

Rabbit Type Classification Using Multi-SVM Based on Feature Extraction

Ery Mintorini*¹, Wildan Mahmud²

STMIK Kadiri, Kediri, Jawa Timur, Telp. (0354) 696770

E-mail : ery_minto@yahoo.co.id*¹, wildan.mahmud@gmail.com²

*Corresponding author

Abstract - Rabbits reputation of being cute, fluffy, cuddly critters lend then to being a popular choice for children pets. But in raising a rabbit is not easy, this depends on the type of rabbit. Rabbits that commonly pet are Rex Rabbits, American Rabbits, and Giant Rabbits. Rex Rabbits itself has some species including Rex Amber and Rex Lilac species, Giant Rabbits rabbits includes Giant Chekered rabbit, Harleyquin, dan American White Rabbit. Classification technology can be used to help the classification process of rabbits are Multi-SVM method and image feature extraction to classify rabbit species. Feature extraction used in this study is mean, variance, skewness, kurtosis, entropy. The five features are classified with Multi-SVM. The data used in this study are 125 images, consisting of 100 training images and 25 test images. The accuracy of this method reached 92%.

Keywords – Classification, Multi-SVM, Rabbit, Feature Extraction

1. INTRODUCTION

Rabbits are mammals that have a body weight of 1.35-7 kg with a length of 40-70 cm including the Leporidae family. Rabbits come from Africa to Europe. Rabbits have been kept as pets in Western countries since the 19th century, but because wild rabbits in Australia are considered pests, domestic rabbits have become illegal for pets in Queensland. But in the United States rabbits are very popular as pets during the Easter season, because of their relationship with the holidays [1].

In Indonesia, there are also two types of rabbits, which originate from Java (*Lepus Negrivicollis*) and Sumatra (*Nesolagus Netseherischlgel*). In Java, rabbits were brought by the Dutch as ornamental cattle in 1835. Until 2016 the rabbit livestock population was 1.1 million. The provinces that have the most rabbits are the provinces of West Java, Central Java and East Java with a population of 894,505 or 80% of the entire rabbit population in Indonesia. The tail is orange with a black tip. Meanwhile, from the neck to the front legs under a white body. While the Sumatran rabbit is the only native rabbit from Indonesia, its habitat is the mountain forests of the island of Sumatra. Its fur is yellowish brown, dark brown striped on the side of its forelegs and hind legs [2].

Rabbits can also be classified as free rabbits (*Lepus curpaeums*) and wild (*Oryctolagus cuniculus*). Through selection and breeding rabbits now consist of various varieties, types, sizes and benefits as livestock. Common types that are bred are American Rabbit, Angora, Belgian, California, Dutch, English, Flemish Giant, Havana, Himalayan, Rex, Harleyquin and many more with their respective characteristics.

Diverse characteristics and types make rabbits difficult to distinguish. The purpose of the identification process in rabbits can be done with the aim of combining 2 types of rabbits to get superior offspring. The classification process can utilize technology, one of which is the image

processing [3] and [4]. In the image processing, there are various algorithms that have been used in previous studies, such as research conducted by Manohar in 2018 [5]. They combine convolutional neural network and SVM for classification of animals. In the first stage, frames are extracted from the animal videos. The extracted animal frames are trained using Alex Net pre-trained convolution neural network. Further, the extracted features are fed into multi-class SVM classifier for the purpose of classification. To evaluate the performance of our system we have conducted extensive experimentation on our own dataset of 200 videos with 20 classes, each class containing 10 videos. Another research by Michael Revina in 2018 [6], was done using Face Expression Recognition (FER) and SVM. The recognition accuracy of the different techniques is from 75% to 99%. The power of algorithms, advantages are discussed elaborately to reach the aim of this survey. ROI segmentation method is used for preprocessing and it gives the highest accuracy 99%. According to feature extraction GF have less complexity which gives the accuracy always between 82.5% and 99%. The highest recognition accuracy of 99% is provided by the SVM classifier and it recognizes the several expressions such as disgust, sad, smile, surprise, anger, fear, neutral effectively.

Support Vector Machine (SVM) is a learning system that uses hypothesis space in the form of linear functions in a high-dimensional feature space, trained with learning algorithms based on optimization theory by implementing learning bias derived from statistical learning theory. SVM has the advantage of being able to find an optimal separator (classifier) function that can separate two data sets from two different classes. But SVM when first introduced by Vapnik can only classify data into two classes (binary classification). While problems in the real world generally have many class. One way to implement SVM Multiclass is by combining several binary SVM.

