International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-2, December 2019

Network-Simulated Generation of Human Faces with Expressions and Orientations for Deep Learning Classification

Kornprom Pikulkaew, Ekkarat Boonchieng, Waraporn Boonchieng, Varin Chouvatut

Abstract*: *Human face recognition is a complex task, and it is important as it can be applied to assist people worldwide, such as those in the medical field or security. For example, human faces can be used for detecting pain or emotion. Nevertheless, a drawback of deep learning methods that need a lot of data to process is key. In this study, a deep-learning-based technique, which is used to classify, that generates a synthetic image of the facial expression and orientation by utilizing the Wasserstein generative adversarial network (WGAN) is presented. The WGAN can improve the performance of the deep learning method. The proposed system certainly generates images with a small number of datasets compared to the large datasets. This research aims to solve the problem of deep learning by increasing the accuracy of the system. The generated output coincides with the real image dataset. The application using ResNet-50 and RetinaNet as a pre-model for the prediction and detection of the human faces revealed a rapid prediction time and accuracy during the assessment test.

Keywords*: *Classification, Deep Learning, Facial Expression, Generative Adversarial Networks, Human Faces, Orientation.

I. INTRODUCTION

In the past, when people felt ill or felt pain, they needed to go to the hospital. The doctor diagnosed patients based on their symptoms. As a result, considerable money is required as payment for diagnosis. Following this method, the self-report method for pain became useful to medical staff. Consequently, it becomes standard to measure pain levels. However, it still exhibits drawbacks. For example, the concept of pain between the people and clinicians is different [1]. Subsequently, the observer rating method apparently improves the self-report. On the other hand, the observer rating cannot be used in some cases, such as cases that require long-term monitoring, such as in the ICU. Moreover, the self-report and observer measurements are extremely subjective. Nevertheless, patients may suffer from various types of pain. Asking patients to describe their pain may not be sufficient to guide medical staff to suitable interventions for specific types of pain. Currently, several techniques can be employed to assess pain, such as pathological examination, neurological examination, or imaging technologies, but they can be invasive or expensive for patients. Thus, for screening of pain [2], some researchers have proposed facial expression as a potential solution for pain detection. The most standard database used to measure pain algorithms is the UNBC-McMaster shoulder pain expression archive. Human face data have been used to classify and recognize pain in this system, which resembles an array of faces that detects moving faces and posture of humans. Patients are shown these images when they receive diagnosis. In the past [3], artificial neural networks used for image detection or security systems can achieve that goal in several years, and several enhancements in different duties. Deep learning is an algorithm that is an improvement from the artificial neural network. Nevertheless, deep learning exhibits disadvantages: It requires a considerable amount of data, and the data must be manually labeled to function correctly. The facial expression of humans is apparently a sensitive, limited resource. For example, it is difficult to predict the shape of the faces in expressions and orientations. Furthermore, the collecting method requires crystal clear pictures. Hence, a generative adversarial network (GAN) model is used to generate image data, which are used for training and testing. GANs constitute one of the most popular topics in deep learning that has been proposed by Goodfellow et al. [4]. Currently, some researchers have employed GAN for performing several things, including flower models or an MNIST database. It functions similar to neural networks, which comprise two deep neural networks: the generator and discriminator [5]. The generator can generate fake images that the discriminator attempts to distinguish with real images [6]. It continues this process until the discriminator cannot classify between the real and generated images. In this study, a method that improves the performance of deep learning for classification and recognition via the generation increased amount of data using GAN is proposed. Fig. 1 shows the structure of the proposed GAN combined with deep learning.

Revised Manuscript Received on December 30, 2019.
* Correspondence Author
Kornprom Pikulkaew, Department of Computer Science, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand
Ekkarat Boonchieng, Center of Excellence in Community Health Informatics, Department of Computer Science, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand
Waraporn Boonchieng, Faculty of Public Health, Chiang Mai University, Chiang Mai, Thailand
Varin Chouvatut*, Department of Computer Science, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand. E-mail: varinchouv@gmail.com

The facial expressions of patients with shoulder pain and several postures are focused. The real images are programmed in GAN during the training process. Next, GAN generates vectors. Finally, GAN results are applied to deep learning for classification and recognition. In the next section, the manner in which the proposed method trains the sequential images of human faces with several orientations is defined.

