

Fruit tree disease classification system using generative adversarial networks

Changsu Kim¹, Hyesoo Lee², Hoekyung Jung³

^{1,3}Department of Computer Engineering, PaiChai University, Daejeon, Republic of Korea

²Korea Institute of Science and Technology Information, Daejeon, Republic of Korea

Article Info

Article history:

Received Jul 3, 2020

Revised Dec 10, 2020

Accepted Dec 26, 2020

Keywords:

Classification system

GAN

Machine learning

Smart farm

ABSTRACT

Smart farm refers to a farm that can remotely and automatically maintain proper growth and management of crops and livestock by integrating technology with agriculture. Currently, smart farms are concentrated in the field of smart horticulture, and although spreading research is being conducted in limited spaces. In addition, it is difficult to obtain a sufficient amount of data to be used for learning, and there is a problem that data imbalance occurs because it is difficult to obtain a similar amount for each class. In this paper, we propose a method to amplify a small amount of data and to solve the problems of imbalance data by using a feature that can learn to mimic the data of a generative adversarial network. The proposed method can create dataset of various crops and also show high hit rate. Dataset generated from crops would be used to solve problems of data imbalance by learning.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Hoekyung Jung

Department of Computer Engineering

Paichai University

155-40 Baejae-ro, Seogu, DaeJeon, Republic of Korea

Email: hkjung@pcu.ac.kr

1. INTRODUCTION

Smart farming is evolving with the application of the 4th industrial revolution technology. It refers to a technology that enables remote control as well as services such as automation and intelligence by integrating new technologies such as ICT, IoT, big data, cloud, and AI into the growth and environment of crops or livestock [1-3]. The need for smart farms is deeply linked to global climate change and food shortages as populations grow. Food demand is on the rise due to an increase in the world's population, but there is a growing shortage of people who will grow because of the shrinking and aging population. In other words, as the population increases, urbanization reduces the area of crop cultivation, and farmers on the production site are aging. Even in Korea, as of 2017, the average age of farmers' chief executives is 67, and this average age is only increasing with time. As a result, smart farms are becoming increasingly important. However, smart farms are concentrated in the field of smart horticulture, and although they are recently spreading, research is being conducted in limited spaces, such as facilities. In addition, it is difficult to obtain a sufficient amount of data to be used for learning, and there is a problem that data imbalance occurs because it is difficult to obtain a similar amount for each class [2].

In this paper, we propose a method to amplify the amount of data through generative neural networks and to solve data imbalance between classes by generating data of various classes. The proposed technique improves the speed by reducing the amount of computation by preprocessing the data and highlights the features to help learning. The preprocessed data uses a generative neural network to generate

new data and filters out the low quality data through filtering for data integrity. In case of using the proposed technique, it is possible to amplify a small amount of data to generate a sufficient amount of data for learning, and to apply it to an existing system as it is. Based on these advantages, the proposed method can be used to compensate for the limitations of the recent smart farm.

2. RELATED RESEARCH

In this chapter, we analyze the region of interest extraction method and the generative neural network used in preprocessing and the problems that can be caused by data imbalance.

2.1. Generative adversarial networks (GAN)

GAN is a paper written by Ian Goodfellow in 2014 and various studies are being conducted based on it [4-7]. The generative hostile neural network learns and produces the result through the competition of two neural networks as the name of the hostile neural network. In (1) represents the formula of the generative host neural network

$$\text{Min}_G \text{Max}_D V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{x \sim p_x(z)} [\log(1 - D(G(z)))] \quad (1)$$

In the above formula, $x \sim p_{data}(x)$ represents an image generated by the probability distribution of the actual image, and $x \sim p_x(z)$ represents an image generated using noise. $D(x)$ is a discriminator, with a value between 0 and 1 indicating the probability that the image is real. In the case of $D(G(z))$, the image generated by the creator is distinguished through the discriminator, which also has a probability between 0 and 1 that the image is real. To maximize the above equation, the value of $D(G(z))$ should be close to 1 and the value of $D(x)$ should be close to 0. In this way, constructors and discriminators learn by solving the Minmax problem.