2. RESEARCH METHOD

At this point we will explain the stages of the classification process that have been carried out. Prepare a rabbit image from taking pictures with a cellphone camera. Furthermore, the process is carried out to transform the color image from color (RGB) to gray (grayscale), for input for the next process. To see the pre-processing steps in more detail can be seen in Figure 1.



Figure 1. Preprocessing stage of our proposed method

Feature extraction is a method of capturing features based on image histogram characteristics. Histograms show the probability of the appearance of the gray degree degree of pixels in an image [7]. Feature extraction is the process of taking the characteristics contained in objects in the image to recognize the object. Feature extraction is the first step in image classification and interpretation [8]. This process is related to the quantization of image characteristics into a group of corresponding characteristic values. General characteristics that

are used to recognize one or several objects in an image are the size [9], position or location, and the orientation or angle of the object to the reference line used [10]. One method used in feature extraction is first-order statistical feature extraction [11]. From the values in the resulting histogram, several characteristic parameters can be calculated, including the mean, variance, skewness, kurtosis, and entropy basen on [12].

2.1. Mean

The Mean (μ) feature indicates the size of the dispersion of an image, as explain in Equation (1).

$$\mu = \sum_{n=0}^N f_n P(f_n) \quad (1)$$

Where:

μ is mean

f_n is gray intensity value

$P(f_n)$ is histogram value

2.2. Skewness

Skewness α_3 indicates the relative slope of the histogram curve of an image, as follow in Equation 2.

$$\alpha_3 = \frac{1}{\alpha_3} \sum_{n=0}^N (f_n - \mu)^3 P(f_n) \quad (2)$$

Where :

α_3 is standard deviation from gray intensity value

f_n is gray intensity value

μ is mean

$P(f_n)$ is histogram value

2.3. Variance

Variance σ^2 indicates the variation of elements in the histogram of an image, as follow in Equation (3).

$$\sigma^2 = \sum_{n=0}^N (f_n - \mu)^2 P(f_n)$$

Where :

σ^2 is variance

f_n is gray intensity value

μ is mean

$P(f_n)$ is histogram value

2.4. Kurtosis

Kurtosis α^4 shows the relative degree of shock of the histogram curve of an image as follow in Equation (4).

$$\alpha^4 = \frac{1}{\sigma^4} \sum_{n=0}^N (f_n - \mu)^4 P(f_n) - 3$$

Where :

α^4 is variance

f_n is gray intensity value

μ is mean

$P(f_n)$ is histogram value

2.5. Entropy

Entropy feature (H) indicates the size of the irregular shape of an image, as shown in Equation (5).

$$H = - \sum_{n=0}^N P(f_n) \cdot 2_{\log P(f_n)}$$

Where:

$P(f_n)$ is histogram value

2.6. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a machine learning method for pattern recognition developed by Vladimir Vapnik. This algorithm is included in the form of supervised learning that can be used to solve classification problems. SVM classifies data into two different classes by making decision boundaries or commonly called hyperlanes that maximize margins. The following linear models are generally used in SVM to produce hyperlane, as follow in Equation (6).

$$y = \text{sign}(W^T x + b) \quad (6)$$

Where:

$X \in R$ is vector whose dimensions depend on n number of features

$y \in \{-1,1\}$ is the target of vector x

$W \in R^n$ is a vector which is a weight

$b \in R$ is a scalar bias error

The resulting hyperlane divides the data into two classes, namely positive and negative classes which are modeled according to Equation (7) and Equation (8).

$$W^T x_i + b \geq 1, \text{ for } y_i \text{ is } 1 \quad (7)$$

$$W^T x_i + b \leq -1, \text{ for } y_i \text{ is } -1 \quad (8)$$

Thus becoming

$$y_i(W^T x_i + b) \geq 1 \quad (9)$$

The resulting optimal hyperlane is a hyperlane that maximizes the minimum distance between the two hyperlanes above. In this study, a Radial Basis Funtion (RBF) kernel was used to facilitate distance searching, according to Equation (9).

$$K(x, x') = \exp\left(-\frac{\|x-x'\|^2}{2\sigma^2}\right) \quad (9)$$

$\|x - x'\|^2$ may be recognized as the squared Euclidean distance between the two feature vectors. α is a free parameter.

3. RESULTS AND DISCUSSION

After the extraction feature values have been obtained. These values will be classified by the Multi-SVM method. Classification with multi-SVM method is one of the methods used to classify data from more than two classes. There are two main approaches that are used in multi-SVM, namely determining and combining several functions to separate class classification problems. The second approach is done directly using all data from all classes in one optimization problem formation (Sembiring 2007). Support Vector Machine (SVM) works by defining the boundary between two classes with the maximum distance from the closest data (Athoillah, Muhammad, 2017).

