MEKATRONIKA
e-ISSN: 2637-0883
VOL. 3, ISSUE 1, 27 – 31
DOI: https://doi.org/10.15282/mekatronika.v3i1.6680

ORIGINAL ARTICLE

The Animal Classification: An Evaluation of Different Transfer Learning Pipeline

Ee Ken-ji[1], Ahmad Fakhri Ab. Nasir[2,3], Anwar P.P. Abdul Majeed[1,3], Mohd Azraai Mohd Razman[1], and Nur Hafieza Ismail[2]

[1]Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang, 26600 Pekan Pahang, Malaysia.
[2]Faculty of Computing, Universiti Malaysia Pahang, 26600 Pekan Pahang, Malaysia.
[3]Centre for Software Development & Integrated Computing, Universiti Malaysia Pahang, 26600 Malaysia.

ABSTRACT – The animal classification system is a technology to classify the animal class (type) automatically and useful in many applications. There are many types of learning models applied to this technology recently. Nonetheless, it is worth noting that the extraction of the features and the classification of the animal features is non-trivial, particularly in the deep learning approach for a successful animal classification system. The use of Transfer Learning (TL) has been demonstrated to be a powerful tool in the extraction of essential features. However, the employment of such a method towards animal classification applications are somewhat limited. The present study aims to determine a suitable TL-conventional classifier pipeline for animal classification. The VGG16 and VGG19 were used in extracting features and then coupled with either *k*-Nearest Neighbour (*k*-NN) or Support Vector Machine (SVM) classifier. Prior to that, a total of 4000 images were gathered consisting of a total of five classes which are cows, goats, buffalos, dogs, and cats. The data was split into the ratio of 80:20 for train and test. The classifiers hyper parameters are tuned by the Grids Search approach that utilises the five-fold cross-validation technique. It was demonstrated from the study that the best TL pipeline identified is the VGG16 along with an optimised SVM, as it was able to yield an average classification accuracy of 0.975. The findings of the present investigation could facilitate animal classification application, i.e. for monitoring animals in wildlife.

ARTICLE HISTORY
Received: 24th March 2021
Revised: 3rd May 2021
Accepted: 17th June 2021

KEYWORDS
Animal Classification
Transfer Learning
VGG Model
SVM
k-NN

INTRODUCTION

Nowadays, artificial intelligence (AI) is a technology that can be found regularly in our daily life. With this technology, human's daily life, safety, and even the workstation can be taken to a new level of efficiency. One of the applications of AI is the animal classification system which is used to classify the animal type or species through the extracted features from the image data. This is a system that generally can be applied to a few fields of study such as zoology, wildlife and conservation biology, and animal behaviour [1] such as the trap camera in the jungle to observe the animal and the security system to warn the human when the predator appears near the camp-side. An animal classification system can be used by invoking the intelligence to the vision sensor such as a camera to work automatically and eventually helps humans to avoid the animal attack due the mistake by human behaviour (i.e. tiredness). In other words, the animal classification system is important as it helps people to gain more knowledge about the animals and be able to protect both people and animals at the same time in better ways.

Another interesting application of animal classification and detection systems is for animal and vehicle collisions which have been a growing concern in Malaysia particularly since the abundant wildlife resources, abandoned livestock animals, as well as the increase of automobiles. Such problems cause hundreds of people deaths, thousands of human injuries, billions of dollars in property damage and countless animal deaths every year. This can be proven by the news in the mainstream media such as in [2][3]. The department of wildlife and national parks (Jabatan Perhilitan) reported that during 2012-2017, 2,444 species of 36 wildlife protected under the Wildlife Conservation Act (Act 716) were killed in this bloody tragedy, including tapir, bear, elephant, beetle tiger, striped tiger, and goat [4]. By looking at the population of these animals and the statistics of road accidents involving wildlife, it is clearly shown that accidents are one of the threats to wildlife and implementation of animal classification and detection system is deemly important to help reduce human errors.