2.2. Crop region of interest (RoI)

The region of interest means an area of interest on the image. When the object is detected or detected by processing an image [8-10], the detected area may be referred to as a region of interest [11]. There are three main reasons for specifying a Region of Interest. The first is to remove unnecessary images of the area around the object, and the second is to reduce the amount of computation and resources. Finally, the accuracy of learning can be improved. By eliminating unnecessary information in advance, only necessary parts are used for learning, which increases accuracy.

The most representative ones in crop region of interest may be referred to as RGB separation and contour extraction. In the case of RGB separation, a region of interest can be obtained by specifying a range of channels and outputting pixel values within a range by using properties having different values for each channel of an image. Contour extraction means extracting a point where the brightness of the image changes from a low value to a high value or vice versa. Contour detection is a technique to find the pixels corresponding to the contour, and calculates the slope based on the calculation of the partial differential operator [12-13].

2.3. Data imbalance

If there is a large difference in the amount of data that each class has in the data, it is said that there is a class imbalance. Most of the data in the real world has this data imbalance problem, and it rebalances the classes by sampling again before training the model. Undersampling is selecting parts of a lot of data, such as the image on the left, and aligning them toward the smaller data. Undersampling can reduce execution time by reducing the size of the data set, but it is also possible that the useful data is not extracted or biased to one side [14-17]. Oversampling involves copying less data and fitting it toward more data. In the case of oversampling, the data set has more oversampling, so learning is better. In other words, it performs better than undersampling. However, there is a possibility of overfitting when the same data is extracted repeatedly [18-19].

2.4. Problems and solutions

Problems between smart farms and data include: It is difficult to obtain enough data to be used for learning, and even the provided data has an adverse effect on learning due to unbalanced data per class [20-23]. Based on these problems, this paper proposes a method of amplifying the amount of data through a generative antagonistic neural network and generating data of various classes to solve the data imbalance between classes. Using a generative adversarial network not only amplifies a small amount of data to generate enough data for learning, but also has the advantage of using oversampling with less redundancy of the data used.

3. SYSTEM DESIGN

3.1. Overall system design

The proposed system can be divided into the data processing stage that manages data preprocessing and postprocessing, and the network stage that defines the model for generating data and classifying images. Figure 1 shows this structure. In the data processing stage, it is divided into preprocessing and postprocessing. In the preprocessing step, image size adjustment, contrast adjustment, and region of interest extraction are performed. Image resizing reduces the size of the image, lowering the memory footprint of the system and reducing computation. The size of the image is changed to the n power of 2 to adjust the size of the image so that it does not give a real value when training the model. Region of interest extraction reduces the amount of computation by reducing data outside the region of interest. Contrast makes the characteristics of the data stand out by making the contrast and hue clearer. Post-processing involves a data filter that either stores the results or filters the data. The data filter filters out low quality data from images generated by genetic adversarial networks to ensure the quality of the generated image. Data storage is responsible for sending data to the classification system when the user gathers more than a predetermined amount.

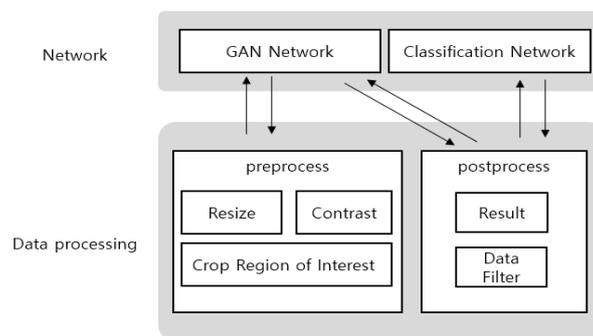


Figure 1. System architecture diagram

At the network level, the generator is divided into a generator, a discriminator network, and a classification network that divides the image. Genetic adversarial networks take preprocessed data as input and generate new images based on it. In the classification system, the generated images are learned and the actual images are determined based on the generated images. In the image discrimination step, the validity of the proposed system is verified by comparing the proposed system with the existing system.