II. METHODOLOGY

A. GAN structure and implantation

The Wasserstein generative adversarial network (WGAN) structure used by Martin et al. [7] was implemented as it can improve our system, and training GAN is complicated and unstable. WGAN used the Earth-Mover distance approximation to solve the optimization problem. One of the benefits of WGAN is that datasets can be trained until the optimal state. When datasets are trained to achievement, it basically offers loss to the generator that can be trained as any other neural network, implying that an equilibrium between the generator and discriminator is no longer required. Furthermore, WGAN is considerably more robust than another GAN. The TensorFlow-based WGAN source code available in the GITHUB project is used as the basis code [8]. However, some researchers [9] have attempted to utilize Graphics processing unit (GPU) on cloud to run GAN because Central processing unit (CPU) alone is slow for training. GPU is considerably more rapid than CPU for most deep learning computations. Currently, Tesla V100 is the most rapid NVDIA GPU on the market [10]. Google has created cloud tensor processing units (TPUs) that speed up the implementation of linear algebra computations, which is considerably used in machine learning applications. TPUs [11] reduce the time to accuracy when huge, complex neural network models are trained. Models that typically take over weeks to train on other hardware platforms can converge in hours on TPUs. The disadvantage of this method is the cost required to pay for the use of service, but some services are free. For example, Google Colaboratory [12] is free. It is a Jupyter notebook system with a good user experience. It mixes with GitHub and Google Drive. Google Colaboratory permits collaborators to leave comments in the notebook. Google Colaboratory is super-fast to get started with Keras or TensorFlow, but it needs some knowledge for the system set-up. Another popular cloud service is Amazon Web Services (AWS) [13] from Amazon. It is difficult to configure the system in this platform. AWS support users with NVDIA Tesla K80 with 4 CPUs and 61 RAM. AWS is scalable, secure, and reliable. The Floydhub[14] is another cloud service in which the knowledge for the system set-up or configuration is less complex than that required for the alternative set-up on AWS or Google Cloud Platform. On the other hand, every cloud that provides GPU cannot support datasets that are created by us. For example, Floydhub can run only some datasets such as MNIST [15], cat and dog [16], or flower [17]. Thus, our PC is used to generate images. Data processing is carried out using PC with a processor Intel Core i7, CPU 3.40 GHz, 16 GB RAM.

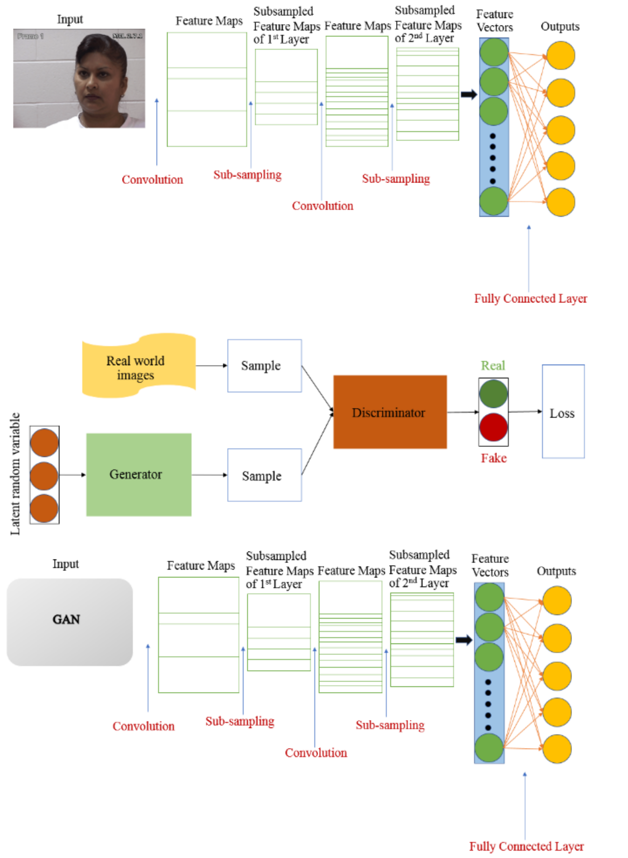


Fig. 1.(top) A standard architecture of a Deep Learning convolutional neural network. (middle) GAN architecture. (bottom) Overall architecture of the proposed model.

B. Image Preparation

The sequential image of the human faces is taken from the UNBC database [18]. Raw images of the facial expression and orientation with a size of 320 × 240 are collected by the two Sony digital video cameras [19] under active and passive conditions. In the active part, patients rotate their shoulders by themselves, while in the passive part, medical staff is required to help movement. Furthermore, under the active condition, researchers focus on the frontal camera orientation, whereas under the passive condition, the camera orientation is set at 70° from the frontal orientation. The database is composed of 129 participants, comprising 63 men and 66 women ($\bar{x}$ age = 42.23 years, SD = 14.48) that are sick from shoulder pain. Moreover, patients are recruited from physiotherapy clinics and by advertisements placed at McMaster University [20]. The University contains almost 50,000 images comprising 25 subjects and 225 directories.

C. Synthetic-image-based Deep Learning method

For predicting human faces, the ResNet50 pre-model by Microsoft Research is used [21], which exhibits a rapid prediction time and high accuracy. ResNet is an abbreviation for residual network. As specified by the name of the network, residual learning is the new terminology presented by this network. Deep convolutional neural networks have led to a series of advances for image classification.

