2022 5th International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT) | 978-1-6654-7647-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/IMPACT55510.2022.10029203

Breast Ultrasound CAD System Based on Efficient Tumour Segmentation Network and Transfer-Learned Features

Nadeem Zaidkilani*
Department of Computer Engineering and Mathematics
Universitat Rovira i Virgili
Tarragona (Catalonia), Spain
nadimkilani@gmail.com

Mohamed Abdel-Nasser
[1]*Universitat Rovira i Virgili, Spain*
[2]*Aswan University, Egypt*
m.nasser@ieee.org

Miguel Angel Garcia
Department of Electronic and Communications Technology
Universidad Autonoma de Madrid
Madrid, Spain
miguelangel.garcia@uam.es

Domenec Puig
Department of Computer Engineering and Mathematics
Universitat Rovira i Virgili
Tarragona (Catalonia), Spain
domenec.puig@urv.cat

***Abstract*—Breast cancer is the second most common type of cancer worldwide after lung cancer and the leading cause of cancer death among women. Over the past few decades, computer-assisted diagnostic (CAD) systems have been implemented to assist physicians. This paper introduces a CAD system to segment tumours in breast ultrasound (BUS) images and classifies them as benign or malignant. The CAD system has two stages: segmentation and classification. In the segmentation stage, we have developed an encoder-decoder network based on different backbones with various loss functions to segment the tumours. We have fine-tuned the MobileNetv2 network in the classification stage to classify the segmented tumours as benign or malignant. Our experiments demonstrate that WideResNet with BCE and Dice loss function outperforms and yields the best tumour segmentation results with a Dice score of 77.32%. The CAD system achieves a classification accuracy of 86%.**

***Index Terms*—Breast Cancer, Image Segmentation, Loss Functions, Deep learning, CAD System.**

I. Introduction

Early detection is the most effective method for diagnosing and treating breast cancer since it increases the likelihood of treatment effectiveness and survival. Early detection and diagnosis of breast cancer expands treatment options and decreases mortality by 25% [1]. Different computer-aided diagnosis (CAD) systems have been proposed to analyze common diseases, including breast cancer [2]–[6].

The precision of some CAD systems depends on the accuracy of the segmentation of the region of interest (ROI), i.e., the tumor region. Sonography uses high-frequency sound waves to create an image of the body. Because ultrasound images are collected in real-time, they can also depict the movement of the body's internal organs and blood moving via the blood arteries. Unlike X-ray imaging, ultrasound imaging does not expose the patient to ionizing radiation. The safety of ultrasound imaging has been exemplary for more than two decades.

Manual inspection determines the lesion's benign or malignant status, sometimes based on its delineation. In contrast, manual inspection is subjective, time-consuming, and prone to error. The segmentation of breast cancer plays a crucial role in determining the tumour's location, shape, and, ultimately, stage. As a result of segmentation, the therapy dosage for breast cancer is dependent on tumour size. Different machine learning-based segmentation approaches have been suggested for medical image segmentation. The commonly used machine learning techniques for this task are fuzzy C-clustering, k-means clustering, and hierarchical k-clustering support vector machines (SVM).

Recently, deep convolutional neural networks (CNNs) have been successfully applied in various fields due to their remarkable ability to extract features. U-Net [7] is frequently used for ultrasound image segmentation because it does not require many annotated images compared with other deep learning models. In addition, previous research demonstrates that it is possible to train a deep learning model without preprocessing or postprocessing. This study employs cutting-edge deep CNNs-based semantic segmentation techniques to segment breast tumors in BUS images. Modern CNNs are trained using various loss functions (Cross-entropy loss, Dice loss, Tversky loss, and Focal loss) to improve the system's accuracy. Furthermore, we extracted deep learning-based radiomics (features) for classifying breast tumours as benign or

*Corresponding Author:Nadeem Zaidkilani E-mail:nadimkilani@gmail.com

Authorized licensed use limited to: Wuhan University. Downloaded on August 31,2023 at 06:41:24 UTC from IEEE Xplore. Restrictions apply.

malignant.