Figure 2 shows the flow chart. The flow of the proposed system proceeds from the size control of the input data. By adjusting the size, it plays a role in speeding up the future work. The resized image is then extracted from the region of interest. This plays a role of improving the speed of learning by reducing the amount of computation later. The final preprocessing is to give the image a contrast. The image is then generated through genetic adversarial networks. The image you create is based on an uninfected image, which allows you to create an image similar to the diseased one from the original. The generated image is saved or discarded according to its quality and is repeated until the image reaches a certain number. When more than a certain number of images are created, the generation stops and the classification model is used to train the classification model.

3.2. Crop region of interest design

The region of interest was extracted using some Sobel mask method based on RGB extraction. Figure 3 shows the flow chart. First, the image is searched and the average value is calculated. The average value is calculated using a Sobel mask. Find the average and start checking every pixel. If the average value is not exceeded, the next pixel is checked. If the average value is exceeded, the next pixel is checked to see if the pixel is stored. If not, the pixel is stored and moved to the pixel. At this time, the inspection direction of the pixel proceeds constantly. If the pixel is stored, the search ends and a mask is created from the stored pixels to extract the point.

3.4. GAN design

The designed generative adversarial networks is designed based on DCGAN. The constructor uses the existing model of generative adversarial networks, but receives $64 \times 64 \times 3$ variables which are the size of the image to be generated as input values [24-25]. Figure 4 shows the model of the discriminator network.

The discriminator was modified according to the image given as input, which has a certain form of leaf, the disease occurring in the leaf, and the size of the image is 64x64x3. The size of the convolutional layer was changed from 5x5 to 3x3 to reduce the amount of computation and save memory. In addition, the activation function is designed to generate images of various classes using softmax rather than sigmoid.