Several other recognition assignments have also significantly profited from deep models. Thus, there is a trend to delve deeper, to answer more difficult tasks, and further improve the classification and recognition accuracy. However, if we go lower, the training of neural networks will become difficult, and the precision will start to saturate and then worsen. Residual learning attempts to solve this problem. Typically, in a deep convolutional neural network, some layers are stacked and trained to perform the task. The network discovers a few level features at the end of its layers. Residual can basically understand as the subtraction of features is learned from the input of that layer. Hence, ResNet50 [22] is a 50-layer residual network. Moreover, it has other variants such as ResNet101 and ResNet152. For the classification of facial expressions, RetinaNet [23] is used, [24] which is a pre-model that exhibits high performance and precision with a long detection time. RetinaNet [25] is a unified network comprising a backbone network and two job-specific subnetworks. The backbone is responsible for calculating a convert feature map over an entire input image. The first subnet works on the classification of the backbone output, while the second subnet works on the convolution bounding box regression.

III. RESULT AND DISCUSSION

A. Synthetic human faces image generation

Synthetic human face images are generated by the proposed GAN architecture. For comparison, four cases are separated: 1. normal datasets with 46 samples, 2. transfer learning for normal datasets, 3. less datasets with 16 samples, and 4. crop only interested areas such as the face and remove the background with 46 samples. Case 4 exhibits good results in various benchmark datasets as well as less time consumption compared to other cases. Moreover, GAN can create a background although it is attempted to be cut during the pre-processing stage, displaying its advantages. However, case 3 affords one of the most disappointing results as realistic images cannot be generated [26]. Fig. 2–Fig. 4 show the input of human faces. WGAN is the type of GAN that improves the stability of learning or reduces the mode collapse [27], [28] which is used for the test. Table 1 summarizes the batch size, epoch, and number of samples. Each combination of these parameters is selected and examined on the basis of previous studies that utilized WGAN [29]. Fig. 5–Fig. 9 show the output of the generated images. The results revealed that case 3 cannot produce any image; only dot data are created even by performing several trials. This problem known as the mode collapse, where only identical images are generated in one test batch, is experimental during the test. Mode collapse [30] is one of the toughest problems to explain in GAN. The whole collapse is not typical, but partial collapse always occurs.

Fig. 8 shows another GAN in which the mode collapses and rotates to another mode when the discriminator catches up. During training, the discriminator is regularly updated to detect adversaries. Hence, it is less possible for the generator to be overfitted. In practice, our understanding of the mode collapse is still limited. However, GAN training is still an exploratory process. Partial collapse is still common.

When an experiment is conducted on cases 1–3, a satisfactory result is not obtained as long time is needed for training, and some images cannot generate high quality. Next, cropping of images of only the interested areas, such as facial areas or facial areas with the hand, is attempted. Fig. 4 shows the example of input for case 4. The result demonstrated that case 4 generates more realistic results than other cases although it uses a smaller number of iterations than other examples. Moreover, case 4 uses less time to generate images during the training phase than every case. Hence, case 4 exhibits the best performance. Cases 1 and 2 generate human faces, but some problems occur. For example, it exhibits a rough overall texture and some brightness areas compared to the real images. Furthermore, it uses more time for generation than other cases. The human faces generated from WGAN are 1024 × 638 pixels, and these faces are added into the deep learning model for the classification and detection.

Fig. 2. Example of input for facial expression and orientation for case 1 and case 2.

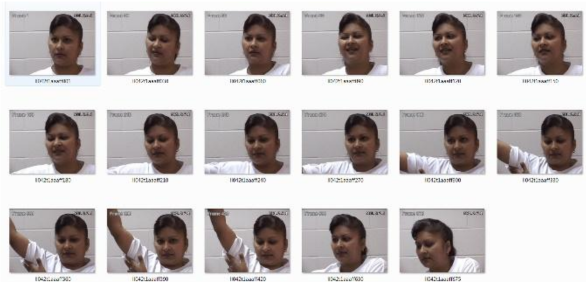


Fig. 3. Example of human faces input for case 3.

Fig. 4.Example of human faces input for case 4.

Retrieval Number: B7491129219/2019©BEIESP
DOI: 10.35940/ijitee.B7491.129219
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

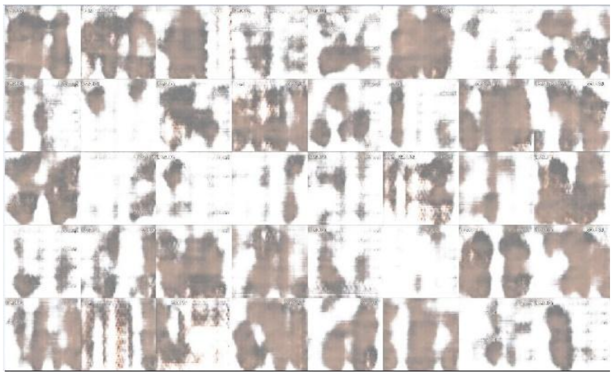


Fig. 5.Example of output from GAN for case 1 at epoch 6500.

Fig. 6. Result from GAN for case 1 at epoch 29800.

Fig. 7. Result from GAN for case 2 at epoch 29800.

Fig. 8. Result from GAN for case 3 at epoch 29800.

Fig. 9.Result from GAN for case 4 at epoch 13800.

