International Conference on Advances in Materials, Computing and Communication Technologies
In Association with International Journal of Scientific Research in Science and Technology
Volume 9 | Issue 1 | Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)

Liver Tumor Detection in CT Images by VGG16 Convolutional Neural Network

Keerthana G[1], Keerthi T[1], Preetha Lekshmi N[1], A. Anisha[2]

[1]UG Student, Department of CSE, St .Xavier's Catholic College of Engineering, Chunkankadai, Nagercoil, Tamil Nadu, India

[2]Assistant Professor, Department of CSE, St .Xavier's Catholic College of Engineering, Chunkankadai, Nagercoil, Tamil Nadu, India

ABSTRACT

Automatic tumor detection and segmentation is essential for computer-aided diagnosis of liver tumors in CT images. However, it is a challenging task in low-contrast images as the low-level images are too weak to detect. In this project, we propose a new method for the automatic detection of liver tumors. VGG16 CNN has been used as a powerful tool for liver cancer analysis. Whereas the CT-based lesion-type definition defines the diagnosis and therapeutic strategy, the distinction between cancer and non-cancer lesions is crucial. It demands highly qualified experience, expertise, and resources. However, a deep end-to-end learning approach to help discrimination in abdominal CT images of the liver between liver metastases of colorectal cancer and benign cysts has been analyzed.

I. INTRODUCTION

SYSTEM ANALYSIS

EXISTING SYSTEM

The Multi-scale candidate generation (MCG) for the liver tumor segmentation approach on CT images. They utilized as an active contour model and 3D fractal residual network in a coarse to fine-tune the liver cancer cells. First, the liver is segmented by 3D U-Net and tumor candidates, which are identified using the MCG method. They suggest 3D FRN in respect of the candidate class. This paper proposes a variation of the superpixel segmentation approach in multiple scales and information in the neighborhood to produce candidates for segmenting the liver tumor that can require more detailed tumor data in the candidate area. It enhances the sensitivity of the network to liver tumor specifics and decreases the computational difficulty arising from redundant data. They have performed 3DIRCADb segmentation tasks, the experiment results, and comparisons with the associated research show that their advanced system can achieve a high segmentation efficiency.

Watershed Transform and Gaussian mixture model (WT-GMM) for liver cancer recognition based on deep learning. This approach relies on the marker-controlled transformation of the watershed and the Gaussian blend model for reliable detection. The algorithm proposed is tested with clinical data from various patients in the clinical set-up in real-time. The essential advantage of this automated detection is that the deep-neural network classifier produced the

best precision of 99.38 percent with negligible validation loss. The first method of application for detecting liver tumors is the use of the DNN model in the detection process. The proposed method is analyzed using an efficient way to identify the region of cancer from liver CT images that will be beneficial for the early diagnosis of symptoms in clinical and decision-making process. The estimation of the volumetric size of the lesion is the key limit of the job, which can be constructed from various image slices in a three-dimensional mesh structure.

A Fully convolutional network for liver and lesion recognition. They are investigating the FCN by way of contrast on a relatively small dataset with patch-based classification systems for CNN and sparsity. They have CT scans of 20 patients, each with a minimum of 67 lesions and 42 livers in a single slice and 22 patients with complete 3D segmentation of the liver. They have carried out 3 times cross-validation, and the results show that the FCN is superior to all other test methods. They have obtained real positive rates of 0.85 and 0.7 false-positive per case using our fully automatic algorithm.

SYSTEM DESIGN

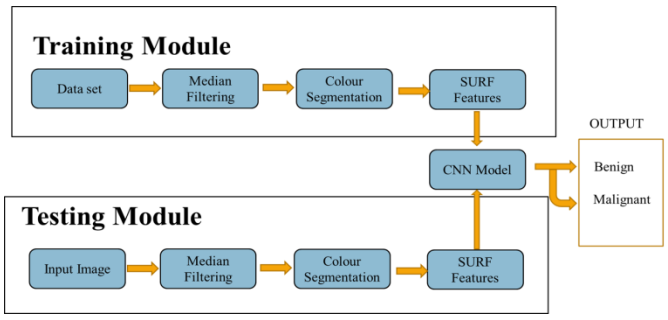

FIGURE 1. Block Diagram of Proposed System

II. METHOD

The flowchart of proposed liver tumor segmentation method, which includes (1) liver region segmentation; (2) liver tumor candidate generation; (3) tumor candidate classification, and (4) liver tumor boundary refinement using an active contour model [27].

The specific steps are as follows.

Step 1: We segment liver regions from CT volumes through 3D U-Net [26].

Step 2: The multi-scale candidate generation method, which combines multi-scale super pixels segmentation method based on linear spectral clustering (LSC) [28]and multiple neighbourhood information, is proposed for segmenting tumor candidates within liver regions.

Step 3: 3D FRN combining the fractal structure [29] and the residual structure [30] is proposed for classifying the liver tumor candidates.

Step 4: The active contour model [27] is introduced to refine the boundary of the liver tumor and the high probability region in the network output is used as the initial contour.

A. TUMOR CANDIDATE GENERATION METHOD IN THE LIVER REGION

Because the size of liver tumors in the CT image varies, accurate liver tumor segmentation remains a challenge. To this end, the region of background needs to be narrowed by segmenting the liver region, which can effectively improve the segmentation efficiency of tumors. In this paper, we use the 3D U-Net for liver segmentation before liver tumor segmentation. Although the image after liver segmentation reduces a large number of non-interested regions, the tumor region of interest is still too small to be segmented in liver regions.

