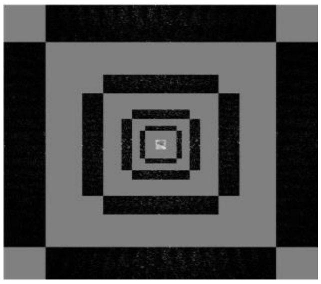
Figure **4**. Conversion Result of Undamaged Mangosteen Image to Grayscale Image

### 4.3.3 Curvelet Transformation

The image in grayscale format is then transformed at a sub-image frequency where the components are generated by decomposing the grayscale image so as to produce curves of matrix values from the image. From the curvelet transformation that has been done, the values of the curvelet transformation results will be obtained. The value of the curvelet transformation contains information about the decomposition of the compressed transformed image.

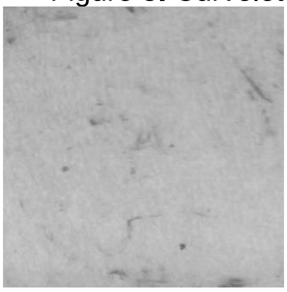


(a) Damaged Grayscale Image

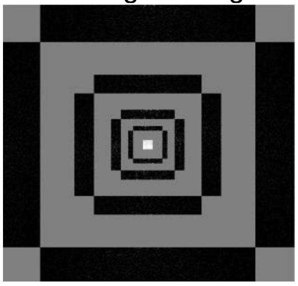


(b) Curvelet Transform Result

Figure **5.** Curvelet Transform Process On Damaged Image



(c) Undamaged Grayscale Image



(d) Curvelet Transform Result

Figure **6.** Curvelet Transform Process On Undamaged Image

Figure 5 (b) is the discrete curvelet transformed image of a defective image, the defective image is decomposed using the curvelet transform and produces curvelet coefficients. The curvelet coefficient is composed of curves that are formed into one. The collection of curves produces a value that will be used to determine the value of feature extraction from the damaged image. Figure 6 (b) is an image of the results of discrete curvelet transformation on an undamaged image, the undamaged image is decomposed using curvelet transformation and produces curvelet coefficients. From the results of curvelet transformation, the damaged and undamaged images look physically the same and it is difficult to distinguish whether the image is damaged or not because it is in the form of slices of curves that are visibly the same. The value of the curvelet transformation result is then used to form the feature extraction value of each image.

#### 4.3.4 Feature Extraction

Characteristic extraction is a process of taking the unique characteristics of the mangosteen surface image into a certain value. Characteristic extraction is calculated from the curvelet coefficient that has been obtained from the curvelet transformation process and will produce the required value. The results of this feature extraction will be used as an analysis material to determine the damaged or undamaged fruit surface.

#### 4.3.5 Classification and Validation

In this research, the classification uses Linear Discriminant Analysis (LDA) and uses 4-Fold Cross Validation. The data used in this research is 120 images. The data is used in two stages in the classification, namely the training stage and the testing stage. The 120 data consisted of 60 damaged images and 60 undamaged images. Then the data is divided into 4 folds with each fold containing 30 images from a combination of 15 damaged images and 15 undamaged images. This class division aims to facilitate the validation process so as to produce maximum accuracy.

After getting the data from the feature extraction, the feature extraction data is stored in the database with the file.mat extension. After the data is stored, the feature extraction data is trained using the LDATrain function to generate training data, then the training data is used to test the test image.

### 4.4 Testing Result

The results were tested on 120 images consisting of 60 damaged images and 60 undamaged images divided into 4-fold using LDA classification and 4-Fold Cross Validation validation. Each fold is tested so as to produce an overall accuracy percentage of the 4 folds. Overall, it can be concluded that testing on 120 test images divided into 4-fold cross validation results in varying accuracy of each feature extraction. In the mean feature extraction, the average obtained from fold¬-1 to fold-4 is 80%, energy produces an average of 77.5%, entropy produces an average of 77.5%, standard deviation produces an average of 91.7%, variance produces an average of 88.4%, sum produces an average of 80%, correlation produces an average of 50%, contrast produces an average of 71.7%, and homogeneity produces an average of 80.8%. Of the nine feature extractions, the standard deviation and variance feature extractions have the highest accuracy value, so they can be used properly to create a mangosteen fruit surface quality detection system.