Table- I: Training hyper-parameters of the GAN

	Case 1	Case 2	Case 3	Case 4
Sample size	46	46	16	46
Batch size	40	40	10	40
Epoch	29800	29800	29800	13800
Number of days to train	14	14	3	7
Train result	o	o	x	o

o Indicate whether the GAN was successfully trained while x was showed mode collapse.

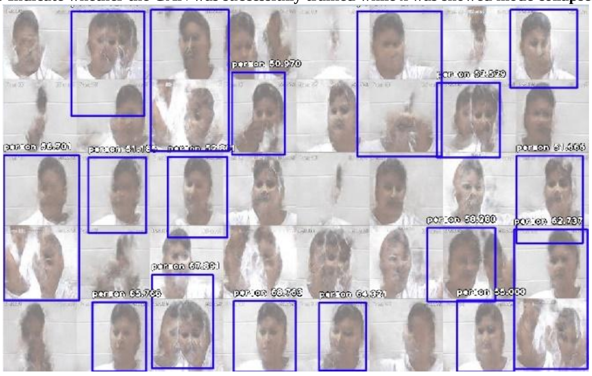


Fig. 10. Example of recognition for case 4.

Table- II: Summary of recognition phase

	Case 1	Case 2	Case 3	Case 4
Number of successful recognitions	11	12	0	17

B. Deep Learning for classification and recognition

Deep learning is extremely beneficial for facial detection [31], and it is more outstanding than the other state-of-the-art methods [32]. Several previous studies have used it to classify emotion [33] due to its acceptable accuracy. In this study, human faces are classified from GAN and recognized by deep learning. Fig. 10 show the results obtained. RetinaNet is used for detection, and it exhibits high performance and accuracy, albeit with a longer detection time. Moreover, the pre-model has more than RetinaNet, such as YOLOv3 or TinyYOLOv3. Some researchers [34] have used YOLOv3 for recognition, but its accuracy and performance are less than those observed by employing RetinaNet, with a moderate detection time. TinyYOLOv3 is another pre-model that is optimized for speed and moderate performance, with a rapid detection time. However, the performance and accuracy are focused; hence, RetinaNet is used. Table 2 summarizes recognition phase.

In addition, a successful rate per round in WGAN for recognition is attempted; hence, data in case 1–case 4 and plot graph about epoch and accuracy are used to consider the convergence trend. Fig. 11–Fig. 13 show the successful detection rate in WGAN.

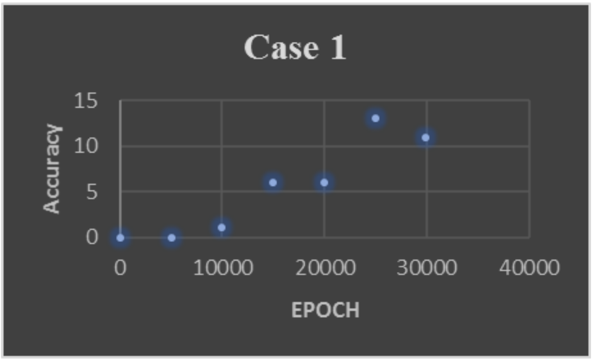


Fig. 11. Successful rate for case 1.

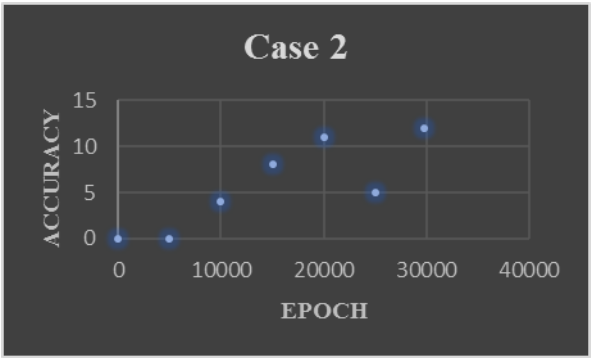


Fig. 12. Successful rate for case 2.

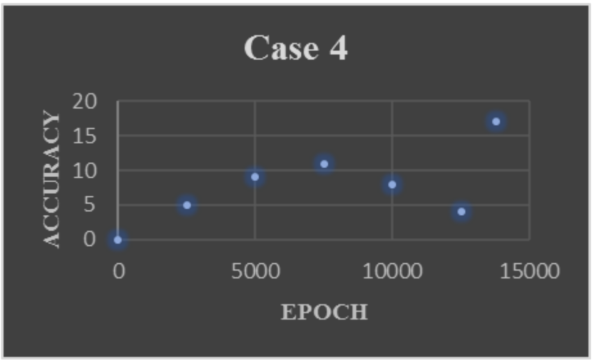


Fig. 13. Successful rate for case 4.

The algorithm used for the classification and recognition in real images needs to be confirmed; hence, it is examined for special cases 1–3. Each of the special cases 1–3 have RGB and grayscale images. Fig. 14–Fig. 15 show examples of special cases 1–3. Special cases 1 and 2 comprise 46 sequential images, while special case 2 is a crop-only interested area to show facial expression and orientation. Special case 3 is different from other cases as we must confirm that the algorithm cannot solve 683 sequential images if collage images such as Fig. 15 are created.