The remainder of this research is organized as follows: Section II provides an overview of the related studies in this field. Section III presents the proposed CAD system. Section IV provides the setup of the experiment and results. Section V presents the conclusions and future work.

II. Literature Review

Recently, deep learning techniques for breast image analysis have demonstrated significant improvement. These efficient, data-driven techniques process input images to discover high-level image representations. Based on these features, it can perform classification and segmentation tasks. Many deep learning architectures have been proposed to segment medical images, such as Fully Convolutional Network (FCN) [8], SegNet, U-NET, and GAN [9]. The U-NET architecture [10] is the state-of-the-art architecture for medical image segmentation. U-NET has an encoder-decoder structure that downsamples the spatial dimension to extract features and then upsamples to restore the spatial extent. It uses skip connections to maintain spatial information, enhancing the segmentation process. The encoder captures deep semantic information through multiple down-sampling processes. At the same time, the decoder part gradually restores the spatial and detailed information of the input image through several up-sampling operations.

Several deep learning-based models for breast tumor segmentation in BUS images have recently been proposed. For example, Yap et al. [11] developed several Fully FCN-based variants for semantic segmentation of breast lesions in BUS images. With a dataset of 113 malignant and 356 benign BUS images, they obtained a dice score of 0.76 with benign lesions and a dice score of 0.54 with malignant lesions. Almajalid et al. [12] modified and enhanced U-Net for lesion segmentation in BUS images with contrast enhancement and speckle reduction. With 221 BUS images, they obtained a dice value of 0.83. Yap et al. [11] applied deep transfer learning to develop a breast mass segmentation model based on FCNs. Xu et al. [13] combined the dilated convolution method with a phase-based active contour (PBAC) model to develop an accurate breast tumor segmentation model. Han et al. [14] proposed a semi-supervised breast tumor segmentation method based on GAN networks to improve the segmentation quality of unannotated BUS images.

This research presents a deep CNN-based CAD system for segmenting and classifying breast tumors.

III. Methodology

In this paper, we develop a complete CAD system for breast cancer tumor segmentation and classification using modern techniques of deep learning architectures. Fig. 1 depicts the proposed CAD system. Firstly, using the segmentation model, the CAD system automatically localizes and segments breast tumors from BUS images. After that, the segmentation mask is multiplied with the input BUS image to extract the breast tumor region (i.e., ROI). Finally, the CAD system uses a deep learning feature extractor to extract texture information from the ROI to classify the segmented tumors as benign or malignant.

To build the CAD system, we built two models: one for the segmentation task and the other for the classification task. We train different encoder-decoder segmentation networks to segment breast tumors. Then, we used the best-trained segmentation network to segment BUS images. Also, we train a CNN network to extract texture features from segmented tumors to classify them as benign or malignant. For the segmentation task, we employ different backbones, such as VGG 13, VGG16, MobileNet, ResNet18, ResNext, and WideResNet, as encoders, and we built a decoder part for tumour segmentation. Below, we briefly explain the architecture for each network.

