Int. J. Advance Soft Compu. Appl, Vol. 12, No. 3, November 2020
ISSN 2074-8523; Copyright © ICSRS Publication, 2020
www.i-csrs.org

Early Detection of Diabetic Retinopathy Cases using Pre-trained EfficientNet and XGBoost

Yunika Laurensia[1], Julio Christian Young[2], and Alethea Suryadibrata[3]

Universitas Multimedia Nusantara, Indonesia
e-mail: yunika.laurensia@student.umn.ac.id [1]; julio.christian@umn.ac.id [2]; alethea@umn.ac.id [3]

Received 1 July 2020; Accepted 28 October 2020

Abstract

Diabetic retinopathy (DR) has been a leading cause of global blindness and its diagnosis through color fundus images requires experienced clinicians which makes this a difficult and time-consuming task. In this paper, a CNN approach is proposed to diagnose DR from digital color fundus images and create a preliminary system as an early detection of DR. Gaussian filter is applied to the images for filtering and resizing purpose before the images are processed any further. Furthermore, EfficientNet and XGBoost are used to extract the images' features and classify the images correspondingly. The network is trained using a high-end GPU on the publicly available Asia Pacific Tele-Ophthalmology Society (APTOS) dataset. On the data set of 3662 Gaussian filtered and resized images used, the proposed model achieves an accuracy score of 98% for binary classification.

Keywords: *diabetic retinopathy, neural networks, efficientnet, xgboost.*

1 Introduction

According to the World Health Organization (WHO), diabetes was the seventh leading cause of death in 2016. In high-income countries, the premature mortality rate due to diabetes decreased from 2000 to 2010 but then increased in 2010-2016. Meanwhile, in lower-middle-income countries, premature mortality due to diabetes increased across both periods [1]. Complications due to the disease are becoming major health issues that require effective health interventions for prevention and treatment. A specific microvascular complication of diabetes is diabetic retinopathy, which affects the vision of 2.6 million people in the world and is responsible for 2.6% of global blindness (0.84 million of 32.4 million people) [2]. The crude global prevalence (all ages) of blindness and vision impairment caused by diabetic retinopathy increased significantly between 1990 and 2015 [3]. Diabetic retinopathy can be classified into two stages, which are non-proliferative and proliferative. The earliest possible signs in proliferative diabetic retinopathy are microaneurysm and retinal hemorrhages [4].

If it could be detected early enough, effective treatment of diabetic retinopathy is available, making the detection a vital process [5]. In order to diagnose diabetic retinopathy at an early stage when it can be treated with the best prognosis that caused visual loss can be delayed or deferred is frequently less than 60%, adherence to regular eye examinations is necessary. This leaves millions of people at risk of diabetes with a potentially preventable visual loss and blindness [6]. However, the diagnosis of diabetic retinopathy through color fundus images requires experienced clinicians to identify the presence and significance of many small features, which along with a complex grading system, causes the diagnosis to be a difficult and time-consuming task. As such, automation for the retinal examination is highly imperative. However, even though automation has improved productivity in many sectors of the economy, in health care, the productivity has remained stagnant in the last 20 years [7]. Thus, it is indeed essential to implement the automation of the retinal examinations which will also contribute to the rise of productivity in the health care sector.

Convolutional Neural Networks (CNNs), a branch of deep learning, have an impressive record for applications in image analysis and interpretation, including medical imaging [5, 17]. Compared with manual extraction techniques, CNNs are able to learn about low-level and high-level features for image classification process more accurately. Based on past researches, in general, the CNNs architecture is created by having a large number of filters in one layer of Neural Network (width), deeper layers (depth), and greater resolution of input image in order to have better performance, such as Xception, DenseNet-201, ResNet-152, VGG-19, and NASNet-Large architectures. In the use of CNN, the appropriate CNN architecture design process will greatly affect the classification performance of the model. The most common method is to scale up the models by its depth or width dimensions. Another less common method is to scale up the models by image resolution. Even though it is possible to do scaling on two or three dimensions at the same time, manual tuning is still required, which results in sub-optimal levels of accuracy and efficiency. It is important to do a balancing of the three dimensions which can be achieved by using a constant aspect ratio in each dimension. One architecture which is able to perform such scaling uniformly in all dimensions by using a coefficient compound is EfficientNet [8].