Fig. 14. Example of recognition for special case 2.

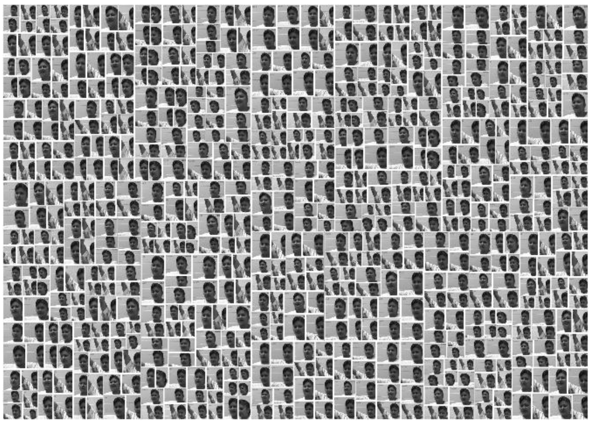


Fig. 15. Example of recognition for special case 3 (greyscale images).

Collage [35] is a sculpture production method, which is predominantly used in graphical arts, where the creation is made from a collection of various forms, creating a new whole. The backgrounds of a collage were discovered hundreds of years ago, but this technique returned in the early 20th century as an art type of innovation. A collage may be seen in magazines, newspapers, photographs, and other objects. In the beginning of the 20th century, Georges Braque and Pablo Picasso invented the term collage, and it became a unique part of contemporary art.

Table 3 show the prediction and detection results for special cases 1–3.

Table- III: Summary of recognition phase for special case

	C1	C1*	C2	C2*	C3	C3*
Number of successful recognitions	**35**	**35**	**37**	**25**	**0**	**0**

*** Mean greyscale images.**

To demonstrated method that increase performance of Deep Learning. We are sampling result of GAN about 15 images mix with 46 real images and tried to classification and recognition. Four extra cases are made: 1. GAN color images with case 1–2 color images 2. GAN color images with case 4 color images 3. GAN grayscale images with case 1–2 grayscale images 4. GAN grayscale images with case 4 grayscale images.

Retrieval Number: B7491129219/2019©BEIESP
DOI: 10.35940/ijitee.B7491.129219
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Fig. 16–Fig. 17 show the results from each category, and Table 4 show the prediction and detection for categories 1–4. Fig. 18 show the full flowchart of our system.

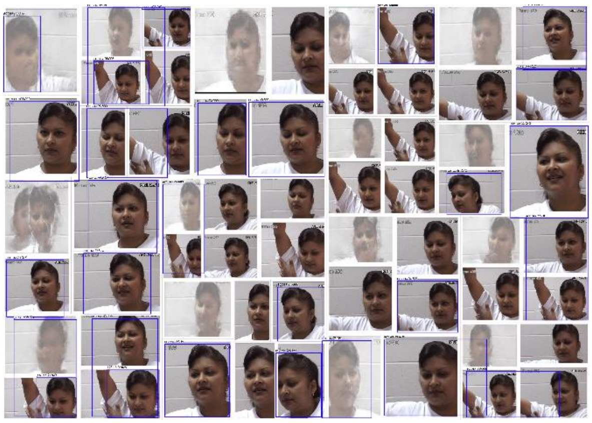


Fig. 16. Example of recognition for extra case 1.

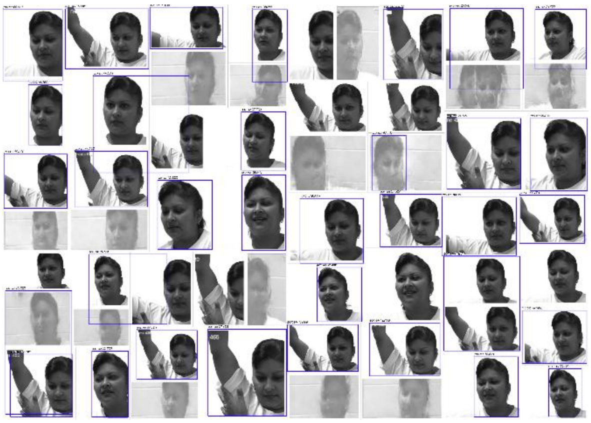
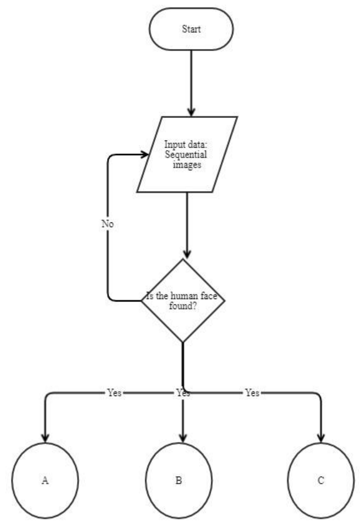
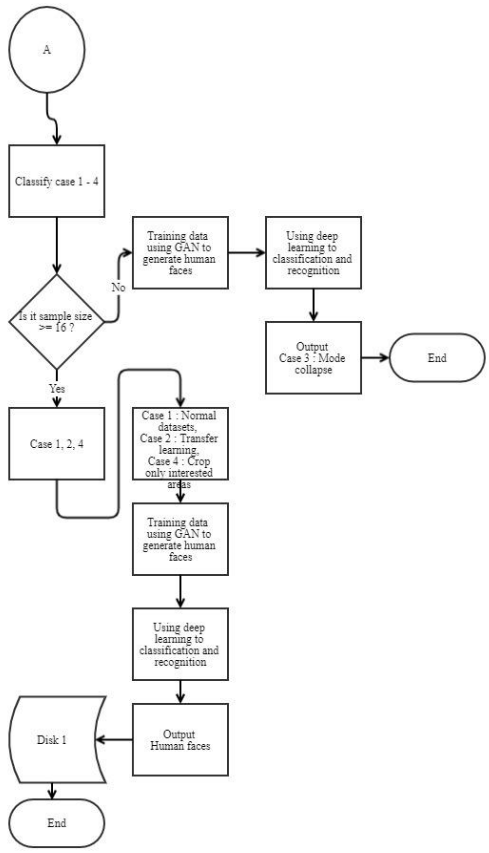