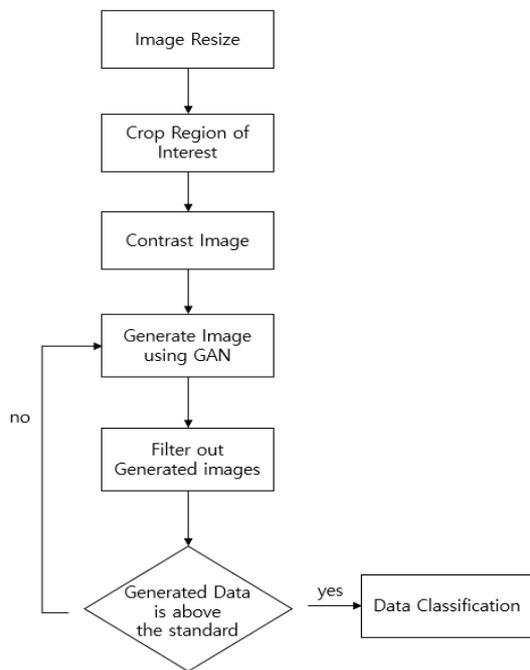


Figure 2. System flow chart

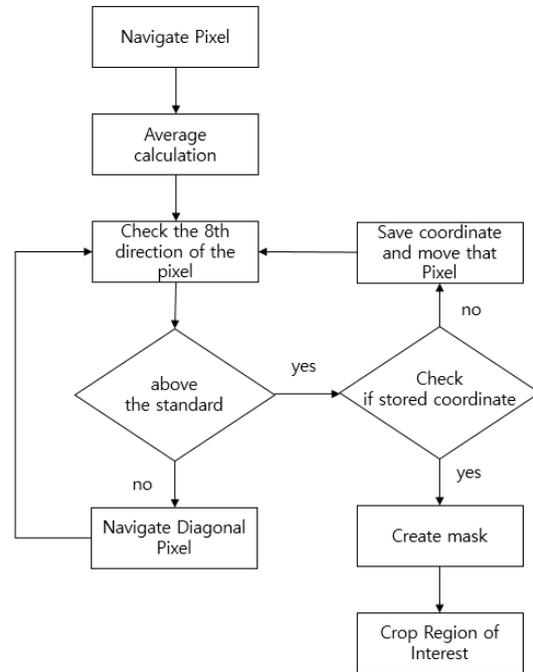


Figure 3. Crop region of interest flow chart

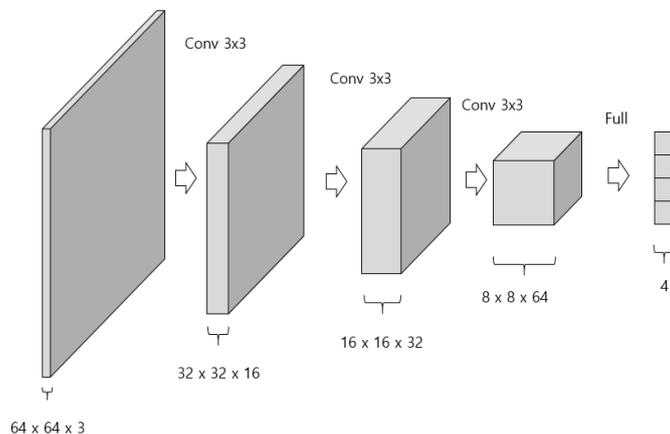


Figure 4. Structure of discriminator

3.5. Classification system design

The designed model was designed based on CNN. First, in the input phase, the image is changed to code and the data for each class is adjusted regularly. There are 5 convolutional layers, which change according to the size of the input image. The smaller the size of the image, the smaller the size of the feature, so there is no need to build a deeper convolutional layer. Therefore, the image in this paper is 64x64x3, so 5 convolutional layers and 1 fully-connected layer are used. The convolutional layer is a filter that is characterized by a 3x3 filter and a Max Pooling layer that ignores minor changes. Max Pooling creates a small output image by extracting only the main values from the filtered image. The Max Pooling size is used to reduce the size of the image by half every 2x2. In the fully-connected layer, the image which has

repeatedly passed the convolutional layer is extracted only from the feature, and it changes the image into one dimension and distinguishes the image in order to transfer the extracted feature to the precombined layer.

4. SYSTEM IMPLEMENTATION

To implement the system, the Windows 10 Pro operating system was installed on a computer equipped with an i5-6500 processor, 16 G memory, and a GeForce GTX 1060 graphics card. We also installed CUDA 10.0 to use the GPU and implemented the system using Python3.5 in Jupyter Notebook.

4.1. Implement preprocess

The crop region of Interest is responsible for finding key parts of the image before the generative hostile neural network proceeds. First, we search the entire pixel to find the average value, and then search through the pixel values of the image as designed. After searching, it prints and saves only the image inside the newly created mask. Figure 5 shows the results of the region of interest.

In Figure 5, from left to right, the original, crop of interest extraction 1 time and interest region extraction 2 times. As can be seen from the Figure 5, it is necessary to proceed two times in order to see the extraction of the region of interest. This occurs when the average value is calculated. The first time, the region of interest is extracted by calculating the average value of the background, the shadow, and the leaves. Get it. In addition, the shadow is caused in the picture, but this problem occurs when the focus is blurred at the edge of the leaf. Therefore, in this paper, we implemented the region of interest twice using the crop region of interest to concentrate the region of interest. Figure 6 shows the preprocessed image.

In Figure 6, From the left, the original, the extracted region of interest, and the contrasted image. In the last image, you can see that the image of the affected area is more intense in color. The image is then generated through a generative hostile neural network. The image you create is based on an uninfected image, which allows you to create an image similar to the diseased one from the original. The generated image is saved or discarded according to its quality and is repeated until the image reaches a certain number. When more than a certain number of images are created, the generation is stopped and the classification model is used to train the classification model.



Figure 5. Cropped region of interest images

Figure 6. Preprocessed images

4.2. Implement GAN

Genetic adversarial networks play a role in learning preprocessed data and generating new images. The generator network has four convolution layers to normalize the output and activation values via batch normalization. The discriminator is implemented with three convolutional layers and one fully-connected layer, and sigmoid is used as an activation function to discriminate images of various classes. Figure 7 is an image created by applying the disease called Black Rot to the blueberry leaves. In Figure 7, the black rot leaves are well formed. In addition, by limiting the region of interest during the pretreatment process it can be seen that the location of the disease is properly seated on the leaves.