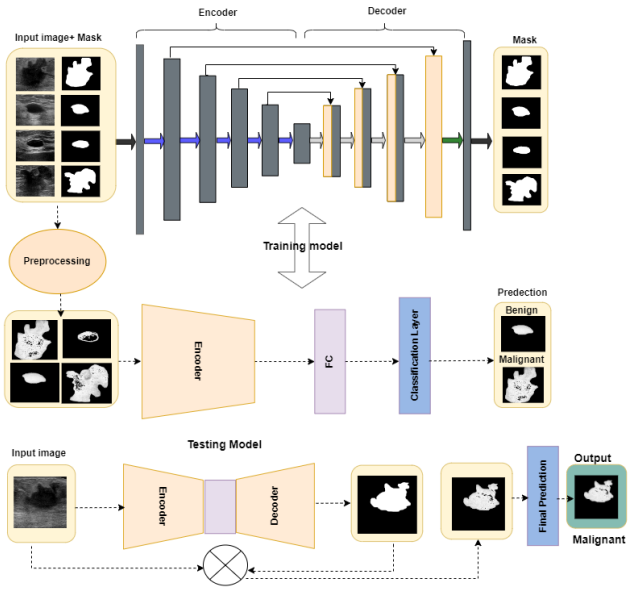

Fig. 1. Architicure of the proposed CAD system

- VGG 13 and VGG16 [15] are conventional CNNs that utilize small 3 x 3 filters.
- Mobile Net [16] employs depthwise separable convolutions. It greatly decreases the number of parameters compared to networks with standard convolutions of the same depth in lightweight deep neural networks.
- ResNet18 [17] utilizes skip connections and incorporates intensive batch normalization..
- ResNext [18] is a version of ResNet; the main concept is that the input is added to the layer's output in a basic residual block. Here, instead of generating the output from a single layer, we concatenate the outputs of multiple layers and then add the input.
- WideResNet [19] is a version of ResNets in which the depth of the residual network is reduced, and the width is enhanced by employing wide residual blocks.

In this study, we employed different combinations of loss functions, such as Cross-entropy loss, Dice coefficient loss,

Authorized licensed use limited to: Wuhan University. Downloaded on August 31,2023 at 06:41:24 UTC from IEEE Xplore. Restrictions apply.

Tversky loss, and Focal loss. More details in the loss functions subsection

A. Dataset and Data Augmentation

In this study, we used a dataset provided by the UDIAT Diagnostic Centre of Sabadell, Spain. The size of the dataset is 163 images, corresponding to 110 benign and 53 malignant breast masses (one mass per image). The ground truths of the tumour regions in the BUS images are already available in this dataset. We have randomly split the dataset into 113 BUS images for training and 50 BUS images for testing. To train the segmentation models, we applied data augmentation techniques such as rotation, width shift, height shift, shear, and horizontal flip. Fig. 2 shows examples of BUS images after applying data augmentation. It should be noted that we applied this technique to the training set only. We kept 50 images out of 163 for testing to evaluate our model in segmentation and classification tasks. The whole training set contains 1400 images.

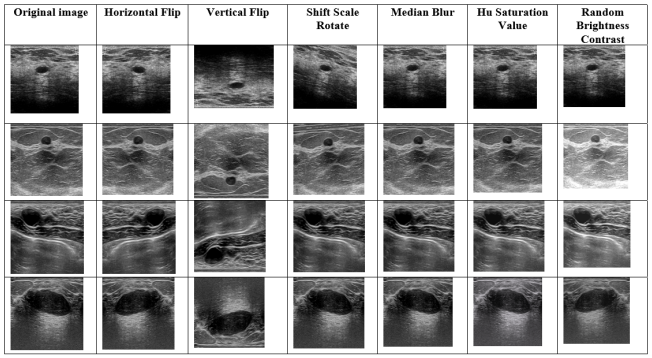

Fig. 2. Examples of BUS images after applying different augmentation techniques.

B. Loss Functions

The loss function is an essential component for CNN training. Each loss function has advantages and disadvantages. Thus, we chose to evaluate the efficiency of several loss functions. The loss functions can be defined as follows:
Cross entropy (CE) measures the difference between two probability distributions for a given random variable or set of events. The CE can be defined as follows:

$$\mathrm{L}\,BCE\,(y, y^{\wedge}) = -\left(y \log\left(y^{\wedge}\right) + (1-y)\log\left(1-y^{\wedge}\right)\right) \quad (1)$$

Here, y and y$\wedge$ are the actual and predicted class, respectively.

The *Dice coefficient* loss function measures the overlap between the predicted and targeted samples. It is used for the binary data; the Dice loss can be defined as follows:

$$\text{Dice Loss } = 1 - 2 \cdot \frac{Precision * Recall}{Precision + Recall} \quad (2)$$

where we defined the Dice (F1 score) in evaluation metrics.

The *Focal loss* is an improved Cross-Entropy Loss (CE) version. It down-weights the contribution of easy examples and enables the model to focus more on learning hard examples. It works well for highly imbalanced class scenarios. The focal can be defined as follows:

$$\text{Focal loss} = -\alpha_t \left(1-p_t\right)^{\gamma} \log\left(p_t\right) \quad (3)$$

where α is a hyperparameter that can be tweaked for further calibration. Here,γ >0 and when γ= 1. Notably, Focal Loss works like the CE loss function.
The *Tversky loss* is a generalization of Dice's coefficient. It adds weight to FP (false positives) and FN (false negatives). The Tversky loss can be defined as follows:

$$\text{Tversky loss} = 1 - (\text{TP} + \zeta)/(\text{TP} + \alpha * \text{FN} + \beta * \text{ FP } + \zeta) \quad (4)$$

In this expression, α stands for a weight constant that penalizes the model for FPs, β is a weight parameter that penalizes the model for FNs, and ζ is a minimal value to avoid division zero.