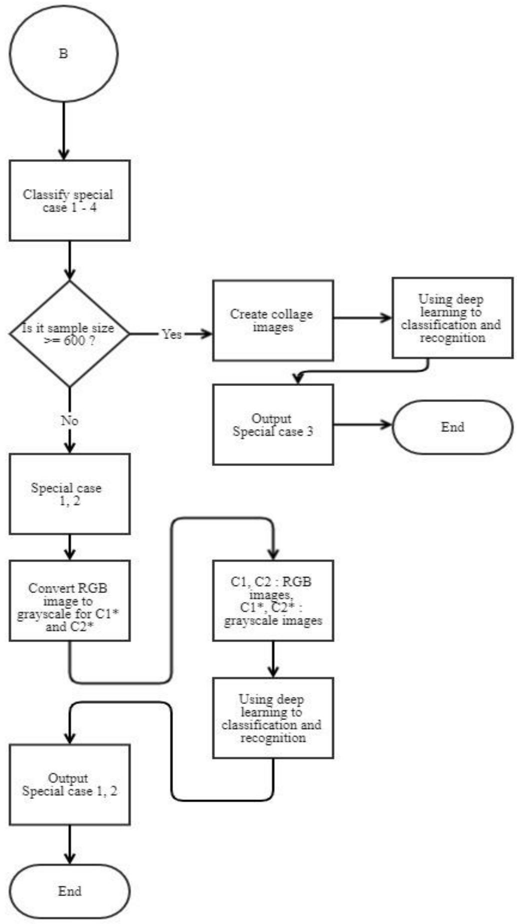
Fig. 17. Example of recognition for extra case 4.

Table- IV: Summary of recognition phase for extra case

	1	2	3	4
Number of successful recognitions	29	35	27	34







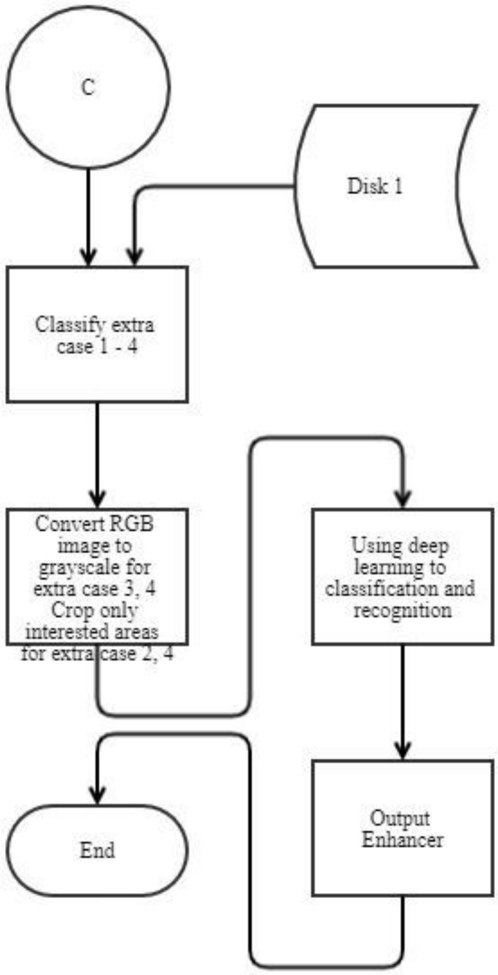


Fig. 18. Flowchart of the system.

