Pig Treatment Classification on Thermal Image Data using Deep Learning

Savina Jassica Colaco[1], Jung Hwan Kim[1], Alwin Poulose[2], Zutphen sanne van[3],
Suresh Neethirajan[3] and Dong Seog Han[1,*]
[1]School of Electronic and Electrical Engineering, Kyungpook National University, Daegu, Republic of Korea
[2]Center for ICT & Automotive Convergence, Kyungpook National University, 80 Daehak-ro, Buk-gu, Daegu 41566, South Korea
[3]Wageningen University & Research, 6700 AH Wageningen, The Netherlands

savinacolaco@knu.ac.kr[1], jkim267@knu.ac.kr[1], alwinpoulosepalatty@knu.ac.kr[2],
sanne.vanzutphen@wur.nl[3], suresh.neethirajan@wur.nl[3], dshan@knu.ac.kr[1*]

***Abstract*—Recently, image classification has gained recognition in several applications like self-driving cars, security surveillance systems, face detection, etc. The conventional methods have been overtaken by deep learning methods which can detect and classify objects in complex scenarios. In this paper, we propose a simple CNN model for pig treatment classification on thermal images. The proposed model is compared with different deep learning models which are widely used for image classification. The models are evaluated with our own thermal dataset collected using a FLIR camera. The experimental results show the thermal images of different pig treatments are better classified with the proposed model. The proposed model can achieve 99.96% accuracy with a few parameters.**

***Index Terms*—Classification, convolutional neural network (CNN), thermal images**

I. Introduction

Despite image classification being used as the oldest technique, it is still crucial for various applications. The human visual system can perform complex detection differentiating various objects and challenging scenarios compared to the computer system [1]. Image classification involves classifying the target into different classes and inferring from the visual identification into further applications. It is used in object identification, medical imaging, facial recognition and more applications. The image classification technique is complicated due to challenges in the images such as occlusion, illumination, pose variations, etc. Deep learning has outperformed conventional techniques in several aspects. Although it has been used in different fields, it is difficult in the field of classification of animal species. Due to challenges in animal images, automatic classification has remained an unsolved problem [2]. Deep learning consists of various hidden layers which can help in learning semantic representations at different abstraction levels. Moreover, it uses raw data for processing networks with the least human interference, which improves the probability of detection in complex cases. In this paper, we conduct a comparison study with deep learning methods with the proposed model. These experiments classify pig treatment images taken from a thermal camera. It is further evaluated with classification metrics such as accuracy, confusion matrix and more.

II. Proposed Model

Pig treatment classification is compared using different models such as a LeNet5 [3], AlexNet [4], VGGNet [5] and the proposed model. LeNet5 neural networks consist of five layers with learnable parameters. This network consists of three sets of convolutional layers and a combination of max-pooling layers followed by fully connected hidden layers. The filter size used is 3×3 and the pooling size is 2×2 with stride 1. The hidden layers have 500 neurons. A softmax classifier is applied to classify the images. AlexNet won the Imagenet large-scale visual recognition challenge in 2012. In this model, the network depth was increased compared to the LeNet5 network. It consists of eight layers with learnable parameters. The model involves five layers with a max-pooling combination followed by three fully connected layers. The layers use rectified linear unit activation (ReLU) as their activation function which accelerated the speed of the training process. The model also uses dropout layers to prevent overfitting. The last layer uses softmax as its activation function.

The visual geometry Group of the University of Oxford (VGGNet) developed a deep convolutional neural network that is popularly used in computer vision fields. It consists of VGG-16 or VGG-19 referring to the 16 and 19 convolutional layers. In this paper, we evaluated the VGG-16 model which is supported by 16 layers for comparison with other models. The convolutional layers are followed by single max pool layers. The layers leverage a minimal receptive field of 3x3 kernel size. These are followed by the ReLU unit which reduces the training time compared to AlexNet. The proposed model is a simple sequential model consisting of several convolutional layers and a batch normalization layer. The convolutional layers are followed by Leaky ReLU, which is based on a ReLU but has a small slope for negative values instead of a flat slope. Max pooling is followed after every even convolution layer to