C. Evaluation Metrics

The segmentation models are evaluated using segmented masks and the corresponding ground truth. The segmented shows were binarized so that pixels of the tumor has the label '1' while pixels in the background have the label '0'. We computed the number of true positive (TP), true negative (TN), false positive (FP), and false-negative (FN) pixels. Then, we computed the evaluation metrics commonly used for the tumour segmentation task in the ultrasound images: the intersection-over-union (IOU) and the Dice coefficient (Dice).

i) *IOU* (Jaccard Index) is the ratio of the intersection between the two masks (original and predicted mask) concerning their union. It is computed as:

$$IOU(A,B) = \frac{A \cap B}{A \cup B} = \frac{TP}{TP + FP + FN} \quad (5)$$

ii) *Dice coefficient* is the harmonic mean of precision and recall. It is also referred to as F-score, which can be computed as:

$$\text{F1 score } = 2 \cdot \frac{\text{Precision. Recall}}{\text{Precision } + \text{ Recall}} \quad (6)$$

iii) *Accuracy* is a metric that generally describes how the model performs across all classes. It is useful when all classes are of equal importance. It is calculated as the ratio between the number of correct predictions to the total number of predictions, which can be computed as:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+FN+TN}} \quad (7)$$

Authorized licensed use limited to: Wuhan University. Downloaded on August 31,2023 at 06:41:24 UTC from IEEE Xplore. Restrictions apply.

IV. Experimental Results

In our experiments, we divided the dataset into 113 images of training data and 50 images of test data. We have used a batch size of 2 and an Adam optimizer with a learning rate of 0.0001 and 20 epochs for training. As mentioned in the above section, we only used data augmentation on the training set.

A. Analyzing the performance of the segmentation models

In our experiments, we tested different augmentation methods. After we applied the data augmentation methods mentioned above to the original dataset, the size of the new dataset was almost 1400 images. It is worth noting that data augmentation significantly improves our model's performance. Before applying the data augmentation, the Dice score of the UNET model was 60%. We started the first experiment to study what is the suitable loss function that leads to the best performance based on the UNET model. Fig. 3 demonstrates that the best loss function is the combination of Dice and BCE loss function. Therefore, we used this fusion loss function in the remainder experiments.

TABLE I
THE PERFORMANCE OF THE SEGMENTATION MODELS USING DIFFERENT BACKBONES.

Evaluation Metrics	Jaccard	Dice
UNET	0.67	0.76
VGG13	0.64	0.72
VGG16	0.65	0.73
Mobile Net	0.61	0.70
ResNet18	0.64	0.71
ResNext	0.63	0.71
WideResNet	0.69	0.77

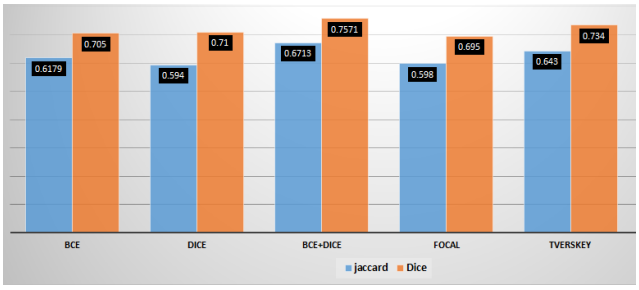

Fig. 3. Impact of different types of loss functions.

In this research, we used various networks like ResNext and WideResNet as a backbone for the encoder. Table 1 presents the evaluation metrics of the segmentation model with different backbones. Fig. 4 presents samples of predicted images of all models to demonstrate their performance. The experimental result demonstrates that WideResNet outperforms other models.