IV. CONCLUSIONS AND FUTURE WORK

In this study, a newly developed technique using GAN and deep learning was described to solve the image generation problem in a small dataset. Human faces were used as an example to confirm that the proposed method effectively works in cases that cannot be solved by using less data. The proposed method led to the enhanced stability and accuracy of deep learning by the addition of generated images not only to situations where there is marginal data but also to other difficult conditions. Overall of this system can present with three block diagrams. The first block diagram is the sequential images that we got from the UNBC dataset. After that, the second block diagram training the input images and generate human faces. Following, the third block diagram is the enhancer. The issue of accuracy was addressed using the proposed method to help decrease the failure rate by deep learning training. As part of our future study, we plan to consider other types of GAN, such as deep convolutional generative adversarial network [36], semi-supervised learning with context conditional generative adversarial networks [37], or improved training of the Wasserstein generative adversarial networks [38], which are used to generate images of human faces. The grayscale image is another factor that must be considered because C. Kanan, G.W. Cottrell [39] have claimed that grayscale images can improve the accuracy of the system. The input of grayscale images into GAN will be advantageous to the process. Moreover, our research can be applied for the detection of pain levels and emotions of humans.

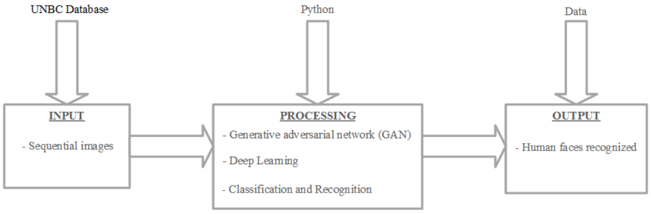


Fig. 19. The block diagram of our system.

ACKNOWLEDGMENT

Financial support from the Thailand Research Fund through the Royal Golden Jubilee Ph.D. Program (Grant No. PHD/0226/2561) to Kornprom Pikulkaew and Varin Chouvatut is acknowledged.

REFERENCES

1. P. Lucey, J.F. Cohn et al., "PAINFUL DATA: The UNBC-McMaster Shoulder Pain Expression Archive Database," Image, Vision, and Computing Journal, vol. 30, pp. 197-205, 2012.W.-K. Chen, *Linear Networks and Systems* (Book style). Belmont, CA: Wadsworth, 1993, pp. 123–135.
2. C.L. Hicks, C.L. von Baeyer et al., "The Faces Pain Scale – Revised: toward a common metric in pediatric pain measurement," Pain, vol. 93, pp. 173-183, August 2001.
3. Ashraf A., P. Lucey et al., "The painful face: Pain expression recognition using active appearance models," Image and Vision Computing, vol. 27, pp. 1788-1796, 2009.
4. I.J. Goodfellow et al. (June 2014). "Generative adversarial networks." [Online]. Available: https://arxiv.org/abs/1406.2661
5. K. Simonyan and A. Zisserman. (Apr. 2015). "Very deep convolutional networks for large-scale image recognition." [Online]. Available: https://arxiv.org/abs/1409.1556
6. I. Goodfellow, J. Pouget-Abadie et al., Generative Adversarial Nets. Neural Information Processing Systems, 2014.
7. M. Arjovsky, S. Chintala et al. (Dec. 2017). "Wasserstein GAN." [Online]. Available: https://arxiv.org/abs/1701.07875
8. GAN Implementation. (Nov. 2017). [Online]. Available: https://github.com/ckmarkoh/GAN-tensorflow/
9. W. Xiao, R. Bhardwaj et al., Gandiva: Introspective Cluster Scheduling for Deep Learning. Operating Systems Design and Implementation, 2018.
10. O.R. Kachare., "Streaming Multiprocessor Architecture, Tensor cores and Nvidia Optimized for Deep Learning and High-Performance Computing (Nvidia Tesla v100 & Nvidia Volta v100)," Electrical and Electronics Engineering, vol. 8, pp. 1-19, 2019.
11. N.P. Jouppi, C. Young et al. (June 2017). "In-Datacenter Performance Analysis of a Tensor Processing Unit." [Online]. Available: https://arxiv.org/ftp/arxiv/papers/1704/1704.04760.pdf
12. T. Carneiro et al., "Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications," IEEE Access, vol. 6, pp. 61677-61685, 2018.
13. V.A. Fusaro, P. Patil et al., "Biomedical Cloud Computing with Amazon Web Services," Computational Biology, vol. 7, pp. 1-6, 2011.
14. J. Salomon, B.S. Phelan. (Apr. 2018). "Lung Cancer Detection using Deep Learning." [Online]. Available: https://www.researchgate.net/publication/324728060_Lung_Cancer_Detection_using_Deep_Learning
15. Y. LeCun, C. Cortes., "The MNIST database of handwritten digits." [Online]. Available: http://yann.lecun.com/exdb/mnist/
16. Dogs vs. Cats Redux: Kernels Edition. (May 2013). [Online]. Available: https://www.kaggle.com/c/dogs-vs-cats-redux-kernels-edition/

