Computer Engineering and Applications Vol. 9, No. 3, October 2020

Utilization of Support Vector Machine and Speeded up Robust Features Extraction in Classifying Fruit Imagery

Muhathir Muhathir, Wahyu Hidayah, Dian Ifantiska

Department of Computer Engineering,Faculty of Computer Science, Universitas Medan Area
**muhathir@staff.uma.ac.id*

ABSTRACT

Indonesia's various types of fruits can be met by the community. Many fruits that contain a source of vitamins are very beneficial to the body, or as an economic source for farmers. It's no wonder that many experts submit discoveries to increase the amount of productivity or just want to experiment with intelligent systems. Intelligent systems are specially designed machines in certain areas to adjust the capabilities made by the creators. This article provides the latest texture classification technique called Speeded up Robust Features (SURF) with the SVM (Support Vector Machine) method. In this concept, the representation of the image data is done by capturing features in the form of keys. SURF uses the determinant of the Hessian matrix to reach the point of interest in which descriptions and classifications are performed. This method delivers superior performance compared to existing methods in terms of processing time, accuracy, and durability. The results showed that the fruit classification by using the extraction of Speeded up Robust Features (SURF) feature and SVM (Support Vector Machine) Classification method is quite maximal and accurate. Result of 3 kinds of classification with SVM kernel function, SVM Gaussian with 72% accuracy, Polynomial SVM with 69.75% accuracy, and Linear SVM with 70.25% accuracy.

Keywords: Fruit, SVM, SURF.

1. INTRODUCTION

Indonesia is a country rich in tropical fruits of choice. All local fruits are available all year round in the market but some are seasonal [1]. Texture classification is a way of grouping similar things together according to common characteristics that allow us to obtain information about the image. This information can be obtained by extracting the image features. With the help of features, we can accurately describe large amounts of data. To acquire the features of the image, textures play an important role because it illustrates the appearance of an object. We can't classify images that overlap separately, but they can be character by using their textures because each image has a specific texture that determines the characteristics of the image. The classification of textures has a wide range of applications in rock classification, the introduction of wood species, facial recognition, geographic landscape segmentation, object detection, and many other image processing applications [2].

Advances in digital technology have created a large collection of digital images that require efficient and intelligent texture classification techniques. The classification of textures is a process for dividing texture features into a texture class. The first phase of the classification is a feature extraction process where the features are extracted from the image using texture. In the second stage, the feature is converted to texture classes by using a classifier. The processing of the SURF method is much faster than the SIFT (Scale Invariant Feature Transform) method. SURF is an improved version of SIFT [3].

SURF is a texture detector and descriptor method that has applications such as object recognition, image registration, classification, 3D scene reconstruction, and tracking objects. The SURF technique handles blurry images and rotations efficiently. The first process in the SURF classification is the calculation of an integral image used to increase performance speed. Furthermore, to extract the features of an image, an interesting place is achieved using the Hessian BLOB detector and the description obtained for each destination. The descriptor is based on the number of responses to the Haar wavelet around the destination. The last stage is the matching descriptors obtained from different images [4]. SURF solves this problem and accelerates calculations with the help of integral images. The Hessian matrix determinant is used in the SURF to detect the location of the flower point and provide stable performance compared to other detectors such as the Harris detector. The SURF feature was extracted in terms of keys found by describing the pixel intensity distribution [5].

In this paper, the vector support machine (SVM), which is a relatively new machine learning technique, was introduced and proposed for the fruit classification system. SVM is a simple and very powerful method. This is because of the optimization and generalization solution [6]. The proposed effectiveness of the system was evaluated experimentally and compared to existing methods, neural networks. The results show that the proposed technique has resulted in a better level of accuracy compared to existing methods. Also, the training time is less than that of the existing methods [7].