RELATED WORK

During the study conducted, at least 30 articles have been reviewed on the combination of animal classification and detection within the last five years. From the review, it is obvious that deep learning is the most common method selected by the researcher. However, one of the interesting findings is that transfer learning becomes popular starting from 2017 onward. Moreover, some researchers also compare the effectiveness of both deep learning and transfer learning. Transfer learning can be more accurate, faster and less data training since it uses a pre-trained model to solve a

*CORRESPONDING AUTHOR | Ahmad Fakhri Ab. Nasir | ✉ afakhri@ump.edu.my

new issue and utilize the knowledge gained from previous tasks to facilitate the generalization in another training. It has been employed in various applications such as EEG signals [5], wafer defects [6] as well as in the diabetic retinopathy diagnosis [7].

A short literature survey on animal classification and detection was gathered from [8] to [12]. Nguyen and et al. in [8] identified the most effective model for animal classification using the deep learning approaches which are Lite AlexNet, VGG16 and ResNet50. The models tested with 18 species of wildlife animals, and the results showed good accuracy which was 82.49%, 83.93% and 84.39% for AlexNet, VGG16 and ResNet50 respectively. Besides, Horn and et al. in [9] showed the result of 80.2% and 95.21% on the top-1 and top-5 accuracy in overall for the ResNetV2 SE model. Although the model showed a better result as the comparison between all the nine models tested, the image dataset acquired should be more relevant as the precision of the model is not enough and their training set amount is unbiased and the models also show an unsatisfied result on accuracy of certain species.

The research in [10] achieved an accuracy of 95% in identifying the general test on the individual giant panda's face. The study conducted under different image pre-processing such as rotated, brighten, and darken. The accuracy shows a highly satisfied result as it is between the range of 90%-95% for all preprocessed images. Abbadi and Alsaadi [11] tested the deep CNN model on the mammals, reptiles, amphibians, birds, and fishes. The overall accuracy on the classification is 97.5% with the best image size to input into the Deep CNN model is 50x50. Moreover, Jamil et al. [12] tested eight transfer learning models on bear and sheep classification and the results could yield around 90% - 98%. This experiment was conducted with four feature extraction models and two classifiers, *k*-NN and SVM. The highest accuracy model was Inceptionv3 + *k*-NN which gives classification accuracy of 98.3%.

To the best of the authors' knowledge, the capability of a hybrid Transfer Learning-conventional classifier pipeline for multi-classes animal classification is still limited. Therefore, the objective of this paper is to appraise the ability of different TL-based models in extracting features that are then classified by an optimised *k*-NN and SVM model. It is hypothesised that the proposed technique could distinguish well the different categories of animal images.

METHODOLOGY

The setup for the animal classification consists of five steps, namely data collection, preprocessing, feature extraction, classification, and finally the performance evaluations. To simulate the recognition of the animal, few resources of animal images were used in this research, which are KTH animal image dataset, Kaggle image dataset, Mendeley Dataset, and Google Image. Figure 1 shows the sample of the animal image data used in the study and Table 1 shows the distribution of data for each class. In total, 4000 images were collected consisting of almost equally distributed images per class; 850 cow images, 850 goat images, 700 buffalo images, 800 cat images, and 800 dog images.

Figure 1. Sample of images.

The images were then resized to 224x224 to fit the input dimension of the transfer learning model of VGG16 and VGG19. In this study, the image data was split into two sets, which are train dataset and test dataset for performance measures. The data splitting ratio was set to 80:20, train:test. 80% of the total image dataset (3200 images) was used to train the system to fit the model. Then the remaining 20% of the image dataset (800 images) was used for the test dataset to provide an unbiased evaluation of the final model fit on the training dataset. The stratified splitting method was chosen which separates the dataset into homogenous classes or subpopulation which is known as strata. This division is based on the specific characteristics of the data. After the dataset is divided into strata, the data will be selected randomly to form the sample.