Figure 7. Generated blueberry black rot images

4.3. Implement proposed system

We implemented and compared the designed system with the existing system. Figure 8 shows the classification per epoch when sorting without preprocessing (system 1) and oversampling by increasing data through rotation (system 2). This graph shows the accuracy comparison.

The axis, epoch, represents 10 times. In Figures 8 and 9, system 1 took a long time to maximize accuracy due to lack of data. System 2 was maximized in the epoch similar to the proposed system, but the accuracy was confirmed to be low. In contrast, the proposed system reached its maximum at a faster rate than system 1 and showed an average 2% accuracy improvement over system 2.

Table 1 shows the comparison of accuracy when classified using the test set. Experimental results show that system 1 has a 77.99% probability of lacking a data set, while system 2 has a 91.14% probability that some data sets are small. On the other hand, the proposed system increased the amount of data used for learning using generative antagonistic neural networks, and provided the same amount of data for each class, showing high accuracy of 98.17%.

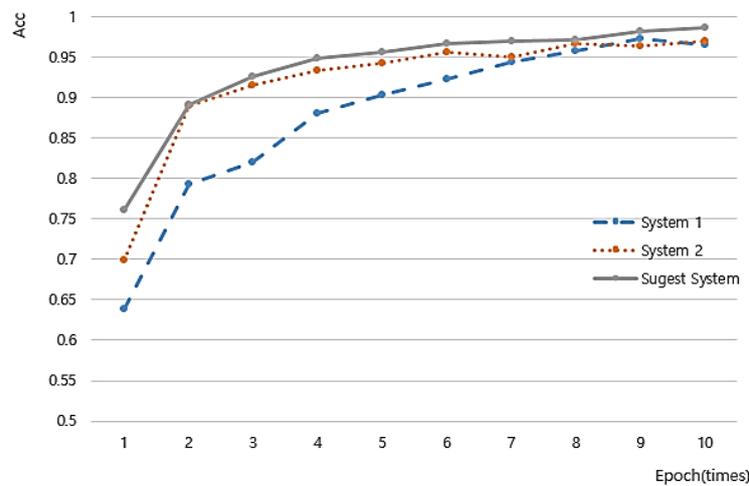


Figure 8. Graph comparing accuracy

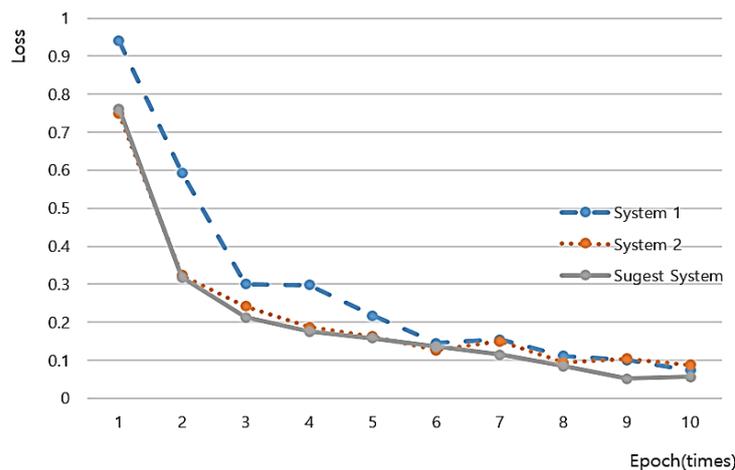


Figure 9. Generated blueberry black rot images

System	Acc1	Acc2	Acc3	Acc4	Acc5	Avg Acc
System1	76.85	78.33	78.12	77.30	79.39	77.99
System2	89.67	91.33	90.65	92.47	91.61	91.14
Proposed system	98.08	97.29	98.78	98.54	98.17	98.17

5. CONCLUSION

Smart farm refers to a farm that can maintain and manage crop and livestock growth conditions appropriately remotely and automatically by using ICT in agriculture. In addition, it was possible to improve the productivity and quality of agricultural products by creating an optimal growth environment based on data on crop growth information and environmental information. However, crops to which smart farms are currently applied are limited to horticulture or house crops, and there is a problem in that a large cost is required to introduce smart farms. In addition, the provided data may show a big difference from the actual data, and since the amount of data is different, there is a problem that the data imbalance problem must be solved in order to learn.