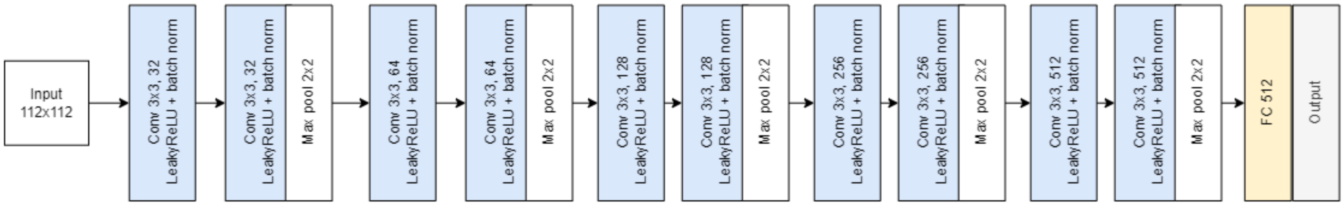
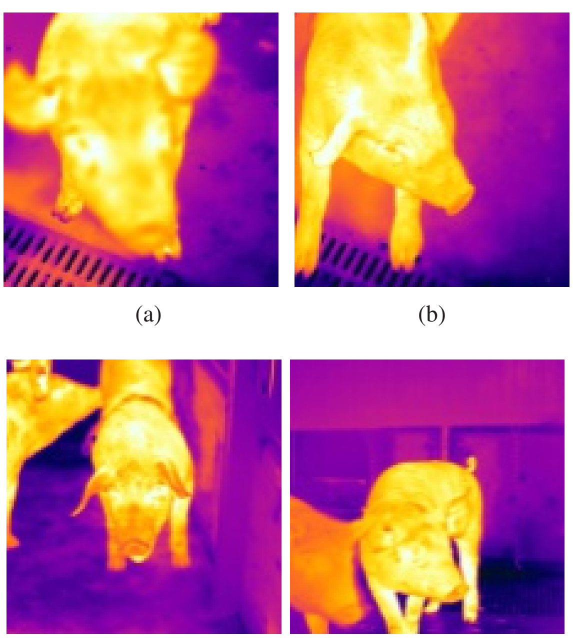


Fig. 1. Proposed model.

reduce the spatial dimension of the feature map. Fig. 1 shows the proposed model used for the classification.

III. Experiment

Thermal videos were collected by Wageningen University & Research with the help of a FLIR camera. Images extracted from videos are resized and converted to grayscale. The total images consist of 62,800 images with four labels such as isolation after feeding (IAF), isolation before feeding (IBF), paired after feeding (PAF) and paired before feeding (IBF) for pig treatment. The thermal images have been resized to 112×112 resolution. The Keras framework is used for model implementation with a batch size of 32 and epochs of 100. The model is optimized with the Adam optimization technique [6] with a learning rate of 10^{-3}. Further, the whole data is split into 60, 20 and 20 ratios for train, test and validation data. For the model training, categorical cross-entropy is used to classify the pig treatments.



(a) (b) (c) (d)

Fig. 2. Thermal images. (a) IAF. (b) IBF. (c) PAF. (d) PBF.