This study used 125 rabbit images consisting of 25 Rex Amber rabbits, 25 Rex Lilac rabbits, 25 Harleyquin rabbits, 25 Giant Checkered rabbits and 25 American White rabbits. The purpose of this study is to prove what percentage of successful introduction of training with the Multi-SVM method as follow in Equation (10).

$$Accuracy = \frac{\text{the amount of correct data}}{\text{the amount of all data}} \times 100\% \quad (10)$$

This research was implemented using MATLAB, where the results of the application implementation can be seen in Figure 2.

Based on the mean, skewness, variance, kurtosis, and entropy formulas, the values presented in Table 1, Table 2, Table 3, Table 4, Table 5, and Table 6. Table 1 contains feature extraction data for Rex Amber images, Table 2 contains feature extraction data for Rex Lilac images, Table 3 contains feature extraction data for American White images, Table 4 contains feature extraction data for Chekered Giant images, Table 5 contains feature extraction data for Harleyquin images, and Table 6 contains feature extraction results on the test image.

Table 1. Rabbit Image Extraction Class Rex Amber

No.	Feature Extraction				
	Mean	Entropy	Variance	Skewness	Kurtosis
1	121.1049	7.0632	1210.5775	0.53686	-0.18855
2	118.7459	6.8471	844.7549	0.3598	0.25566
3	107.3068	6.6397	611.3544	-0.22557	0.77583
...
...
20	169.5842	7.3807	4534.102	-0.46137	-1.2741

Table 2. Rabbit Image Extraction Class Giant Checkered

No.	Feature Extraction				
	Mean	Entropy	Variance	Skewness	Kurtosis
1	126.0385	7.5451	7505.4589	-0.14991	-1.4921
2	81.4198	7.4388	3298.9377	1.3135	1.1252
3	74.8043	7.1757	3772.1296	0.59203	-1.0908
...
...
20	85.8772	7.5322	3668.0889	1.0269	0.14674

Table 3. Rabbit Image Extraction Class Harleyquin

No.	Feature Extraction				
	Mean	Entropy	Variance	Skewness	Kurtosis
1	115.9059	7.6479	2663.8582	0.49234	-0.3869
2	121.5746	7.8503	3503.7669	0.13371	-0.84792
3	145.0735	7.5685	2211.2559	-0.35267	-0.22866
...
...
20	145.9057	7.4396	1936.1252	-0.56065	-0.029005

Table 4. Rabbit Image Extraction Class Rex Lilac

No.	Feature Extraction				
	Mean	Entropy	Variance	Skewness	Kurtosis
1	145.3362	7.7889	3689.6573	0.082347	-0.82303
2	148.3409	7.7258	3456.055	0.032015	-0.84195
3	143.5635	7.6089	2517.1395	0.10888	-0.38233
...
...
20	161.722	7.6364	3039.7677	-0.1718	-0.80292

Table 5. Rabbit Image Extraction Class American White

No.	Feature Extraction				
	Mean	Entropy	Variance	Skewness	Kurtosis
1	117.0787	5.7528	8911.3197	-0.11388	-1.8933
2	157.5912	7.7596	4765.5939	-0.40505	-0.93258
3	146.2312	7.7821	4992.1862	-0.13526	-1.0982
...
...
20	165.1577	6.9089	944.1182	-0.1301	-0.68184

Table 6. Rabbit Image Extraction Class of all Class

No.	Feature Extraction				
	Mean	Entropy	Variance	Skewness	Kurtosis
1	110.7532	7.2087	1336.8963	0.1807	-0.53105
...
6	101.1174	7.7958	5146.8173	0.40677	-1.1227
...
11	150.8536	7.38	1661.742	0.051104	-0.11933
...
16	144.2981	7.7085	3167.5972	0.17156	-0.80548
...
...
25	238.9433	4.0107	868.907	-3.32	14.1268

The results of the study using the Multi-SVM algorithm, the test results are matched between the original data of the input test image and the results of the Multi-SVM algorithm by using the 25 test images.