Also, SURF is very sensitive to changes in perspective (viewpoint) imagery, for example, changes such as left or right rotation or often referred to as planar changes [8]. SURF is the development of the Scale Invariant Feature Transform (SIFT) feature extraction algorithm [7]. Feature extraction is a method for generating traits from an image. SURF can detect local features of an image precisely, rapidly, and invariant on scale changes, rotation, and transformation [9]. SURF uses an integral mix of imagery and blob detection based on the determinant of the Hessian matrix on its feature-extraction algorithm. The extraction phase begins by specifying an interesting point that contains a lot of information and is stable to interference with imagery, where the interest point value is obtained using blob detection. In the interest point stage, there are two stages: scale-space representation and feature description [10].

SURF (Speeded Up Robust Features) is threaded to detect the features of an image quickly and reliably. The SURF method is inspired by the SIFT (Scale-Invariant Feature Transform) method, which makes this method inspired by the SIFT method is when it comes to the Scale-space representation phase. The SURF method uses two combined methods: an integral image method and also a blob detection based on the determinant of the Hessian matrix [11].

In this study, we analyzed using the SVM method in the classification of the fruit with the extraction of the SURF features, with some samples of existing fruit, and also the feature SURF that is quite capable of working faster and reliable [12]. It is hoped that this study can provide new technology concepts that facilitate the farmers in classifying the introduction of imagery patterns in the fruit. Therefore, this

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

research implements the SVM (Support Vector Machine) method with the extraction combination of SURF features (Speeded Up Robust Features).

2. MATERIAL AND METHODE

2.1 FRUIT

The diversity of fruit genetic sources that grow scattered in various regions in Indonesia is a priceless wealth. Fruits grown in Indonesia and planted by Indonesian people are named local fruit. A wide variety of fruits has nutrients, vitamins, minerals, and fibers that are very necessary to be consumed every day [1]. Therefore, researchers are interested to classify some fruit such as apples, avocado, apricot, and banana.

a) Apple

Apples (Pyrus malus L) is the fruit that produced apple plants. Apple fruit is usually red, green, or yellow, according to the type of apples. The fruit skin is slightly mushy, the flesh is hard. Various kinds of apples that are not difficult to find in the market include Red Delicious, Apple Fuji Jingle, Apple Golden Delicious, Gala, Granny Smith, Manalagi, and Malang [13][14].

Apples are generally consumed as fresh fruit. The important component of Apple fruit is pectin, which is about 24%. The content of pectin on the Apple fruit is found around the seeds, under the skin and liver. The pectin will form a gel when added sugar in a certain pH range. Pekin plays an important role in the making of juice (fruit juices), jellies, jams, and Dodol. Apple fruit (Malus sylvestris Mill) In addition to having substance pectin compounds also contain other nutrients [15].

b) Avocado

Avocado is the only fruit that is rich in fats. It is more than twice the fat content of durian. Nevertheless, avocado fats include healthy fats, as they are dominated by monounsaturated fatty acids that are powerful antioxidants [1]. Avocado is considered to be one of the main tropical fruits, since it contains fat-soluble vitamins that are less common in other fruits, in addition to high levels of protein, potassium, and unsaturated fatty acids. Avocado porridge contains varied oil content and is widely used in the pharmaceutical and cosmetic industries, and in commercial oil production is similar to olive oil. This fruit has recognized the facility for health benefits, mainly because of the compounds present in the lipidic fraction, such as Omega fatty acids, phytosterols, tocopherols, and Squalene [17].

c) Apricot

According to the article [18], apricot fruit is rich in nutrients and contains antioxidants (carotenoids) which is great for supplements because of its nature that inhibits oxidation and protects cells from the harmful effects of free radicals. Substances found in dried apricots are proteins, carbohydrates, fruit sugars, dietary fibers, fats, vitamin A, vitamin C, vitamin B9, vitamin K, calciumpotassium, iron, and lycopene. Apricot fruit can be obtained from May to August. Apricot fruit can be consumed as fresh fruit, dried fruit, in juice form, or made jam. Ripe apricot fruit has the highest antioxidant content [18].

d) Banana

Banana is an agricultural commodity and is one of the superior fruits of Indonesia that is very popular because of its many benefits [19]. The banana fruit contains quite high nutrition, low cholesterol as well as vitamin B6, and vitamin C high. The largest nutrient in ripe banana fruit is potassium amounting to 373 milligrams per 100 grams of bananas, vitamin A 250-335 grams per 100 grams of bananas, and clenbuterol at 125 milligrams per 100 grams of bananas. Bananas are also a source of carbohydrates, Vitamin A and C, as well as minerals. The biggest carbohydrate component in the banana fruit is the starch in the flesh and will be converted into sucrose, glucose, and fructose when ripe bananas [20].