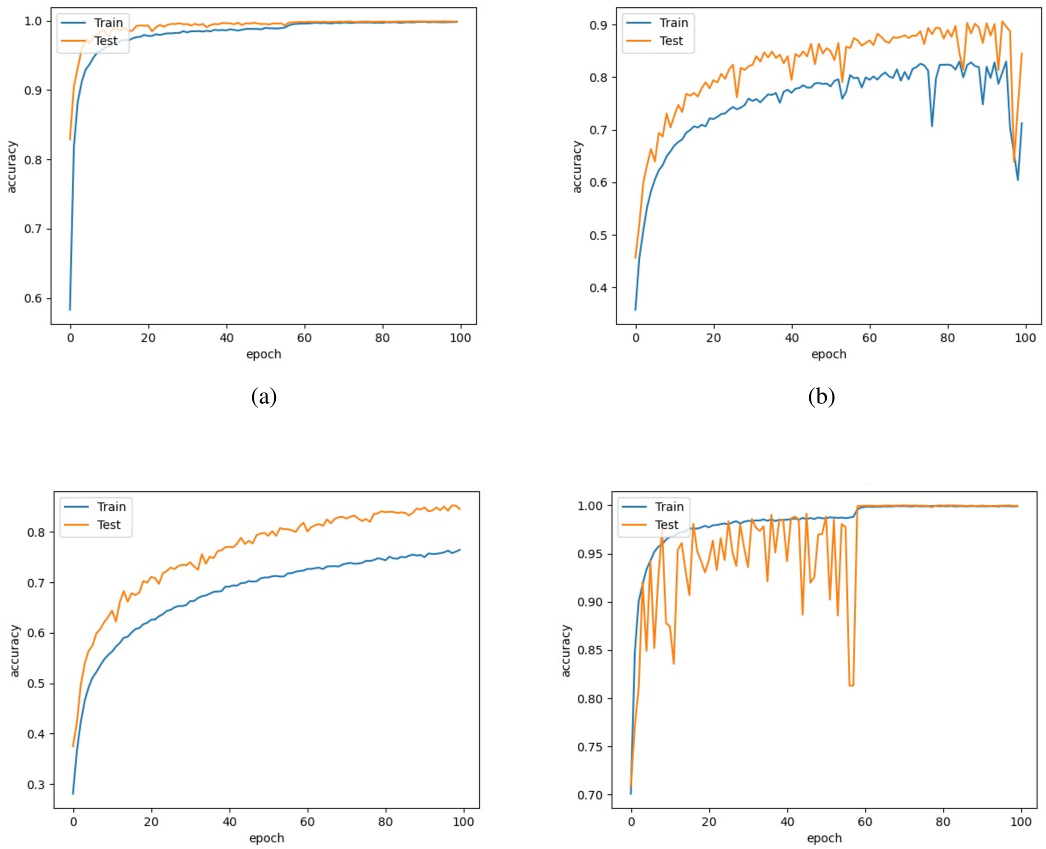
IV. Discussion

The comparison of the different models is evaluated with accuracy, F1 score, precision, recall and number of parameters. Accuracy on validation data measures how often the classifier correctly predicts. Precision metric explains how many of the correctly predicted cases actually turned out to be positive. It is useful where false positive is a more serious concern than false negatives. Recall describes how many of the actual positive cases we were able to predict correctly with the model. It is helpful where false negative is of deeper concern than false positive. F1 score combines both precision and recall metrics. The confusion matrix is a widely used performance measurement for classification problems where output can be two or more classes. We evaluated our results with an accuracy plot and confusion matrix in Figs. 3 and 4.

LeNet5 provides an accuracy of 99.90% with 19.6 M parameters. The model converges well after a certain epoch. According to the Lenet5 confusion matrix, the model is able to classify the paired before feeding treatment class accurately compared to other classes. Though AlexNet provides an accuracy of 90.22% with a large number of parameters, it has several images that are misclassified. The model is also overfitting for the validation data. The model is able to classify isolation before feeding class the most compared to other classes. VGGNet showed decreasing performance compared to other models with a large convergence gap between train and test data. The model provided 85.43% of accuracy with 17 M parameters. Like AlexNet, the model is overfitting due to the wide distribution of data. The proposed model showed similar accuracy of 99.96% with Lenet5 but with a 7.2 M number of parameters. The model showed fluctuations during the training due to the wide distribution of data but converged well after certain epochs. It also classified most of the pig treatment classes with increased performance. With the help of an L2 regularizer, the overfitting in the proposed model is reduced.

TABLE I: Comparison with different CNN models with the proposed model

Model	Parameters	Accuracy	F1 score	Precision	Recall
LeNet5	19.6 M	99.90446	99.90446	99.90455	99.90446
AlexNet	23.3 M	90.22293	90.25812	90.53939	90.22293
VGGNet	17 M	85.43790	85.50911	86.31478	85.43790
Proposed Model	7.2 M	99.96815	99.96815	99.96816	99.96815



(a) (b)

(c) (d)

Fig. 3. Accuracy. (a) LeNet5. (b) AlexNet. (c) VGGNet. (d) proposed model.

Table 1 shows the comparison of different models. The accuracy of validation data is reduced with VGGNet compared to other models. Though AlexNet shows better performance compared to VGGNet, it consists of 23.3 M parameters. The proposed model shows improved performance compared to other models with only 7.2 M parameters.

V. Conclusion

In this paper, we have compared various models such as LeNet5, AlexNet, VGGNet and the proposed model. Instead of raw images, we collected thermal images from a FLIR camera and trained them with different deep learning models. The comparison shows the classification performance of different pig treatments. The different models are evaluated with metrics such as accuracy, F1 score, precision and recall. Compared to other deep learning models, the modified CNN shows better classification performance on thermal images. The future work is concentrated on building a better model considering real-time scenarios.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (2021R1A6A1A03043144).

References

[1] G. Algan and I. Ulusoy, "Image Classification with Deep Learning in the Presence of Noisy Labels: A Survey," ArXiv, vol. abs/1912.05170, 2021.