Table 1. Data distribution for each class

Class	Resource				Total
	KTH	Kaggle	Google Image	Mendeley	
Cow	97	730	23	0	850
Goat	99	50	21	680	850
Buffalo	0	375	0	325	700
Cat	0	800	0	0	800
Dog	0	800	0	0	800
Total	196	2755	44	1005	4000

Feature Extraction: Transfer Learning (VGG16 and VGG19)

There are few models for the transfer learning, and only two models were selected for this study which are VGG-16 and VGG-19. The VGG models proposed by the Visual Geometric Group from Oxford University. Both VGG-16 and VGG-19 used three types of layers, which are convolution layers, maximum pooling layers, and the fully connected layers. The fully connected layers will not be discarded (freezed) when the model is used for feature extraction while the fully connected layers in a layer undergo classification. For the VGG16 model, there are the trainable parameters included in these 16 weighted layers which are 13 convolution layers and 3 fully connected layers after the last pooling layer, however these fully connected layers were not used in this study and substitute with conventional classifier (*k*-NN and SVM). For the VGG19 model, the difference is it has extra three layers of the weighted convolutional layers compared to the VGG16 model but the model retains the same amount of pooling layers [13].

Classifier: *k*-Nearest Neighbour (*k*-NN) and Support Vector Machine (SVM)

K-nearest neighbour (*k*-NN) is a kind of simple algorithm that is easy to implement supervised learning. The *k*-NN algorithm can be used to solve both the problem of classification and regression in supervised learning. This algorithm is used to assume the similar things that exist in close proximity or the similar things that are near to each other. *k*-NN works can be done by finding the distance between a query and all the examples in the data with the selection of the specified number examples, the *k*, closest to the query. The *k*-NN classifier has two types of parameters that can be tuned, which is the amount or number of *k*, and the method of the measurement for the distance. For the *k* value, it has few thoughts while picking, the first thought is there's no physical or biological way to determine the best value of *k*, it needs to be a try and error method. Secondly, low values of *k* may be noisy and subject to the effect of outliers. The example of distance measurements are Euclidean, Manhattan, and Chebyshev measurement or distance.

Support Vector Machine (SVM) is a technique which represents an algorithm that is used for the classification and regression analysis in supervised learning for the purpose of analysing the data. The classification that is conducted under this classifier will be performed by finding the hyper-plane or line that can differentiate the classes. As the hyper-plane or line is found, SVM will separate between the classes by the maximisation of the margin. The success of the SVM classifier depends on its kernel parameters. This kernel is a method to prepare for SVM to compute the dot product of two vectors for the features. The main kernels for the SVM are linear, sigmoid and radian basic function (RBF).

To perform optimization of hyperparameters tuning using the grid search approach, the hyperparameters and the classifier have to pre-define as the dictionary and passed to the library, GridSearchCV as well. This allows the system to try all the combinations of hyperparameters and provide the result of each combination to make selection on the best parameter. Table 2 tabulates the hyperparameters that were tuned using the grid search approach in the present study through the five-fold cross-validation technique.

Table 2. List of hyperameters values in *k*-NN and SVM classifier.

Clasifier	Hyperparameters	Values
k-NN	No. of Neighbour, *k*	1-20
	Metrics / Distance	Euclidean, Manhattan, Chebyshev
SVM	Kernel Type	Linear, RBF, Sigmoid
	Kernel Parameter	Degree, d = 2, 3; Gamma = 0.1, 1, 10; Cost, c = 0.1, 1, 10

Performance Measure

A stratified 80:20 ratio hold-out strategy was used for splitting the training and testing, respectively. In the present study, different performance measures were used to evaluate the developed pipelines, namely classification accuracy (CA), precision, recall, and F1-score, apart from the confusion matrix. The models were developed and evaluated on a Python IDE, i.e. Spyder3.7 with associated Keras and sklearn libraries.

RESULTS AND DISCUSSIONS

From the obtained result shown in Table 3, all the four models have no presence of over-fitting in the result for classification accuracy (CA). Three pipelines shows good CA on its training dataset which are VGG16 + SVM model, VGG19 + SVM model and VGG19 + *k*-NN model with the accuracy of 100%, and the result is shows that the two pipelines of SVM sharing the same CA. By comparing the two pipelines, the SVM pipelines show the CA of 95% on its testing dataset while the VGG19 + *k*-NN model shows much lower CA of 72%. Hence the SVM model has resulted in better CA. As the two models of SVM have the same CA, the consideration to select the best model is required to look for the training time factor.