2.2 SPEED-UP ROBUST FEATURE (SURF)

The air image retrieval process must be repeated to obtain an image with a wide coverage area [21,22]. The result of the aerial image retrieval is a separate image between one image and another. Images that have been captured can be linked using the SURF (Speed Up Robust Features) algorithm [2]. This algorithm works by specifying a key point of an image whose value is fixed when experiencing scale changes, rotation, blurring, lighting, and shape changes [24].

Although the SURF algorithm has a good capability in key point detection, most of the results of such detection are not necessarily considered as interesting key points to the imagery that is tested for that reason, the key point will only increase the compute time on the temporary SURF algorithms in the process, at least only 4 pairs of matching key point between imagery to make the image layer process possible. These unappealing key points need to be eliminated [22].

The final stage of the SURF algorithm is classification. Here, we've used a multi-level SVM classifier that converts image features into texture classes. It compares the test set with the training set and provides the correct classification of objects. SVM divides two classes using hyperplane [8]. Speed up Robust features (SURF) used to extract texture information from images [10] SURF approaches the Hessian matrix of two localized descriptor features. It is extracted in an interesting region in an image [9].

2.3 SUPPORT VECTOR MACHINE (SVM)

Speed up Robust features (SURF) used to extract texture information from images [10] SURF approaches the Hessian matrix of two localized descriptor features. It is extracted in an interesting region in an image [9]. SVM has been used in various applications built upon categorization such as classifying points into separate planes, text categorization, and pattern recognition. In today's world with enormous data there is a need for multigrade classification. This is especially for the target category greater than two. The article [7] proposed the ECOC framework to transform Multiclass into several binary issues. SVM when combined with ECOC enriches the system of failure when completing a multi-screen classification [5].

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

ECOC reduces the multiclass issue to a classifications binary group. It consists of two schemes. "Coding Schema: Design encoding presents ways through which multiclass problems are reduced to a group of binary class issues. That explains the classes that are trained by the binary learner on the ". "Re-Encoding scheme: it presents a way to combine the results obtained from binary learners. A detailed explanation is available with Support Vector Machine (SVM). The steps are as follows: Step 1: Load the fruit data set. Step 2: Extraction of color, shape, and texture features. Step 3: Specify the Predictor record name and data response name. Step 4: Create an SVM template and specify the order of the predictor. Step 5: Train the ECOC classifier using the SVM binary learner with the coding design and specify the order of the classes. Step 6: Cross-validation classification. The ECOC classification is used KFold. Langkah7: Predict the accuracy of classification for Test Data [7].

In this algorithm to move the best value/pattern, depending on the data, the high value/pattern can minimize the noise effect at the time of classification, but this algorithm will cause the constraints between each classification to be less good. The optimization of the parameters can be used when the value/pattern has good quality. In a value/pattern that is arguably good, can be selected using the optimization of parameters, which is by using cross-validation [7].

2.4 DATASET

The datasets used in this study were taken from the Fruits 360 dataset or were accessible to https://www.kaggle.com/moltean/fruits. Image of a JPEG formatted with a resolution of 100x100. Samples for each of the fruit used amounted to 556 samples.

2.5 RESEARCH SCHEME

This research step is modeled into the diagram plot as in Figure 1.