17. M. Seeland, M. Rzanny et al. (Feb. 2017). "Plant species classification using flower images—A comparative study of local feature representations." [Online]. Available: https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0170629&type=printable
18. P. Lucey, J.F. Cohn et al., Painful data: The UNBC-McMaster Shoulder Pain Expression Archive Database. Automatic Face and Gesture Recognition, 2011.
19. Z. Hammal, J.F. Cohn., Automatic detection of pain intensity. Multimodal interaction, 2012.
20. K.M. Prkachin, P.E. Solomon., "The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain," Pain, vol. 139, pp. 267-274, 2008.
21. H. Mikami, H. Suganuma et al. (Nov. 2018). " ImageNet/ResNet-50 Training in 224 Seconds." [Online]. Available: https://arxiv.org/abs/1811.05233
22. K. He, X. Zhang et al. (Dec. 2015). "Deep Residual Learning for Image Recognition." [Online]. Available: https://arxiv.org/abs/1512.03385
23. T. Lin, P. Goyal et al. (Feb. 2018). "Focal Loss for Dense Object Detection." [Online]. Available: https://arxiv.org/abs/1708.02002
24. T.M. Hoang, P.H. Nguyen et al., "Deep RetinaNet-Based Detection and Classification of Road Markings by Visible Light Camera Sensors," Sensors, vol. 19, pp. 281-291, 2019.
25. Y. Wang, W. Li et al., Ship Detection by Modified RetinaNet. Pattern Recognition in Remote Sensing, 2018.
26. I. Goodfellow. (Apr. 2017). "NIPS 2016 Tutorial: Generative Adversarial Networks." [Online]. Available: https://nurture.ai/papers/nips-2016-tutorial-generative-adversarial-networks
27. M. Arjovsky, S. Chintala et al. (Dec. 2017). "Wasserstein GAN," [Online]. Available: https://arxiv.org/abs/1701.07875
28. M. Arjovsky, L. Bottou. (Jan. 2017). "TOWARDS PRINCIPLED METHODS FOR TRAINING GENERATIVE ADVERSARIAL NETWORKS." [Online]. Available: https://arxiv.org/pdf/1701.04862.pdf
29. G.K. Dziugaite, D.M. Roy et al. (May 2015). "Training generative neural networks via Maximum Mean Discrepancy optimization," [Online]. Available: https://arxiv.org/abs/1505.03906
30. T. Salimans, I. Goodfellow et al. (June 2016). "Improved Techniques for Training GANs." [Online]. Available: https://arxiv.org/pdf/1606.03498.pdf
31. F.Q. Lauzon., An introduction to deep learning. Information Science, Signal Processing and their Applications, 2012.
32. X. Zhou, W. Gong et al., Application of deep learning in object detection, Computer and Information Science, 2017.
33. J. Huang, X. Xu et al., Emotion classification using deep neural networks and emotional patches, Bioinformatics and Biomedicine, 2017.
34. J. Redmon, A. Farhadi. (Apr. 2018). "YOLOv3: An Incremental Improvement." [Online]. Available: https://arxiv.org/abs/1804.02767
35. M.H. Lee, N. Singhal et al., Mobile photo collage, Computer Vision and Pattern Recognition, 2010.
36. A. Radford, L. Metz et al. (Jan. 2016). "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks." [Online]. Available: https://arxiv.org/abs/1511.06434
37. E. Denton, S. Gross et al., Semi-Supervised Learning with Context-Conditional Generative Adversarial Networks. Learning Representations, 2019.
38. I. Gulrajani, F. Ahmed et al. (Dec. 2017). "Improved Training of Wasserstein GANs." [Online]. Available: https://arxiv.org/pdf/1704.00028.pdf
39. C. Kanan, G.W. Cottrell., "Color-to-Grayscale: Does the Method Matter in Image Recognition?," Plos One, vol. 7, pp. 1-7, 2012.

AUTHORS PROFILE

Kornprom Pikulkaew received the B.S. degree in computer engineering (scholarship) with Bangkok University and the M.S. degree in computer engineering with King Mongkut's University of Technology Thonburi. Now, he is currently pursuing the Ph.D. degree in computer science with Chiang Mai University. His research includes computer vision, computer graphics, human computer interface, 3D-Game, animation and data science.

Ekkarat Boonchieng received the B.S. degree in computer science with Khon Kaen University, the M.S. degree in computer science with University of New Haven and the Ph.D. degree in computer science with Illinois Institute of Technology. Now, he is an associate professor with Chiang Mai University of computer science, and director of Center of Excellence in Community Health Informatics CMU. His current research interests include network, image processing and data science.

Waraporn Boonchieng received the B.S. degree and the M.S. degree in Health Promotion with Chiang Mai University and the Ph.D. degree in Health Education and Behavioral Science with Mahidol University. Now, she is an associate professor with Chiang Mai University at faculty of public health, and head of school in public health.

Varin Chouvatut received the B.S., M.S., and Ph.D. degree in computer engineering from the King Mongkut's University of Technology Thonburi (KMUTT), Thailand. Now, she is an assistant professor with Chiang Mai University of computer science. Her current research interests include image processing, computer vision, computer graphics and machine vision.