In order to solve this problem, This paper proposes a method that learns the data through genetic adversarial networks and generates data similar to the actual data by using the input data. In order to reduce the resources used when learning data, this system proposed a preprocessing method to reduce the amount of computation and speed it up. We then used genetic adversarial networks to increase the amount of data used for learning and to resolve data imbalances by class. This enabled them to generate data for smart farms across a wide range of crops, increasing the accuracy of learning by 98 percent. Future research should generate data that can be used not only in the classification system used in this paper but also in logistics and automation systems. As a result, it is considered that various smart farms can be constructed.

ACKNOWLEDGEMENTS

This work was supported by the research grant of PaiChar University in 2021.

REFERENCES

- [1] N. Suma, S. R. Samson, S. Saranya, G. Shanmugapriya, and R. Subhashri, "IOT Based Smart Agriculture Monitoring System," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 5, pp. 99-109, Feb. 2017.
- [2] C. J. Kim, J. H. Jeong, C. W. Jo, J. K. Yoo, "A Performance Evaluation Analysis of Product Recommendation Techniques," *Journal of Knowledge Information Technology and System*, vol. 14, no. 5, pp. 515-525, Oct. 2019.
- [3] S. H. Lee and J. Y. Bae, "Prediction Crop Production for Agricultural Consultation Service," *Jorunal of Information and Communication Convergence Engineering*, vol. 17, no. 1, pp. 8-13, Mar. 2019.
- [4] H. S. Kim and H. S. Lee, "Generative Adversarial Networks based Data Generation Framework for Overcoming Imbalanced Manufacturing Process Data," *Journal of Korean Institute of Intelligent Systems*, vol. 29, no. 1, pp. 1-8, Feb. 2019.
- [5] I. J. Goodfellow, et al., "Generative Adversarial Networks," *Cornell University*, 2016.
- [6] P. H. Huynh, V. H. Nguyen, and T. N. Do, "Enhancing Gene Expression Classification of Support Vector Machines with Generative Adversarial Networks," *Jorunal of Information and Communication Convergence Engineering*, vol. 17, no. 1, pp. 14-20, Mar. 2019.
- [7] W. Wan and H. J. Lee, "Perceptual Generative Adversarial Network for Single Image De-Snowing," *KTSDE*, vol. 8, no. 10, pp. 403-410, Aug. 2019.
- [8] G. Y. Heo, "ROI Extraction for Automatic Placard Recognition," *Journal of the Korea Institute of Information and Communication Engineering*, vol. 23, no. 4, pp. 374-380, Apr. 2019.
- [9] K. J. Chae, Y. R. Lee, Y. J. Cho, and J. H. Park, "Development of a Gangwon Province Forest Fire Prediction Model using Machine Learning and Sampling," *The Koea Journal of Bigdata*, vol. 3, no. 2, pp. 71-78, Feb. 2018, doi: 10.36498/kbigdt.2018.3.2.71.
- [10] G. Wang and S. Y. Shin, "An Improved Text Classification Method for Sentiment Classification," *Jorunal of Information and Communication Convergence Engineering*, vol. 17, no. 1, pp. 41-48, Mar. 2019.
- [11] Z. Zhu, D. Liang, S. Zhang, X. Huang, B. Li, and S. Hu, "Traffic-Sign Detection and Classification in the Wild," *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2110-2118, doi: 10.1109/CVPR.2016.232.
- [12] J. H. Shin, K. H. Chung, and K. H. Choi, "Destructive Test of a BLDC Motor Controller Utilizing a Modified Classification Tree Method," *KTSDE*, vol. 3, no. 6, pp. 201-214, Mar. 2014.
- [13] S. T. Oh and B. H. Jun, "Contour Extraction Method using p-Snake with Prototype Energy," *Journal of the Institute of Electronics and Information Engineers*, vol. 51, no. 4, pp. 101-109, Apr. 2014.
- [14] J. H. Oh and J. G. Paik, "Clustering-based Undersampling for Imbalanced Data Classification," *Journal of Korean Institute of Industrial Engineers*, pp. 1910-1916, Nov. 2017.
- [15] S. R. Saluti, "A survey of big data and machine learning," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 1, pp. 575-580, Feb. 2020.
- [16] H. J. Shin and C. H. Oh, "Machine Learning based on Approach for Classification of Abnormal Data in Shop-floor," *Journal of the Korea Institute of Information and Communication Engineering*, vol. 21, no.11, pp. 2037-2042, Nov. 2017.

