Performance analysis of mAlexnet by training option and activation function tuning on parking images

S Rahman¹, M Ramli²*, F Arnia³, R Muharar³, A Sembiring⁴

¹Doctoral Program, School of Engineering, Universitas Syiah Kuala, Banda Aceh
²Department of Mathematics, Universitas Syiah Kuala, Banda Aceh
³Department of Electrical and Computer Engineering, Universitas Syiah Kuala, Banda Aceh
⁴Department of Informatics Technology, Universitas Harapan Medan

*E-mail: marwan.math@unsyiah.ac.id

Abstract. Information on the availability of parking spaces is one of the main needs in urban areas. This information can minimize the impact of vehicle growth, including exhaust gas emissions, traffic jams, and fuel use. In general, the detection of parking space availability can be done in two ways, namely the sensor system and computer vision. Computer vision monitoring of parking spaces is more promising for use in the future. A single-camera can monitor multiple parking spaces by making use of computer vision. CNN is a technique for parking space classification. CNN's pre-trained dedicated to parking space classification is mAlexnet. mAlexnet can classify parking spaces well, but not perfect. In this paper, we try to observe and improve the performance of mAlexnet. We tried training options and activation function tuning. From the results of testing the combination of SGDM training option and the LeakyReLU activation function, the performance of mAlexnet improves.

1. Introduction

In recent years, information on the availability of parking spaces has been urgently needed [1]. The high growth in the number of cars raises several problems that must be resolved immediately. With the availability of information on parking spaces, it can overcome traffic congestion problems caused by drivers looking for the nearest parking location. Information on the availability of parking spaces can also save fuel consumption by vehicles. Fuel is a world problem that must be saved. In addition, by not moving around the car looking for a parking space will minimize the impact of vehicle exhaust emissions [2,3].

The space availability information is the result of real-time information extraction from parking spaces. Various ways were developed to convey this information, including by using sensor systems and smart cameras. The use of ultrasonic sensors can solve this problem, but in its implementation, it requires expensive installation and maintenance costs [4]. Using a smart camera or computer vision is a way to hope for the future. Smart cameras are able to recognize parking spaces simultaneously with a single camera. However, in the implementation of the smart camera, there are more challenges such as changes in light in parking spaces, the presence of obstacles such as trees and poles, and the use of a large amount of computing.

Extraction of information on the availability of parking spaces using cameras has been developed, in general, it is divided into two, namely by marking the parking slot and detecting directly without a
marker. Recognize the mark given on the parking slot [5,6], if there is, the slot is still free and if the mark is not detected then it is busy. This method is the same as using a sensor, it needs installation and maintenance so that the mark remains visible on the camera. The most promising step is how to recognize parking spaces without markers. Convolutional neural network (CNN) is a technique in machine learning that can extract features and classify data [4,7]. Parking space classification is very suitable using CNN to distinguish between busy and free spaces. CNNs specifically developed for parking space classification are mLenet and mAlexnet which are Lenet and Alexnet mini networks [4,7]. Testing of parking space images recommends mAlexnet because it is more accurate in parking space classification. In this study, we analyzed the performance of mAlexnet if the training options and their activation functions were changed. Changing training options and activation functions is expected to improve the performance of mAlexnet.

2. Convolutional Neural Network (CNN)
CNN is part of Deep Learning. CNN is a multilayer perceptron which is designed to process image data. CNN is often used to extract features and classify objects in an image. The way CNN works is to extract the features on the raw image at the convolutional network layer and the activation function followed by pooling. The features of this layer are sent to the next layer and so on. The last layer is assigned to classify features, features are classified as fully connected [8]. An example of the CNN architecture is shown in Figure 1.

![CNN Architecture](image)

**Figure 1.** CNN Architecture.

2.1 Training option CNN
Training options that are often used in deep learning are sgdm (stochastic gradient descent with momentum), adam (adaptive moment estimation), and RMSProp (root mean square propagation). Each training option both SGDM, Adam and RMSProp has its advantages [9,10]. In this study, a comparison of the performance of each training option was carried out. Training options are best used to improve the performance of mAlexnet. After the training options are determined, the next step is to compare the use of activation functions on the CNN architecture.

2.2 Activation function CNN
Artificial neural networks generally have an activation function. The activation function is useful for directing the input to the desired output. CNN is part of deep learning which has layer end activation function. Activation functions that can be used include ReLU, ELU, LeakyReLU and tanh [11–13]. After we found the best training option, we tried each activation function. The combination of training options and activation functions is expected to improve the performance of mAlexnet. The following activation function graphs are shown in Figure 2.
2.3 mAlexnet
mAlexnet is CNN pre-trained specifically for the needs of parking space classification. MAlexnet architecture has three convolutional layers namely conv-1, conv-2 and conv-3. Each layer followed by ReLu, LRN and maxpolling which are useful for feature extraction. In the feature classification section, there are two fully connected ones given ReLu between them [7]. MAlexnet architecture is shown in Figure 3.