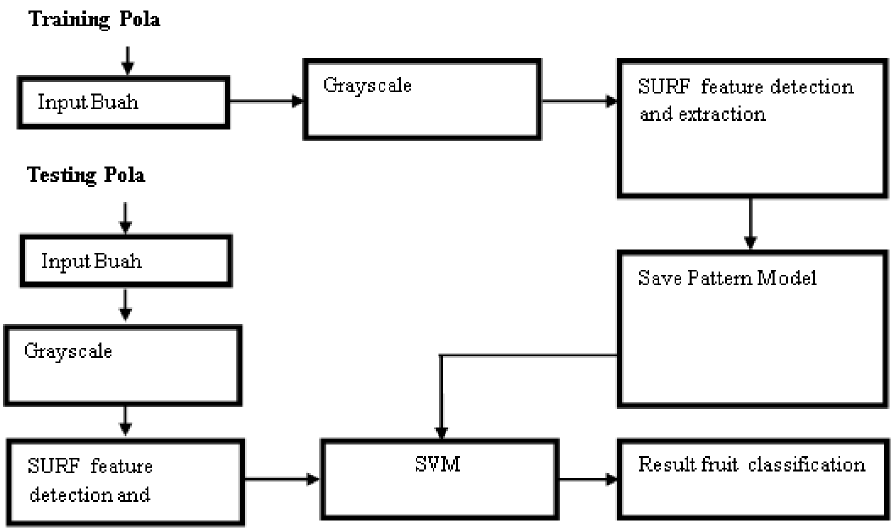


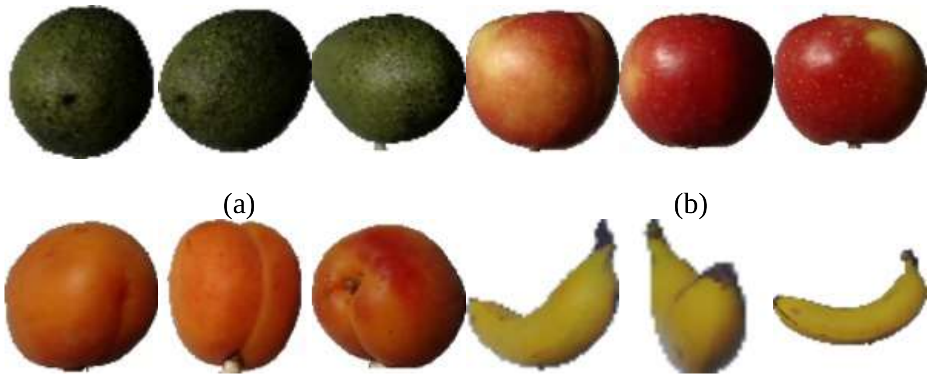
FIGURE 1. Research Step [25, 26]

Figure 1 shows the research steps that will be done by two processes, the first training process, the training process is a process in the data extraction (starting with minimizing the color space on the image of the three color spaces R, G, B into one color space that is grayscale as well as extracting by utilizing the SURF feature extractions), as well as extracellular results, are used in the second testing process is the process of matching the model of the pattern that has been in training by utilizing the SVM method as classification.

3. RESULTS AND DISCUSSION

3.1 FRUIT SAMPLE

The fruit samples used in this study were taken from the Fruits 360 dataset. Figure 2 Enclose some fruit samples.



(a) (b)

(c) (d)

FIGURE 2. Fruit Sample (a) Avocado, (b) Apple, (c) Apricot, and (d) Banana

3.2 SURF FEATURE DETECTION

The SURF detection feature is characterized by a circle mapping on the image, Figure 3 shows the results of the features of the SURF in avocado, apples, apricot, and banana.