## 5. Conclusion

Characteristic extraction of standard deviation and variance has a higher accuracy value with an accuracy value of 91.7% and 88.4%. Characteristic extraction in the form of standard deviation and variance values can be used well to build a mangosteen fruit surface quality detection system because it is proven to be able to distinguish damaged and undamaged images. Reading errors occur due to lighting intensity factors at the time of data collection that affect image feature extraction.

## Acknowledgments

The authors acknowledge the support of Universitas Muhammadiyah Yogyakarta in Conducting this research on fruit classification using digital image processing techniques.

## References

[1] Kusuma, H. A. (2016). Introduction to Matlab Program Window. Retrieved 07 17, 2017, from Hollanda Cocobear: https://hollandakusuma.wordpress.com/2016/06/01/pengenalan-jendela-program-matlab/
[2] P, M. N. (2014). Utilization of Artificial Sparse Imagery for Sharpening Remote Sensing Imagery. Remote Sensing Data and Technology Center, LAPAN.
[3] Kusumaningsih, I. (2009). Color, Shape and Texture Feature Extraction for Animal Image Retrieval. Bogor: Department of Computer Science, Faculty of Mathematics and Natural Sciences, Bogor Agricultural University.
[4] Al-Marzouqi, H., & AlRegib, G. (2012). Curvelet Transform with Adaptive Tiling. *School of Electrical and Computer Engineering*.
[5] Jianwei, M., & Plonka, G. (2010). The Curvelet Transform : A review of recent application. *IEEE SIGNAL PROCESSING MAGAZINE*.
[6] Kazemi, F. M., Samadi, S., Poorreza, H. R., & Akbarzadeh-T, M.-R. (2007). Vehicle Recognition Using Curvelet Transform and SVM. *International Conference on Information Technology (ITNG'07), IEEE*.
[7] Khoje, S. A., Bodhe, S. K., & Adsul, A. (2013). Automated Skin Defect Identification System for Fruit Grading Based on Discrete Curvelet Transform. *International Journal of Engineering and Technology (IJET)*.
[8] Nugroho, A. (2015). A prototype object following system for quadcopter using digital image processing. Yogyakarta: Electronics and Instrumentation Study Program, Department of Computer Science and Electronics, Faculty of Mathematics and Natural Sciences, Gadjah Mada University.
[9] Prabha, S., & Sasikala, M. (2013). Texture Classification Using Curvelet Transform. *International Journal of Advancements in Research & Technology*.
[10] Priyambodo, K. B. (2007). Study of Tranducer Pressure Force on Ultrasonic Wave Velocity in Mangosteen Fruit. Bogor: Department of Agricultural Engineering, Faculty of Agricultural Technology, Bogor Agricultural University.
[11] Rahulkar, A., Jadhav, D., & Holambe, R. (2010). Fast Discrete Curvelet Transform Based Anisotropic Feature Extraction for Iris Recognition. *ICTACT Journal On Image And Video Processing*.
[12] Simanjuntak, F. (2009). Digital Image Processing. http://digilib.ittelkom.ac.id/index.php?option=com_content&view=article&id=73:pengolahan-citra-digital&catid=15:pemrosesansinyal&Itemid=14.
[13] Taufiq, I. H. (2010). Analysis and Implementation of Content Based Image Retrieval Using Curvelet Transfrom Method. Telkom University Bandung.
[14] Taufiq, I. H., Adiwijawa, & Wirayuda, T. A. (2010). Analysis And Implementation Of Content Based Image Retrieval. Informatics Engineering, Faculty of Informatics Engineering, Telkom University.
[15] Thielicke, W., & Stamhuis, E. (2014). PIVlab – Towards User-friendly, Affordable and Accurate Digital Particle Image Velocimetry in MATLAB. *Software Sustainability Institute Journal of Open Research Software*.
[16] *Weisstein, Eric W*. (1999). Retrieved 07 31, 2017, from WolframMathWorld: http://mathworld.wolfram.com/Sum.html