3. Result and discussion
Comparative testing of the use of training options and activation functions using the CNR-Park dataset which was taken randomly as many as 4000 images measuring 150 x 150 pixels. The dataset is labeled busy and free, each with 2000 images. The busy label is for images of parking spaces filled with cars and free for parking spaces that are not filled with cars. Dataset with a size of 150 x 150 pixels is shown in Figure 4.
3.1 Training options tuning
The training model in neural networks is one of the determining factors in the success of a neural network architecture in object classification. In CNN deep learning there are several training options that are often used, namely SGDM, Adam, and RMSProp. We tried to set training options to measure the performance of mAlexnet, which is a CNN pre-trained specifically for parking space classification. In this experiment, we did not change the architecture of mAlexnet. The experimental results show that the accuracy rate of validation using SGDM is 97.77%, ADAM is 97.45%, and RMSProp is 96.73%. From these results, SGDM is better for mAlexnet training for parking space classification. Figure 5 shows the validation graph in each iteration.

3.2 Activation functions tuning
From the previous test, we used SGDM in the mAlexnet training process. We experimented using several activation functions. The activation functions used are ReLU, ELU, LeakyReLU, and tanh. From the experimental results, we note that the performance of ReLU is 97.77%, ELU is 97.30%, LeakyReLU is 98.34%, and tanh is 93.04%. Overall, the results of each experiment using the activation function are shown in Table 1.
Table 1. The performance of the activation function.

| Activation Function | Iteration 1 | Iteration 30 | Iteration 60 | Iteration 90 | Iteration 120 | Iteration 150 | Iteration 175 | Final |
|---------------------|-------------|--------------|--------------|--------------|--------------|--------------|--------------|-------|
| ReLU                | 78.79%      | 92.36%       | 95.58%       | 96.57%       | 96.00%       | 97.56%       | 97.51%       | 97.77% |
| ELU                 | 78.85%      | 91.37%       | 94.44%       | 96.31%       | 97.40%       | 97.30%       | 97.09%       | 97.30% |
| Leaky ReLU          | 76.72%      | 94.85%       | 95.95%       | 97.25%       | 97.19%       | 97.92%       | 98.13%       | 98.34% |
| Tanh                | 74.17%      | 86.85%       | 90.75%       | 92.67%       | 94.13%       | 92.52%       | 92.83%       | 93.04% |

The use of the LeakyReLU activation function gets better results than other activation functions as shown in Table 1. LeakyReLU is the only one that gets stable training results and increases with each iteration. The following is a graph of the performance of the activation function shown in Figure 6.

Figure 6. The performance of the activation function in improving validation accuracy.

4. Conclusion
Experiments were carried out using the CNRPark dataset which was taken 4000 images randomly. The test was conducted to review the performance of training options and activation functions. From the observation that the use of SGDM training options is better than the others. Then we tried to change the activation function trained with sgdm. the LeakyReLU activation function has better performance than other activation functions. It can be concluded that the use of the LeakyReLU activation function and training with SGDM can improve the performance of mAlexnet.

5. Acknowledgments
This research is funded by Professor Research Grant, Universitas Syiah Kuala 2020, with contract number 3/UN.11.2.1/PT 01.03/PNBP/2020.

References
[1] Lin T, Rivano H and Le Mouël F 2017 IEEE Transactions on Intelligent Transport. Syst., 18 3229–3253.
[2] Sukarno I, Matsumoto H and Susanti L 2016 Future Cities Environ., 2 1-11.
[3] Kakouei A, Vatani A and Idris A K B 2012 *J. Environ. Health Sci. Eng.*, 9 1-5.

[4] Amato G, Carrara F, Falchi F, Gennaro C, Meghini C and Vairo C 2017 *Expert Syst. Applicat.*, 72 327–334.

[5] Delibaltov D, Wu W, Loce R P and Bernal E A 2013 *16th International IEEE Conference on Intelligent Transport. Syst. (ITSC 2013)*, 2387–92.

[6] Rahman S, Ramli M, Arnia F, Muharar R, Luthfi M and Sundari S 2020 *J. Physics: Conf. Ser.*, 1566 12092.

[7] Amato G, Carrara F, Falchi F, Gennaro C, and Vairo C 2016 *2016 IEEE Symp. Comput. Communic. (ISCC)*, 1212–17.

[8] Yamashita R, Nishio M, Do R K G, and Togashi K 2018 *Insights Imag.*, 9 611–629.

[9] Dogo E M, Afolabi O J, Nwulu N I, Twala B and Aigbavboa C O 2018 *Int. Conf. on Comput. Tech. Electr. Mech. Syst. (CTEMS)*, 92–99.

[10] Jiang X, Hu B, Satapathy C S, Wang S H and Zhang Y D 2020 *Sci. Program.*, 2020 1-36.

[11] Dewa C K 2018 *Procedia Computer Sci.*, 144 124–132.

[12] Zhang Y, Hua Q, Xu D, Li H, Bu Y and Zhao P 2019 *IGARSS 2019-2019 IEEE Int. Geosci. Remote Sensing Symp.*, 10023–26.

[13] Dubey A K and Jain V 2019 *Appl. Comput. Auto. Wireless Syst. Electric. Eng.*, 553 873–880