Table 3. The best classification accuracy on each pipeline

Pipeline	Classification Accuracy		
	Training	Testing	Average
VGG16 + *k*-NN (k = 14, Euclidean)	0.75	0.68	0.715
VGG19 + *k*-NN (k = 1, Euclidean)	1.00	0.72	0.860
VGG16 + SVM (c = 0.1, d = 2, gamma = 0.1, Linear)	1.00	0.95	0.975
VGG19 + SVM (c = 0.1, d = 2, gamma = 0.1, Linear)	1.00	0.95	0.975

Referring to Table 4, it is obvious that the VGG16 + SVM model needs 57.15 seconds to train while the VGG19 + SVM model needs 59.75 seconds. Although it only has a difference of about two seconds, the VGG16 + SVM model still has a shorter training time. The reason that makes the VGG16 having shorter training time may be due to the simpler architecture of the VGG16 as it has fewer convolution layers compared to VGG19 Hence, VGG16 + SVM can be concluded as the best pipeline to classify five categories of animals.

Table 4. Training time of SVM pipeline.

Pipeline	Training Time (sec)
VGG16 + SVM	57.15
VGG19 + SVM	59.75

From the Table 3, it is obvious that VGG16 + SVM model is needed 57.15 s to training while VGG19 + SVM model is needed 59.75s. Although it only has the difference of around of 2s, but the VGG16 + SVM model has shorter training time. The reason that make the VGG16 having shorter training time may be due to the simpler architecture of the VGG16 as it has fewer convolution layers compare to VGG19 Hence, VGG16 + SVM model is the model with the best pipeline.

Table 5. Recall of each class for VGG16+SVM pipeline.

Class	Recall		
	Training	Testing	Average
Cow	1.00	0.93	0.965
Goat	1.00	0.91	0.955
Cat	1.00	0.99	0.995
Buffalo	1.00	0.96	0.980
Dog	1.00	0.96	0.980
Average	1.00	0.95	0.975

The accuracy of each class of the VGG19 + SVM model is shown in Table 5 or known as recall in the classification report. The goat class shows lowest accuracy as it has 91% of the accuracy in the test set. It is followed by the cow class which is 93% and the buffalo and dog class had the same score which is 96%. Lastly, the cat class in this research gives 99% which is the highest recall accuracy. The summary of performance of the best pipeline (VGG16+SVM) is shown in Table 6.

		Predicted				
		Cow	Goat	Cat	Buffalo	Dog
Actual	Cow	158	5	0	6	1
	Goat	12	154	1	0	3
	Cat	0	0	159	0	1
	Buffalo	2	3	1	134	0
	Dog	0	3	4	0	153

Figure 2. Confusion matrix for VGG16+SVM pipeline on test dataset.

Table 6. The performance summary for VGG16+SVM pipeline.

Dataset	Multiclass	Performance Metric			
		Precision	Recall	F1 Score	Accuracy
Train	Macro Average	1.00	1.00	1.00	1.00
	Weight Average	1.00	1.00	1.00	
Test	Macro Average	0.95	0.95	0.95	0.95
	Weight Average	0.95	0.95	0.95	

CONCLUSION

The present study evaluated different TL-optimised k-NN and SVM pipelines in the classification of animals. It was shown from the preliminary investigation carried out that the VGG16+SVM pipeline is the best and could attain a CA of 100% for the training dataset as well as 95% for the test dataset. Besides, the pipeline took about 57.15 seconds to train all of the data. The outcome of the study is non-trivial, mainly towards the realisation of a larger animals classifications implementation. Future studies shall attempt on the evaluation of other TL pipelines, classifiers as well as optimisation techniques. Besides, in realisation of largest animal classes classification, the further studies shall add massive animals images from much more classes to further generalise the pipelines.

REFERENCES

[1] S. N. Mohammad, N. Anh, K. Margaret, S. Alexandra, S. P. Meredith, P. Craig, and C. Jeff, “Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning”, *Proceedings of the National Academy of Sciences*, vol. 115 no. 25, pp. 5716-5725, 2018, doi: 10.1073/pnas.1719367115.