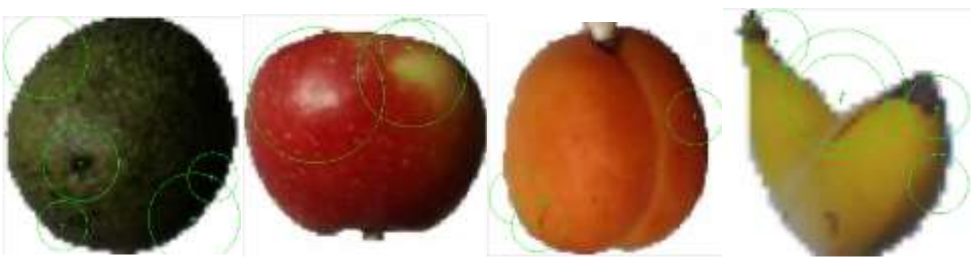


FIGURE 3. Result SURF Detection

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

3.3 CLASSIFICATION RESULT

Fruit classification test Result using the SVM method with SURF extraction combination. Testing SVM by utilizing the Gaussian, Polynomial, and Linear kernel functions. Gaussian SVM Classification results are presented in Table 1.

TABLE 1.
SVM test result using Gaussian kernel function

	Avocado	Apple	Apricot	Banana
Avocado	**49%**	8%	41%	2%
Apple	4%	**53%**	35%	8%
Apricot	0%	2%	**97%**	1%
Banana	8%	2%	1%	**89%**

Table 1 shows the SVM classification results using the Gaussian kernel function and the extraction of the SURF features, in the avocados the percentage is recognized at 49%, in the Apple the percentage is recognized at 53%, the fruit of the apricot percentage is recognized at 97% and in the banana fruit, the percentage is recognized by 89%. With an average total accuracy of 72%. Figure 4 illustrates the SVM classification results using Gaussian kernel function and SURF feature extraction.

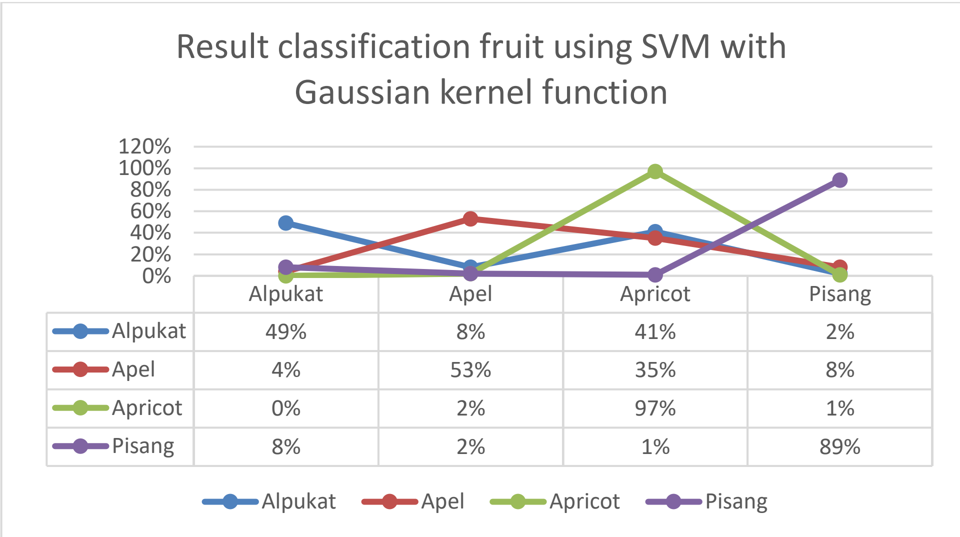


FIGURE 4. Result classification fruit using SVM with Gaussian kernel function

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

TABLE 2.
SVM test result using Gaussian kernel function

	Avocado	Apple	Apricot	Banana
Avocado	**47%**	5%	43%	5%
Apple	4%	**52%**	38%	5%
Apricot	1%	2%	**96%**	1%
Banana	5%	3%	7%	**84%**

Table 2 shows the SVM classification results using the polynomial kernel function and the extraction of the SURF features, in the avocados the percentage is recognized at 47%, in the Apple the percentage is recognized at 52%, the fruit of the apricot percentage is recognized at 96% and in the banana fruit, the percentage is recognized by 84%. With an average total accuracy of 69.75%. Figure 5 illustrates the SVM classification results using polynomial kernel function and SURF feature extraction.