The aim of this research is to create a binary classification method of diabetic retinopathy as an early detection system. This research utilizes a resized and filtered version of the dataset consisting of 3662 labelled images of which the original version of dataset is provided by the Asia Pacific Tele-Ophthalmology Society (APTOS) [9]. EfficientNet and XGBoost are used to extract the images' features and classify the images correspondingly. This paper will describe the method as well as the findings of the research as a solution to alleviate the problem of the absence of a reliable and efficient method to perform retinal examinations for the early detection of diabetic retinopathy.

2 Fundamentals

2.1 Diabetic Retinopathy

As one of diabetes mellitus complications, diabetic retinopathy is one of the major causes of blindness on working age population [10]. The prevalence of diabetic retinopathy increases within the duration of diabetes and almost all patients with type 1 diabetes and more than 60% of those with type 2 diabetes have diabetic retinopathy after 20 years. Diabetic retinopathy can be classified into two stages: non-proliferative and proliferative. The earliest visible signs in non-proliferative diabetic retinopathy are microaneurysms and retinal hemorrhages. Progressive capillary nonperfusion is accompanied by the development of cotton-wool spots, venous beading, and intraretinal microvascular abnormalities. Proliferative diabetic retinopathy is occurred with further retinal ischemia and characterized by the growth of new blood vessels on the surface of the retina or the optic disc [4]. Regular screening of diabetic retinopathy for patients with diabetes is a cost-effective and important aspect of their treatment.

The accuracy and timing of the treatment are significantly important to both cost and effectiveness of the treatment. Diagnosis of diabetic retinopathy is commonly achieved by imaging the fundus either by retinal photography, by direct or indirect ophthalmoscopy, or by slit lamp bio-microscopy. If detected early enough, effective treatment for diabetic retinopathy is available, making this a vital process. Given the criticality of early diagnosis, opportunistic diagnosis of diabetic retinopathy during routine eye examinations is insufficient [10]. The classification of diabetic retinopathy involves the weighting process of numerous features and the location of such features, which is highly time consuming for clinicians [5].

2.2 Artificial Neural Networks

Inspired by the way biological nervous systems such as human brain processes information, an artificial neural network (ANN) is an information processing system that contains a large number of highly interconnected processing neurons. These neurons work together in a distributed manner to learn from the input information, to coordinate internal processing, and to optimize its final output [11]. Image patch classification is an important task in many different medical imaging applications. In image classification problems, the descriptiveness and discriminative power of features extracted are critical to achieve good classification performance. ANN has been studied for many years to solve complex classification problems including image classification [12].

Their penetration and involvement are almost comprehensive for all medical problems due to the fact that neural networks have the nature of adaptive learning from input information and by using a suitable learning algorithm, neural networks can improve themselves in accordance with the variety and the change of input content. Furthermore, neural networks have the capability of optimizing the relationship between the inputs and outputs via distributed computing, training, and processing, leading to reliable solutions desired by specifications, and medical diagnosis often relies on visual inspection, and medical imaging provides the most important tool for facilitating such inspection and visualization.

ANN has been applied to medical images to deal with the issues that cannot be addressed by traditional image-processing algorithms or by other classification techniques. Compared to other machine learning approaches, neural networks have more positive characteristics. The variety of different network architectures and learning paradigms available, coupled with a theoretically limitless number of combinations of layers amounts, connections topologies, transfer functions, and neuron amounts, make ANNs incredibly flexible processing tools. They can be applied to data with almost any number of inputs and outputs and are well supported in different programming languages and software suites [11].

2.3 Convolutional Neural Networks

Convolutional Neural Network (CNN) is one of deep learning architecture that is used for image processing, both semantic segmentation, instance segmentation, synthetic image generation, and image classification. CNN is a successful example of attempts to model mammal visual cortex using ANN. The architecture of CNN has strong biological plausible evidence support from Hubel and Wiesels early work on the cats' visual cortex. CNN can extract low-level features from the raw inputs, then increasingly more global dan high-level features in deeper layers. For examples, CNN can learn to detect complex visual features like edges, simple shapes, and complete objects from raw images [13]. CNNs are designed to process data that come in the form of multiple arrays, for example a color image composed of three 2D arrays containing pixel intensities in the three color channels. The architecture of a typical CNN is structured as a series of stages. The first few stages are composed of two types of layers: convolutional layers and pooling layers. The role of the convolutional layer is to detect local conjunctions of features from the previous layer, meanwhile the role of the pooling layer is to merge semantically similar features into one [14].