- [17] S. T. Nam, S. Y. Shin and C. Y. Jin, "A Reconstruction of Classification for Iris Species Using Euclidean Distance Based on a Machine Learning," *Journal of the Korea Institute of Information and Communication Engineering*, vol. 24, no. 2, pp. 225-230, Feb. 2020.
- [18] D. Oh, B. T. Oh and J. Shin, "Up-sampling method of depth map using weighted joint bilateral filter," *The Journal of Korean Institute of Communications and Information Sciences*, vol. 40, no. 6, pp. 1175-1184, Jun. 2015.
- [19] S. Aich, S. Chakraborty and H. C. Kim, "Convolutional neural network-based model for web-based text classification," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 6, pp. 5185-5191, Dec. 2019.
- [20] X. T. Yang, J. S. Lee and H. K. Jung, "Fault Diagnosis Management Model using Machine Learning," *Jorunal of Information and Communication Convergence Engineering*, vol. 17, no. 2, pp. 128-134, Jun. 2019.
- [21] H. K. Kim, S. O. Lee, and H. K. Jung, "Human activity recognition by using convolutional neural network," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 6, pp. 5270-5276, Dec. 2019, doi: 10.11591/ijece.v9i6.pp5270-5276.
- [22] J. K. Rhee, "Prediction for Periodontal Disease using Gene Expression Profile Data based on Machine Learning," *Journal of the Korea Institute of Information and Communication Engineering*, vol. 23, no. 8, pp. 903-908, Oct. 2019.
- [23] S. Khan, K. V. Kirubanand "Comparing machine learning and ensemble learning in the field of football," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 5, pp. 4321-4325, Oct. 2019, doi: 10.11591/ijece.v9i5.pp4321-4325.
- [24] A. Radford, L. Metz and S. Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks," *Cornell University*, 2016.
- [25] S. Aich, S. Chakraborty and H. C. Kim, "Convolutional neural network-based model for web-based text classification," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 6, pp. 5185-5191, Dec. 2019.

BIOGRAPHIES OF AUTHORS



Changsu Kim He received the B.S., M.S., and Ph.D. degrees from the Department of Computer Engineering of Paichai University, Korea, in 1996, 1998, and 2002, respectively. From 2005 to 2012, he worked for the Department of Internet at Chungwoon University as a professor. Since 2013, he has worked in the Department of Computer Engineering at Paichai University, where he now works as a professor. His current research interests include multimedia document architecture modeling, Data mining, and the semantic web.



Hyesoo Lee She received the B. S degree in 2014 and M. S degree in 2016 from the Department of Computer Engineering of Paichai University, Korea. From 2013 to 2014, she completed her internship at National Security Research Institute and she works for Supercomputing Infrastructure Center of Div. of National Supercomputing in Korea Institute of Science and Technology information (KISTI). She is in charge of contract of usage and account management for super computer, and provides technical support to a user of super computer.



Hoekyung Jung He received the M.S. degree in 1987 and Ph. D. degree in 1993 from the Department of Computer Engineering of Kwangwoon University, Korea. From 1994 to 1995, he worked for ETRI as a researcher. Since 1994, he has worked in the Department of Computer Engineering at Paichai University, where he now works as a professor. His current research interests include multimedia document architecture modeling, information processing, embedded system, machine learning, bigdata, and IoT.