Table 7. Data extraction features of rabbit

No	Name of image	Target	Result	Information
1	amber_uji1.png	Rex Amber	Rex Amber	True
2	amber_uji2.png	Rex Amber	Rex Amber	True
3	amber_uji3.png	Rex Amber	Rex Amber	True
4	amber_uji4.png	Rex Amber	Rex Amber	True
5	amber_uji5.png	Rex Amber	Rex Amber	True

6	checkered_uji1.png	Giant checkered	Giant checkered	True
7	checkered_uji2.png	Giant checkered	Rex Amber	False
8	checkered_uji3.png	Giant checkered	Giant checkered	True
9	checkered_uji4.png	Giant checkered	Giant checkered	True
10	checkered_uji5.png	Giant checkered	Giant checkered	True
11	harley_uji1.png	HarleyQuin	HarleyQuin	True
12	harley_uji2.png	HarleyQuin	HarleyQuin	True
13	harley_uji3.png	HarleyQuin	HarleyQuin	True
14	harley_uji4.png	HarleyQuin	HarleyQuin	True
15	harley_uji15.png	HarleyQuin	HarleyQuin	True
16	lilac_uji1.png	Rex Lilac	Rex Lilac	True
17	lilac_uji2.png	Rex Lilac	Rex Lilac	True
18	lilac_uji3.png	Rex Lilac	Rex Lilac	True
19	lilac_uji4.png	Rex Lilac	Rex Lilac	True
20	lilac_uji5.png	Rex Lilac	HarleyQuin	False
21	white_uji1.png	American White	American White	True
22	white_uji2.png	American White	American White	True
23	white_uji3.png	American White	American White	True
24	white_uji4.png	American White	American White	True
25	white_uji5.png	American White	American White	True

Based on table 7, the results of the classification of 25 rabbit images contained 2 errors of detection using the Multi-SVM algorithm and using the Accuracy formula, an accuracy of 92% was obtained. The results of this study still need to be developed again given the accuracy was still less than 100%.

4. CONCLUSION

Based on the results of trials that have been carried out in this study, it is evident that the proposed method can be used to calcify the type of rabbit. However, the level of accuracy is still less than 100%, namely 92%. So for subsequent, research methods can be developed again. Maybe it can be maximized again in the process of preprocessing and image feature extraction.

REFERENCES

- [1] J. Cervantes, F. G. Lamont, A. López-Chau, L. R. Mazahua dan J. S. Ruíz, "Data selection based on decision tree for SVM classification on largedata sets," *Applied Soft Computing*, p. 787–798, 2015.
- [2] B. Baum dan M. Hewicker-Trautwein, "Classification and Epidemiology of Mammary Tumours in Pet Rabbits (*Oryctolagus cuniculus*)," *Journal of Comparative Pathology*, pp. 291-298, 2015.
- [3] X. Zeng, W. Ouyang, B. Yang, J. Yan dan X. Wang, "Gated Bi-directional CNN for Object Detection," dalam *European Conference on Computer Vision*, The Netherlands, Springer, 2016, pp. 354-369.
- [4] S. A. Medjahed , "A comparative study of feature extraction methods in images classification," *I.J. Image, Graphics and Signal Processing*, pp. 16-23, 2016.
- [5] M. Manohar, S. Kumar, H. Kumar dan R. Rani, "Deep Learning Approach for Classification of Animal Videos," dalam *Data Analytics and Learning*, Springer, 2018, pp. 421-431.
- [6] I. M. Revina dan W. S. Emmanuel, "A Survey on Human Face Expression Recognition Techniques," *Journal of King Saud University - Computer and Information Sciences*, pp. 1-9, 2018.
- [7] X. Wang dan L. Chen, "An effective histogram modification scheme for image contrast enhancement," *Signal Processing: Image Communication*, pp. 187-198, 2017.
- [8] S. Bahrampour, A. Ray, S. Sarkar, T. Damarla dan N. M. Nasrabadi, "Performance comparison of feature extraction algorithms for target detection and classification," *Pattern Recognition Letters*, pp. 2126-2134, 2013.
- [9] S. Kumar dan S. K. Singh, "Monitoring of pet animal in smart cities using animal biometrics," *Future Generation Computer Systems*, pp. 553-563, 2018.

- [10] K. Sruthitha dan S. Padmavathi, "Performance of SVM classifier for image based soil classification," *2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES)*, 2016.
- [11] F. Liantoni, R. I. Perwira, L. D. Putri, R. T. Manurung, M. S. Kahar, J. Safitri, L. T. Muharlisiani, D. Chamidah, A. Ghofur dan P. S. Kurniawan, "Watermelon classification using k-nearest neighbours based on first order statistics extraction," *Journal of Physics: Conference Series*, pp. 1-6, 2019.
- [12] I. Usta dan Y. M. Kantar, "Mean-Variance-Skewness-Entropy Measures: A Multi-Objective Approach for Portfolio Selection," *Entropy*, pp. 117-133, 2011.