CNN has a great record in application for image analysis and interpretation, includes the medical images [5]. For example, CNN reduced the error rates on the ImageNet image-recognition challenge, where 1.2 million images must be classified into 1000 different classes, from above 26% to below 4% within 4 years. Advantages of using CNN include that it is well suited for end-to- end learning, that is, learning from the raw data without any a priori feature selection, that they scale well to large datasets, and that they can exploit hierarchical structure in natural signals [13].

2.4 Extreme Gradient Boosting Trees

Extreme Gradient Boosting (XGBoost) is a scalable machine learning system for tree boosting that proposed by Chen in 2016 [15]. Given gradient boosting as the original model of XGBoost, it combines weak base learning models into a stronger learner in an iterative fashion. At each iteration of gradient boosting, the residual will be used to correct the previous predictor that the specified loss function can be optimized. As an improvement, regularization is added to the loss function to establish the objective function in XGBoost for measuring the model performance [16].

The most important factor behind the success of XGBoost is its scalability in all scenarios. The system runs more than ten times faster than existing popular solutions on a single machine and scales to billions of examples in distributed or memory-limited settings. The scalability of XGBoost is due to several important systems and algorithmic optimizations. Parallel and distributed computing makes learning faster which enables quicker model exploration. More importantly, XGBoost exploits out-of-core computation and enables data scientists to process hundred millions of examples on a desktop. [15]

3 Related Works

Researches on Diabetic Retinopathy classification using CNN has been carried out with encouraging results. CNN model with 6 layers is used by Ghosh et Al. Kernel size of 2x2 was used in the model for the Maxpooling process and the network is flattened to one dimension after the final convolutional block. The activation layer used PreLU to efficiently train the deep neural network and effective regularization technique known as dropout is used. It resulted in 85% of accuracy for the five-class classification (no DR, mild, moderate, severe, and proliferate DR), 95% accuracy for the two-class classification (DR or no DR), and achieved a kappa score of 0.74 which was further increased to 0.754 using ensemble [10].

Pratt et Al also used CNN to classify DR in the color fundus imageries. The model used perceived increased in convolutional layers to learn deeper features and the network starts with convolution blocks with activation and then batch normalization after each convolution layer. All maxpooling is performed with kernel size 3x3 and 2x2 strides. To avoid overfitting weighted class weights relative to the amount of images in each class is used. The network was also initialized with Gaussian initialization to reduce initial training time. The loss function used to optimize was the widely used categorical cross-entropy function. The number of patients correctly identified as not having DR out of the true total amount not having DR is defined as specificity, meanwhile the number of patients correctly identified as having DR out of the true total amount with DR is defined as sensitivity, and the amount of patients with a correct classification is defined as accuracy. The final trained network achieved, 95% specificity, 75% accuracy and 30% sensitivity. It shows that the five-class problem for DR screening can be approached using a CNN method [5].

4 Research Experiment

The dataset used is provided by APTOS [9] and contains over 3000 images of color fundus images. Those images were resized to 224x224 pixels images and filtered using Gaussian filter. The images were given labels according to their severity of DR on a scale zero to four which are no DR, Mild, Moderate, Severe, and Proliferate correspondingly.