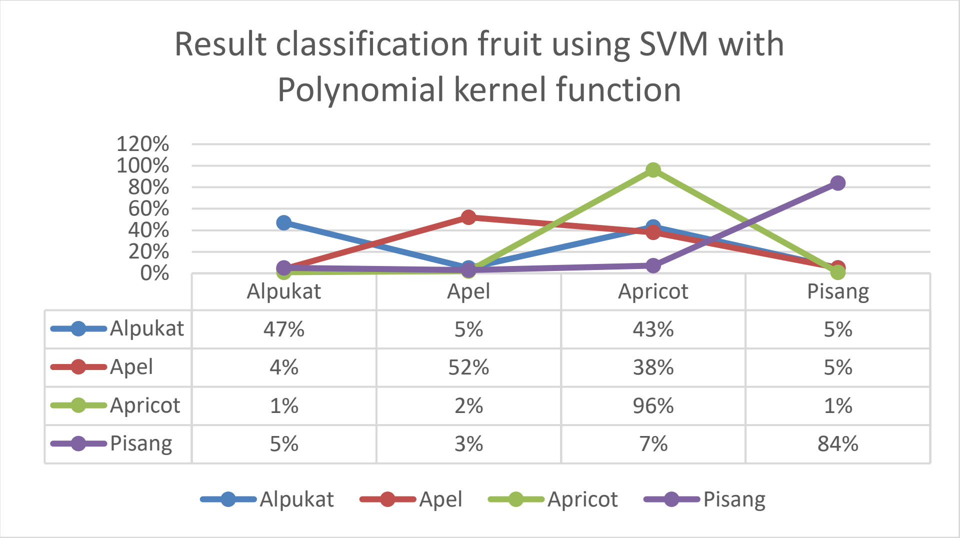


FIGURE 5. Result classification fruit using SVM with Polynomial kernel function

TABLE 3.
SVM test result using Linear kernel function

	Avocado	Apple	Apricot	Banana
Avocado	**32%**	19%	44%	5%
Apple	0%	**60%**	31%	9%
Apricot	0%	4%	**95%**	1%
Banana	2%	1%	3%	**94%**

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

Table 3 shows the SVM classification results using the Linear kernel function and the extraction of the SURF features, in the avocados the percentage is recognized at 32%, in the Apple the percentage is recognized at 60%, the fruit of the apricot percentage is recognized at 95% and in the banana fruit, the percentage is recognized by 94%. With an average total accuracy of 70.25%. Figure 6 illustrates the SVM classification results using Linear kernel function and SURF feature extraction.