B. Classification results

In this experiment, we trained our classification model using the same training set we used on the segmentation model. We tested our model on the same test set. In this experiment, we

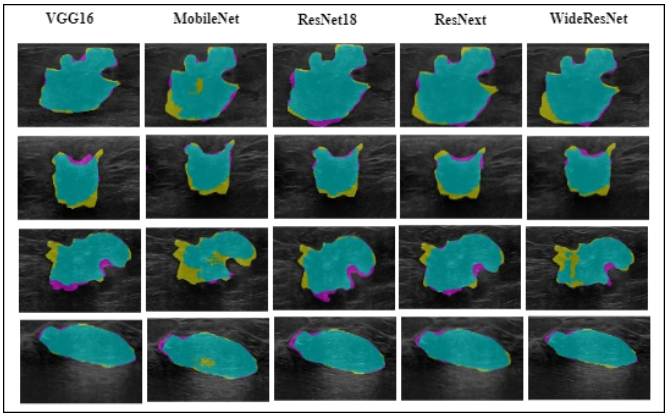

Fig. 4. Segmentation results.

used fine-tuned MobileNetv2 with 50 epochs and batch size 4 to classify the segmented breast tumors as benign or malignant. When we trained the classification model using the whole BUS image, we achieved an accuracy of 68%. In turn, our method achieves an accuracy of 86%. It is worth mentioning that our model achieves an accuracy of 78% when we classify the predicted image from the segmentation model in the CAD system. Fig. 5 presents a sample of the BUS images and the predicted class: malignant or benign.

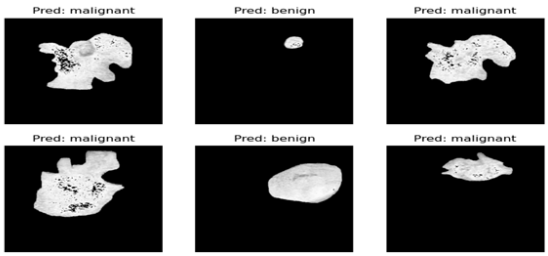

Fig. 5. Classification results.

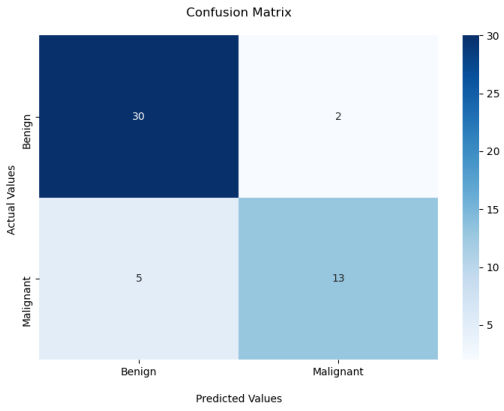

Fig. 6. Confusion matrix of the classification model.

Figures 6 and 7 show the confusion matrix and ROC curve of the proposed CAD system. The confusion matrix

Authorized licensed use limited to: Wuhan University. Downloaded on August 31,2023 at 06:41:24 UTC from IEEE Xplore. Restrictions apply.

summarizes the number of correct and incorrect predictions made by a classifier of the benign and malignant classes. In the case of benign tumors, two predictions are wrong out of 32 images, whereas in the case of malignant tumours, five predictions are wrong out of 18 images. The ROC curve shows the trade-off between sensitivity (or TPR) and specificity (1 – FPR). The area under the ROC curve (AUC) of the proposed CAD system is 83%.

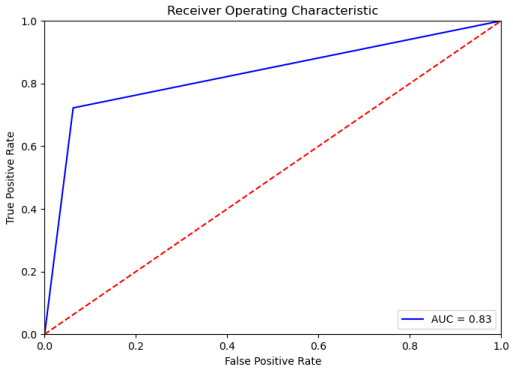

Fig. 7. ROC curve of the classification model.