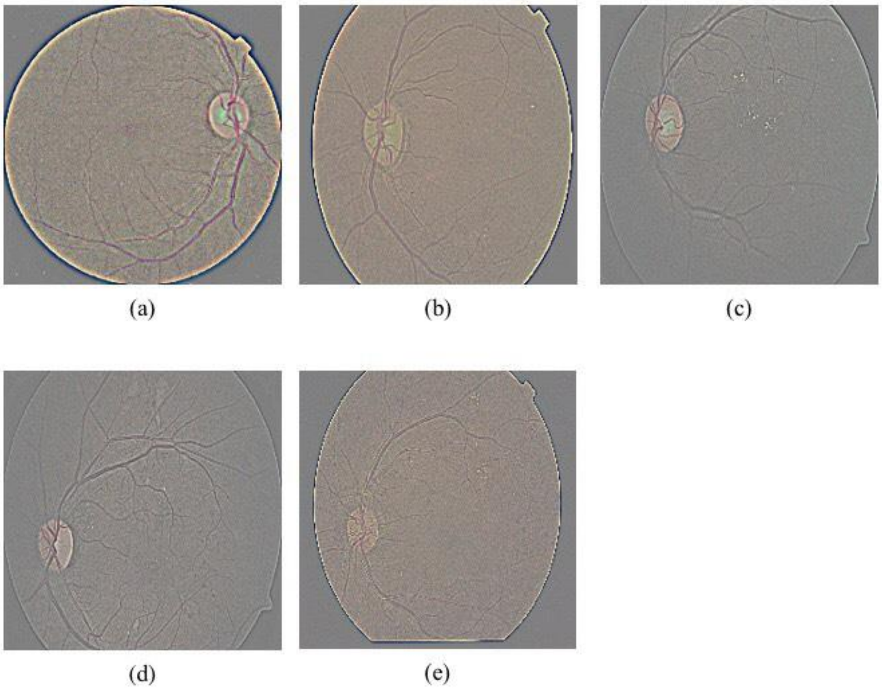

Figure 1: Stages of Diabetic Retinopathy with Increasing Severity (a to e)

To know whether a patient has a diabetic retinopathy, there are some symptoms that could be spotted. Those symptoms include hemorrhages, hard exudates, abnormal growth of blood vessels, microaneurysm, and “cotton wool” spots.

In order to extract the key features of the images, transfer learning is applied to the bottom part of EfficientNetB3 that had been pre-trained for ImageNet’s dataset classification challenge. EfficientNetB3 is used because it has better performance on the particular image size used in this research as compared to the other architectures.

When compared with the size of the dataset, the resulting features obtained from EfficientNetB3 will give rise to a severe over-fitting. Thus, GlobalAveragePooling2D is implemented to resolve this issue. GlobalAveragePooling2D computes the average value of all values across the entire matrix for each of the input channels. As such, it results in a significantly smaller number of features which simultaneously reduces the number of parameters.

Since the main focus of this research is to create a preliminary system for the early detection of diabetic retinopathy, the dataset is then divided into binary classes. The images which do not show any symptoms of diabetic retinopathy are grouped into class 0, while the rest of the images which display the symptoms of diabetic retinopathy are grouped into class 1.

In the subsequent step, the result of the feature extraction is then classified by operating XGBoost. The hyperparameters used include *learning_rate* (the boosting learning rate from one tree to another), *min_child_weight* (the minimum sum of instance weight needed in child), *max_depth* (the maximum depth of a tree), *colsample_bytree* (the fraction of features to be subsampled in a given tree), *gamma* (the minimum loss reduction required), and *n_estimator* (the number of gradient boosted trees). To optimize the aforementioned hyperparameter, BayesSearchCV is utilized. In contrast to GridSearchCV, by using BayesSearchCV, not all parameter values are being fitted, instead, a fixed number of parameters are being sampled from the specified distributions. The range of values of each parameter are as follow.

Table 1: Range of Value for Each Parameter

Parameter	**Range of Value**
learning_rate	0.01 – 1.00
min_child_weight	1 – 10
max_depth	3 – 11
colsample_bytree	0.1 – 0.8
Gamma	1e-9 – 0.5
n_estimators	100 - 250

The parameters of the estimator used are optimized by cross-validated search over parameter settings. Data division by cross-validation generally results in a less biased or less optimistic estimate of the model as compared to other methods, such as a simple train/test split. For the purpose of this research, Stratified KFold is used as the cross-validation splitting strategy. Meanwhile, ROC AUC is used for the scoring. It measures the capability of the model to distinguish between classes, assessed based on the model’s sensitivity and specificity. The sensitivity of the model is recorded as the True Positive Rate (TPR) meanwhile the specificity is recorded as the False Positive Rate (FPR). The ROC Curve is then obtained by plotting the True Positive Rate (TPR) as the y-axis against the False Positive Rate (FPR) as the x-axis.