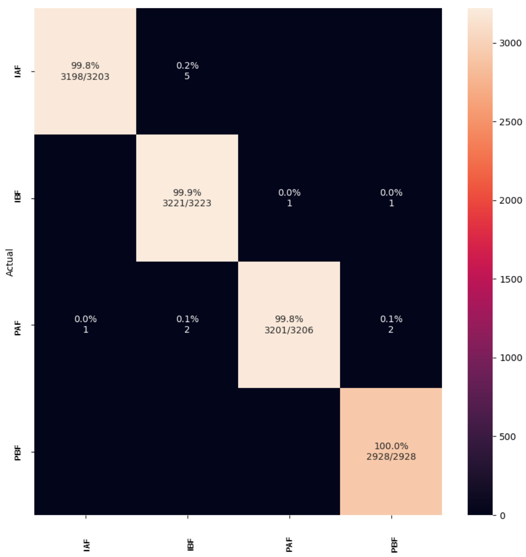
[2] Manohar, N, S. Chalathadka Subrahmanya, R. K. Bharathi, Sharath Kumar Y. H and Hemantha Kumar G. "Recognition and classification of animals based on texture features through parallel computing." 2016 Second International Conference on Cognitive Computing and Information Processing (CCIP) (2016): 1-5.

[3] Y. LeCun, "LeNet-5, convolutional neural networks," URL: http://yann. lecun. com/exdb/lenet, vol. 20, p. 5, 2015.

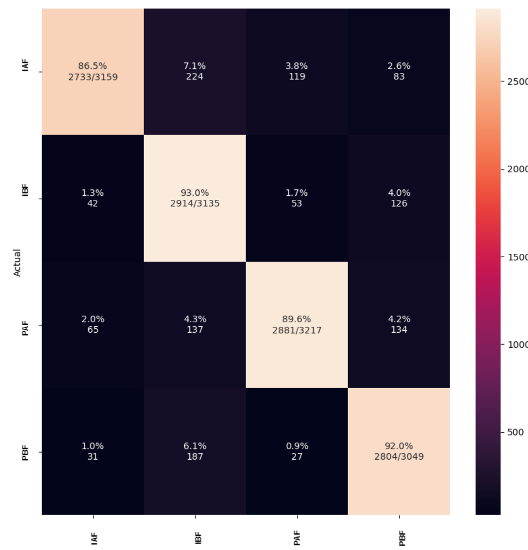
[4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Communications of the ACM, vol. 60, pp. 84 - 90, 2012.

[5] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

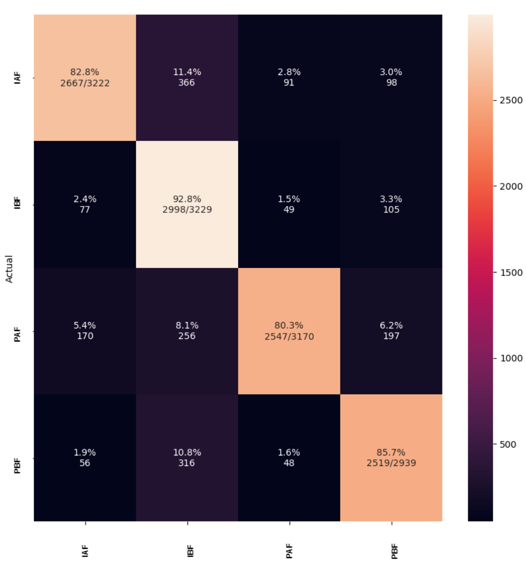
[6] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.



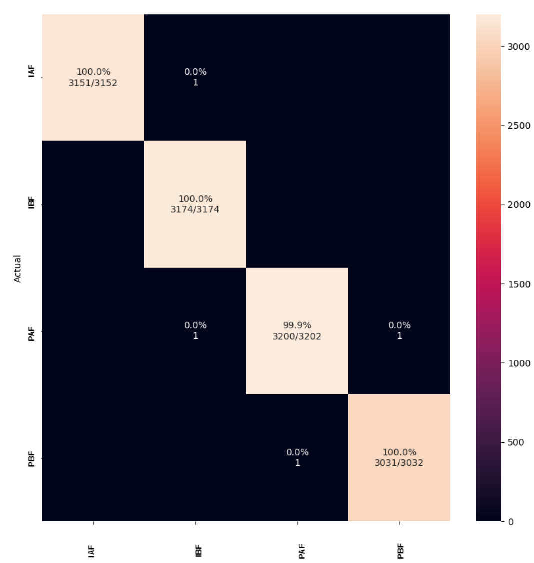
(a)



(b)



(c)



(d)

Fig. 4. The confusion matrix results. (a) LeNet5. (b) AlexNet. (c) VGGNet. (d) proposed model.