[2] A. R. Zulkifli, 2019, Accessed 16 July 2021, “Gajah mati dirempuh lori”, Harian Metro <https://www.hmetro.com.my/mutakhir/2019/07/480469/gajah-mati-dirempuh-lori-metrotv>.

[3] R. Omar, 2019, Accessed 16 July 2021, “Tiga pemandu nyaris maut, kereta langgar kerbau”, Sinar Harian, <Available at: https://www.sinarharian.com.my/article/63391/EDISI/Terengganu/Tiga-pemandu-nyaris-maut-kereta-langgar-kerbau>.

[4] M. A. L. Abu Bakar, 2019, Accessed 16 July 2021, “Sikap pemandu dan kadar kematian hidupan liar di musim perayaan”, Astro Awani, <http://www.astroawani.com/berita-malaysia/sikap-pemandu-dan-kadar-kematian-hidupan-liar-di-musim-perayaan-209371>.

[5] J. L. Mahendra Kumar, M. Rashid, R. Muazu Musa, M. A. Mohd Razman, N. Sulaiman, R. Jailani, and A. P.P. Abdul Majeed, A. “An Evaluation of Different Fast Fourier Transform - Transfer Learning Pipelines for the Classification of Wink-based EEG Signals”, *MEKATRONIKA*, vol. 2, no. 1, pp. 1–7, 2020, doi: 10.15282/mekatronika.v2i1.4881.

[6] J. A. Mat Jizat, A. P.P. Abdul Majeed, A. F. Ab. Nasir, Z. Taha, and E. Yuen, “Evaluation of the machine learning classifier in wafer defects classification”, ICT Express, 2021, doi: 10.1016/j.icte.2021.04.007.

[7] F. N. Mohd Noor, W. H. Mohd Isa, I. Mohd Khairuddin, M. A. Mohd Razman, J. A. Mat Jizat, A. F. Ab Nasir, R. Muazu Musa, A. PP Abdul Majeed, “The Diagnosis of Diabetic Retinopathy: A Transfer Learning with Support Vector Machine Approach”, Advances in Robotics, Automation and Data Analytics: Selected Papers from ICITES 2020, vol. 1350, pp. 391-398, 2021.

[8] H. Nguyen, S. J. Maclagan, T. D. Nguyen, T. Nguyen, P. Flemons, K. Andrews, E. G. Ritchie, and D. Phung, “Animal Recognition and Indentification with Deep Convolution Neural Networks for Automated Wildlife Monitoring”, *2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA),* 2017, pp. 40-49, doi: 10.1109/DSAA.2017.31.

[9] G. Van Horn, O. Mac Aodha, Y. Song, Y. Cui, C. Sun, A. Shepard, H. Adam, P. Perona, and S. Belongie, “The iNaturallst Species Classification and Detection Dataset”, *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition,* pp. 8769-8778, 2018, doi: 10.1109/CVPR.2018.00914.

[10] J. Hou, Y. He, H. Yang, T. Connor, J. Gao, Y. Wang, Y. Zeng, J. Zhang, J. Huang, B. Zheng, and S. Zhou, “Identification of animal individuals using deep learning: A case study of gaint panda”, *Biological Conservation,* vol. 242, pp. 1-6, 2020, doi: 10.1016/j.biocon.2020.108414.

[11] N. K. El Abbadi and E. M. T. A. Alsaadi, “An Automated Vertebrate Animals Classification Using Deep Convolution Neural Networks”, *2020 International Conference on Computer Science and Software Engineering (CSASE),* pp. 72-77, 2020, doi: 10.1109/CSASE48920.2020.9142070.

[12] S. Jamil, Fawad, M. S. Abbas, F. Habib, M. Umair and M. J. Khan, “Deep Learning and Computer Vision-based a Novel Framework for Himalayan Bear, Marco Polo Sheep and Snow Leopard Detection”, *2020 International Conference on Information Science and Communication Technology (ICISCT),* pp. 1-6, 2020, doi: 10.1109/ICISCT49550.2020.9080021.

[13] C. K. Yang, Y. Zheng and A. Merkulov, “Breast cancer screening using convolutional neural network and follow-up digital mammography”, *Computational Imaging III,* 2018, doi: 10.1117/12.2304564.