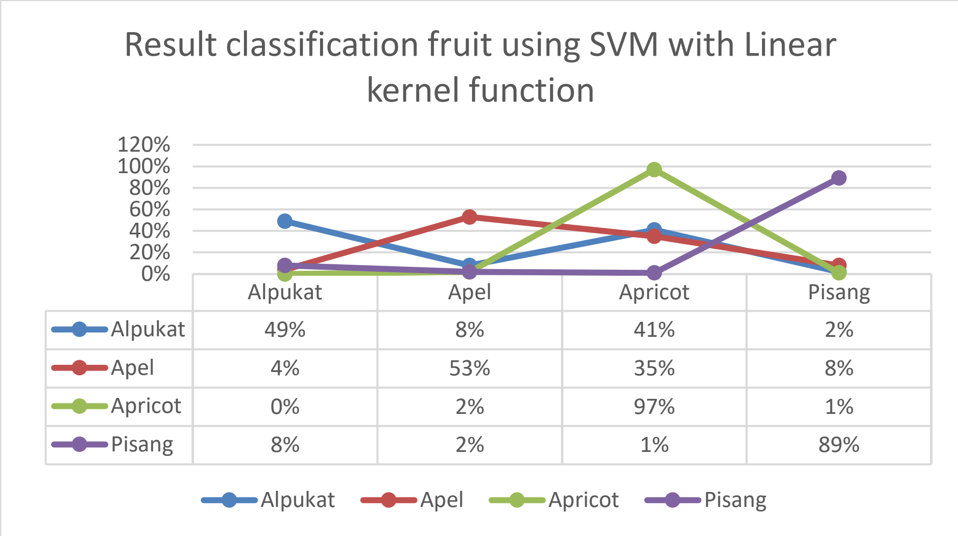


FIGURE 6. Result classification fruit using SVM with Polynomial kernel function

4. CONCLUSION

The introduction of fruit imagery by utilizing surf feature extraction and classification method SVM has not achieved maximum results, so it needs to be researched using various classification methods and other feature extractions. The gaussian SVM kernel function is more optimal than Linear and Polynomial kernel functions in the case of image recognition of the fruit.

REFERENCES

[1] S. Komarayanti, “Ensiklopedia Buah-buahan Lokal Berbasis Potensi Alam Jember Encyclopedia Of Local Fruits Based On Natural Potential Jember,” *Jurnal Biologi dan Pembelajaran Biologi*, vol. 2, no. 1, pp. 61–75, 2017.

[2] M. Muhathir, " Klasifikasi Ekspresi Wajah Menggunakan Bag of Visual Words," *JITE (Journal Of Informatics And Telecommunication Engineering)*, vol. 1, no. 2, pp. 73-82, 2018.

[3] D. Recognition, B. On, and M. Features, “Durian recognition based on multiple features and linear discriminant,” Malaysian Journal of Computer Science, pp. 57–72, 2018.

[4] W. Astuti, S. Dewanto, K. Edi, and N. Soebandrija, “Automatic fruit classification using support vector machines : a comparison with artificial

neural network Automatic fruit classification using support vector machines : a comparison with artificial neural network," IOP Conference Series: Earth and Environmental Science PAPER, pp. 0–10, 2018, doi: 10.1088/1755-1315/195/1/012047.

[5] Y. Zhang and L. Wu, "Classification of Fruits Using Computer Vision and a Multiclass Support Vector Machine," Sensors 2012, pp. 12489–12505, 2012, doi: 10.3390/s120912489.

[6] M. A. M. Nayak, "Fruit Recognition using Image Processing," International Journal of Engineering Research & Technology (IJERT), vol. 7, no. 08, pp. 1–6, 2019.

[7] S. Fatima and M. Seshashayee, "Healthy Fruits Image Label Categorization through Color Shape and Texture Features Based on Machine Learning Algorithm," International Journal of Innovative Technology and Exploring Engineering (IJITEE), no. 3, pp. 34–40, 2020, doi: 10.35940/ijitee.B7740.019320.

[8] Muhathir, R. A. Rizal, J. S. Sihotang and R. Gultom, "Comparison of SURF and HOG extraction in classifying the blood image of malaria parasites using SVM," *International Conference of Computer Science and Information Technology (ICoSNIKOM)*, 2019.

[9] M. S. Mahalakshmi, H. Srinivas, S. Meghana, and C. S. Ashwini, "Fruit Identification and Classification Techniques A Review Using Neural Networks Approach," *International Journal of Advanced Research in Computer and Communication Engineering,* vol. 4, no. 12, pp. 345–349, 2015, doi: 10.17148/IJARCCE.2015.41280.

[10] N. Kumari, D. K. Bhatt, R. K. Dwivedi, and R. Belwal, "Performance Analysis of Support Vector Machine in Defective and Non Defective Mangoes Classification,"International Journal of Engineering and Advanced Technology (IJEAT) , no. 4, pp. 1563–1572, 2019.

[11] N. N. Rade, "Image Texture Classification: SURF with SVM," International Journal of Innovations in Engineering Research and Technology [IJIERT], vol. 4, no. 5, pp. 43–47, 2017.

[12] D. A. Putra, "Klasifikasi Jenis Apel Menggunakan Support Vector Machine (SVM)," pp. 0–1, 2017.

[13] N. Wijaya and A. Ridwan, "Klasifikasi Jenis Buah Apel Dengan Metode K-Nearest Neighbors," *J. SISFOKOM*, vol. 08, no. 01, pp. 74–78, 2019.