V. Conclusions and Future Work

This paper introduced a two-stage CAD system for breast tumor segmentation and classification based on the latest deep learning techniques. With a dataset of 163 ultrasound images, the proposed CAD system achieved a Dice score of 77.32% in the segmentation task and an accuracy of 86% in the classification task. In future work, we will employ deep attention mechanisms and a metaheuristic algorithm like WOA [20] to enhance the performance of the proposed CAD system. Moreover, we will use recent deep learning architectures like ConvNext, Transformers, and CoAtNet.

Acknowledgements

The Spanish Government partly supported this research through Project PID2019-105789RB-I00.

References

[1] Hela B, Hela M, Kamel H, Sana B, Najla M. Breast cancer detection: A review on mammograms analysis techniques. In: 10th International Multi-Conferences on Systems, Signals & Devices 2013 (SSD13). IEEE; 2013. p. 1-6.

[2] Abdel-Nasser M, Moreno A, Puig D. Temporal mammogram image registration using optimized curvilinear coordinates. Computer methods and programs in biomedicine. 2016;127:1-14.

[3] Mahmoud H, Abdel-Nasser M, Omer OA. Computer aided diagnosis system for skin lesions detection using texture analysis methods. In: 2018 International Conference on Innovative Trends in Computer Engineering (ITCE). IEEE; 2018. p. 140-4.

[4] Abdel-Nasser M, Saleh A, Moreno A, Puig D. Automatic nipple detection in breast thermograms. Expert Systems with Applications. 2016;64:365-74.

[5] Abdel-Nasser M, Melendez J, Moreno A, Puig D. The impact of pixel resolution, integration scale, preprocessing, and feature normalization on texture analysis for mass classification in mammograms. International Journal of Optics;2016.

[6] Saffari N, Rashwan HA, Abdel-Nasser M, Kumar Singh V, Arenas M, Mangina E, et al. Fully automated breast density segmentation and classification using deep learning. Diagnostics. 2020;10(11):988.

[7] Soukup D, Huber-Mörk R. Convolutional neural networks for steel surface defect detection from photometric stereo images. In: International Symposium on Visual Computing. Springer; 2014. p. 668-77.

[8] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2015. p. 3431-40.

[9] Fu Y, Lei Y, Wang T, Curran WJ, Liu T, Yang X. A review of deep learning based methods for medical image multi-organ segmentation. Physica Medica. 2021;85:107-22.

[10] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. Springer; 2015. p. 234-41.

[11] Yap MH, Goyal M, Osman FM, Martí R, Denton E, Juette A, et al. Breast ultrasound lesions recognition: end-to-end deep learning approaches. Journal of medical imaging. 2018;6(1):011007.

[12] Almajalid R, Shan J, Du Y, Zhang M. Development of a deep-learning-based method for breast ultrasound image segmentation. In: 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE; 2018. p. 1103-8.

[13] Hu Y, Guo Y, Wang Y, Yu J, Li J, Zhou S, et al. Automatic tumor segmentation in breast ultrasound images using a dilated fully convolutional network combined with an active contour model. Medical physics. 2019;46(1):215-28.

[14] Han L, Huang Y, Dou H, Wang S, Ahamad S, Luo H, et al. Semi-supervised segmentation of lesion from breast ultrasound images with attentional generative adversarial network. vol. 189. Elsevier; 2020.

[15] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556. 2014.

[16] Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:170404861. 2017.

[17] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 770-8.

[18] Xie S, Girshick R, Dollár P, Tu Z, He K. Aggregated residual transformations for deep neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. p. 1492-500.

[19] Zagoruyko S, Komodakis N. Wide residual networks. arXiv preprint arXiv:160507146. 2016.

[20] Mohamed F, Abdel-Nasser M, Mahmoud K, Kamel S. Economic dispatch using stochastic whale optimization algorithm. In: 2018 International Conference on Innovative Trends in Computer Engineering (ITCE). IEEE; 2018. p. 19-24.

Authorized licensed use limited to: Wuhan University. Downloaded on August 31,2023 at 06:41:24 UTC from IEEE Xplore. Restrictions apply.