After the hyperparameter tuning has been implemented by utilizing BayesianSearchCV, the optimum value of each parameter is as follow.

Table 2: Optimum Value of Each Parameter

Parameter	**Range of Value**
learning_rate	0.19
min_child_weight	7
max_depth	3
colsample_bytree	0.1
Gamma	0.5
n_estimators	250

The number of patients correctly identified as not having DR out of the true total amount not having DR is defined as specificity (FPR), meanwhile the number of patients correctly identified as having DR out of the true total amount with DR is defined as sensitivity (TPR). Below is the result of the model prediction in the form of ROC AUC plots.

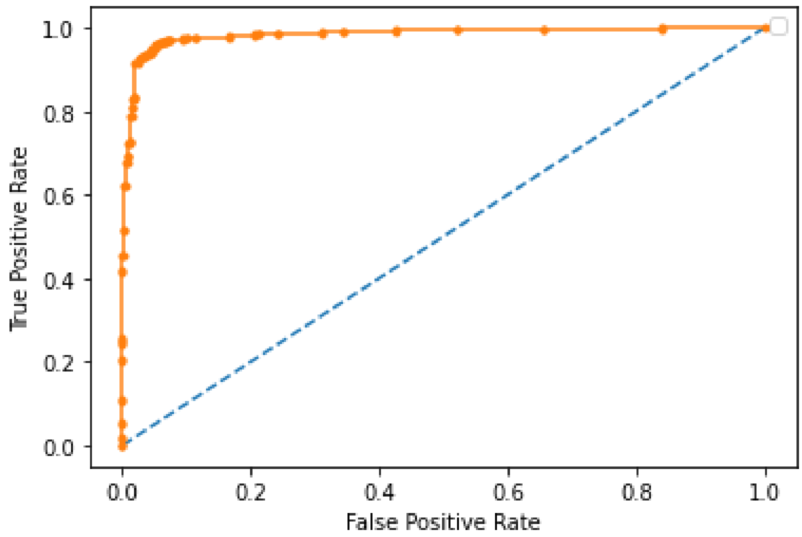

Figure 2: ROC AUC Plots of the Model Prediction

The experiment conducted shows that without large computational resources to train a deep learning model CNN from scratch, a combination of pre-trained EfficientNet-B3 and XGBoosted trees can be utilized as an early detection of diabetic retinopathy. It can detect early case of diabetic retinopathy accurately with ROC AUC score of 98%. Development stage of the model using more minimum computing resource is possible due to transfer learning mechanism that is owned by the CNN model.

5 Conclusion

To conclude, this paper had described the research to create a binary classification method of diabetic retinopathy to alleviate the problem of the absence of a reliable and efficient method to perform retinal examinations for the early detection of diabetic retinopathy. The binary classification of the datasets of images provided by APTOS was carried out using the method as described above. The optimum value of parameters obtained after optimization using BayesSearchCV are 0.19 for the *learning_rate*, 7 for the *min_child_weight*, 3 for the *max_depth*, 0.1 for the *colsample_bytree*, 0.5 for *gamma*, and 250 for the *n_estimators*. Meanwhile, the obtained model has a score of 98%. Thus, it can be concluded that the method as described above is indeed a good method to perform binary classification as an early detection of diabetic retinopathy.

References

[1] World Health Organization. (2020, June 8). "Diabetes". Retrieved from: https://www.who.int/news-room/fact-sheets/detail/diabetes.

[2] Salamanca, O., Geary, A., Suárez, N., Benavent, S., & Gonzalez, M. (2018). Implementation of a diabetic retinopathy referral network, Peru. Bulletin of the World Health Organization, 96(10), 674.

[3] Flaxman, S.R., Bourne, R. R. A., Resnikoff, S., Ackland, P., Braithwaite, T., Cicinelli, M. V., Das, A., Jonas, J. B., Keeffe, J., Kempen, J. H., Leasher, J., Limburg, H., Naidoo, K., Pesudovs, K., Silvester, A., Stevens, G. A., Tahhan, N., Wong, T. Y., & Taylor, H. R. (2017). Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis. *The Lancet Global Health*, 5(12), e1221-e1234.