[14] Muhathir and M. H. Santoso, "Analysis Naïve Bayes In Classifying Fruit by Utilizing Hog Feature Extraction," *JITE (Journal Of Informatics And Telecommunication Engineering),* vol. 4, no. 1, pp. 151-160, 2020.

[15] M. D. Y. Hapsari and T. Estiasih, "Variasi Proses dan Grade Apel (Malus sylvestris mill) pada Pengolahan Minuman Sari Buah Apel : Kajian Pustaka Processing and Grade Variation Apple (Malus Sylvestris mill) in Apple Extract Drink Processing : A Review," *J. Pangan dan Argoindustri*, vol. 3, no. 3, pp. 939–949, 2015, [Online]. Available: https://jpa.ub.ac.id/index.php/jpa/article/view/216/223.

[16] S. Rahman, "Studi Pendahuluan Pengaruh Alpukat Terhadap Profil Lemak di Poli Penyakit Dalam Klinik Iman," pp. 1–9.

[17] H. Saputra, N. Nazir, and D. Sylvi, "Ekstraksi dan karakterisasi minyak alpukat (persea americana, mill) solok, sumatera barat," *J. Teknol. Ind. Pertan.*, pp. 47–53, 2014.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

[18] M. B. Sartiah, P. Studi, P. Tata, F. Teknik, and U. N. Jakarta, “Pengaruh Penggunaan Masker Buah Aprikot (Prunus armeniaca) Kering Terhadap Kelembapan,” pp. 24–30.

[19] A. Z. Q, “Pisang Buah (Musa Spp): Keragaman Dan Etnobotaninya Pada Masyarakat Di Desa Sri Kuncoro Kecamatan Pondok Kelapa Kabupaten Bengkulu Tengah,” Semirata 2013 FMIPA Unila, no. 1995, pp. 33–40, 2013.

[20] M. D. Y. Ambarita, E. S. Bayu, and H. Setiado, “Identifikasi Karakter Morfologis Pisang (Musa Spp.) Di Kabupaten Deli Serdang,” *Jurnal Agroekoteknologi*, vol. 4, no. 1, pp. 1911–1924, 2016, doi: 10.32734/jaet.v4i1.12404.

[21] S. Naik and B. Patel, “Machine Vision based Fruit Classification and Grading - A Review,” *Int. J. Comput. Appl.*, vol. 170, no. 9, pp. 22–34, 2017, doi: 10.5120/ijca2017914937.

[22] Z. Hamizan and R. Sumiharto, “Sistem Pentautan Citra Udara Menggunakan Algoritme SURF dan Metode Reduksi Data,” *IJEIS (Indonesian J. Electron. Instrum. Syst.*, vol. 7, no. 2, p. 127, 2017, doi: 10.22146/ijeis.18240.

[23] J. P. Tanjung and Muhathir, "Classification of facial expressions using SVM and HOG," *JITE (Journal Of Informatics And Telecommunication Engineering)*, vol. 3, no. 2, pp. 210-215, 2020.

[23] F. P. Hilman and U. Telkom, “Perbandingan Metode Surf Dan Sift Dalam Sistem Identifikasi Tanda Tangan,” *e-Proceeding of Engineering,* vol. 2, no. 2, pp. 2467–2481, 2015.

[25] Muhathir, T. T. S. Sibarani and A.-K. Al-Khowarizmi, "Analysis K-Nearest Neighbors (KNN) in Identifying Tuberculosis Disease (Tb) By Utilizing Hog Feature Extraction," *Al'adzkiya International of Computer Science and Information Technology (AIoCSIT) Journal,* vol. 1, no. 1, May 2020.

[26] Muhathir. H. Mawengkang and M. Ramli, "Kombinasi Z-Fisher Transform dan Bray Curtis Distance Untuk Pengenalan Pola Huruf Jar Pada Citra Al-Quran," *Jurnal Bisman Info,* vol. 4, no. 1, 2017.