[4] Mohamed, Q., Gillies, M. C., & Wong, T. Y. (2007). Management of diabetic retinopathy: a systematic review. *JAMA*, 298(8), 902-916.

[5] Pratt, H., Coenen, F., Broadbent, D. M., Harding, S. P., & Zheng, Y. (2016). Convolutional neural networks for diabetic retinopathy. In *Procedia Computer Science*, 90, 200-205.

[6] Abràmoff, M. D., Folk, J. C., Han, D. P., Walker, J. D., Williams, D. F., Russell, S. R., Massin, P., Cochener, B., Gain, P., Tang, L., Lamard, M., Moga, D. C., Quellec, G., & Niemeijer, M. (2013). Automated analysis of retinal images for detection of referable diabetic retinopathy. *JAMA ophthalmology*, 131(3), 351-357.

[7] Abràmoff, M. D., Lou, Y., Erginay, A., Clarida, W., Amelon, R., Folk, J. C., & Niemeijer, M. (2016). Improved automated detection of diabetic retinopathy

on a publicly available dataset through integration of deep learning. *Investigative Ophthalmology & Visual Science*, 57(13), 5200-5206.

[8] Tan, M., & Le, Q. V. (2019, June). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In *Proceedings of the 36th International Conference on Machine Learning on* (Vol. 97, pp. 6105-6114). PMLR.

[9] Asia Pacific Tele-Ophthalmology Society. (2019). “Blindness Detection”. Retrieved from: https://www.kaggle.com/c/aptos2019-blindness-detection/overview.

[10] Bolster, N. M., Giardini, M. E., & Bastawrous, A. (2016). The diabetic retinopathy screening workflow: potential for smartphone imaging. *Journal of Diabetes Science and Technology*, 10(2), 318-324.

[11] Jiang, J., Trundle, P., & Ren, J. (2010). Medical image analysis with artificial neural networks. *Computerized Medical Imaging and Graphics*, 34(8), 617-631.

[12] Li, Q., Cai, W., Wang, X., Zhou, Y., Feng, D. D., & Chen, M. (2014, December). Medical image classification with convolutional neural network. In *13th International Conference on Control Automation Robotics & Vision (ICARCV)* (pp. 844-848). IEEE.

[13] Schirrmeister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggensperger, K., Tangermann, M., Hutter, F., Burgard, W., & Ball, T. (2017). Deep learning with convolutional neural networks for EEG decoding and visualization. *Human Brain Mapping*, 38(11), 5391-5420.

[14] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.

[15] Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794).

[16] Zhang, D., Qian, L., Mao, B., Huang, C., Huang, B., & Si, Y. (2018). A Data-Driven Design for Fault Detection of Wind Turbines using Random Forests and XGboost. *IEEE Access*, 6, 21020-21031.

[17] Prayogo, K. A., Suryadibraya, A., & Young, J. C. (2020). Classification of pneumonia from X-ray images using siamese convolutional network. *Telkomnika*, 18(3).

[18] Young, J. C., & Suryadibrata, A. (2020). Applicability of various pre-trained deep convolutional neural networks for pneumonia classification based on X-Ray Images. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(3), 2649-2654.

Notes on contributors

Yunika Laurensia is an excellent student at the Department of Informatics, Universitas Multimedia Nusantara, Tangerang, Indonesia. She puts interest in image processing, data analysis, and machine learning field. Other than that, she works as junior project manager in a company that engages on data management and consultant.

Julio Christian Young is lecturer at the Department of Informatics, Universitas Multimedia Nusantara, Tangerang, Indonesia. His main teaching and research interests include digital image processing, declarative programming, bioinformatics and applied artificial intelligence and machine learning. He has published several research articles in International Journal of Scientific Technology and Research; International Journal on Telecommunication, Computing, Electronics and Control; and International Journal of Emerging Trends in Engineering Research.

Alethea Suryadibrata is lecturer at the Department of Informatics, Universitas Multimedia Nusantara, Tangerang, Indonesia. Her main teaching and research interests include digital image processing, reliable software engineering, and applied artificial intelligence and machine learning. She has published several research articles in International Journal of Advanced Trends in Computer Science and Engineering; Journal of Information and Communication Convergence Engineering; and International Journal of New Media Technology