To solve this problem, we decided to cut each segmented liver image into tumor candidates and then classify the tumor candidates to obtain the segmentation results. Nowadays, the effective candidate generation method is to use a sliding window to generate candidate regions pixel by pixel, which is a type of pixel-level classification. Although the neighbourhood information of each pixel can be fully considered in this method, it also causes a series of problems, such as, large amount of redundant information and significant requirement of calculation. Also, some scattered points are caused by

false positives, which can be misleading to doctors. Therefore, for the above problems, we propose a tumor candidate generation method (MCG) for dividing liver regions into tumor candidates. The method includes multi-scale superpixel method and multiple neighborhood information.

1) MULTI-SCALE SUPERPIXEL METHOD

The superpixel method, linear spectral clustering (LSC), which maps images to high-dimensional spaces to find relationships between pixels, putting the same type of pixels into the same candidates. LSC can generate good candidates with prior knowledge which benefits tumor segmentation. Therefore, LSC is used to generate tumor candidates.

The traditional superpixel classification method is to directly classify superpixels as classification units and class them by extracting superpixel features. However, it has been experimentally found that the single-scale segmentation results heavily depend on the superpixel segmentation results. If the segmentation result of a single-scale superpixel segmentation method cannot segment the tumor boundary accurately, it will also affect the final segmentation result. The single-scale superpixel segmentation method has low tolerance to false positive and is based on a single-scale candidate region. Consequently, only limited local information is used, which results in a higher false positive rate and lower sensitivity. Therefore, a multi-scale superpixel method is chosen as the method to generate tumor candidates.

For $I_k \in R_{c\times c}$, the kth slice of input image I which only includes the liver region, we calculate superpixel results of s scales. The formula for the single-scale superpixel result is as follows.

$$F_n(I_k) = LSC(I_k, n) \quad (1)$$

where LSC (•) stands for linear spectral clustering operation. Parameter n represents the number of superpixels we segment Ik by LSC in one scale. Fn (Ik) is the superpixel mask of Ik and lth superpixel region in Fn (Ik) is B(k,l). Some examples of superpixel segmentation are shown in Figure 2.

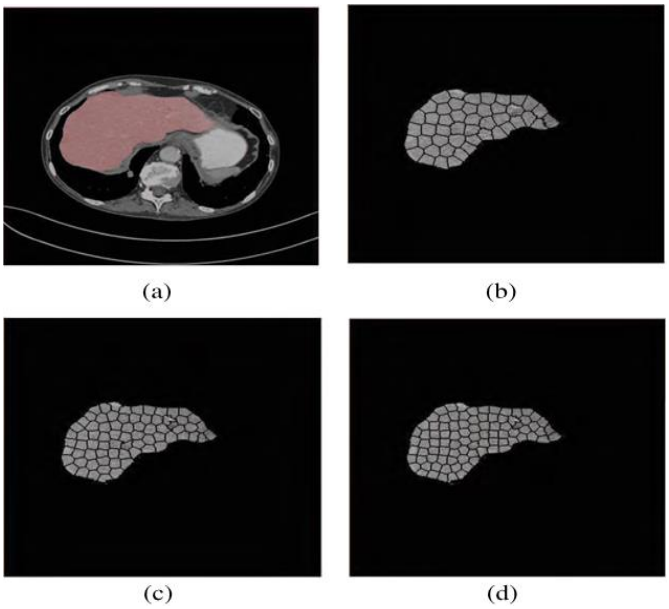

FIGURE 2 . Examples of three-scale superpixel segmentation results.

(a) Input CT image with the red region indicating the liver region.
(b) Superpixel segmentation result for scale with nD 1000.
(c) Superpixel segmentation result for scale with nD 1500.
(d) Superpixel segmentation result for scale with nD 2000.

2) MULTIPLE NEIGHBORHOOD INFORMATION

Although using the multi-scale super pixel segmentation method can generate good tumor candidates, these tumor candidates ,which contain only single slice image information and ignore background information, are not sufficient for effective segmentation. Therefore, a novel approach (MCG) which includes both context information and multi-scale information is developed to increase multiple neighbourhood information in super pixel results of each scale. Let function T represent the transformation from the liver region volume data to the input of the proposed network and T –1 stand for the transformation from the output of the proposed network to the segmentation results:

Xin = T (Ik,Fn (Ik)), Xin ∈ Rm×i×L×L×s (2)

Y = T −1 (Xout,Fn (Ik)), Y ∈ Rm×512×512 (3)

where m is the number of slices, T −1 is the inverse operation of T , Xin is the input of the 3D fractal residual network (3D FRN) and Xout is the output of the 3D FRN. T includes a series of operations and the detailed process of T is as follows.

First, we use LSC to generate Fn (Ik) of Ik and calculate the area S(k,l) of B(k,l). Second, if B(k,l) is not all 0, the center point P(k,l) of B(k,l) is used as the center of the square tumor candidates. Third, the size of the tumor candidates L(k,l,i) is determined according to S(n,k). There are i kinds of size we choose, and each size is calculated as follows,

$$L(k,l,i) = \begin{cases} \lfloor \sqrt{S_{(k,l)}} \rfloor, & \text{if } i = 1 \\ L(k,l,1) + (i-1)L(k,l,1), & \text{otherwise} \end{cases} \quad (4)$$

Then, i square areas, which take P(k,l) as center point, L(k,l,i) as side length, are taken from Ik and resized to the mean size among all candidates. They are then stacked into an i channel tumor candidate.

Finally, the context information is also utilized to introduce more sufficient neighbourhood information. For an image Ik, Ik−1 and Ik+1 are involved to generate 3-channel tumor candidates. The class of a tumor candidate can be predicted through the 3D FRN and the prediction result is also regarded as the class of the corresponding super pixel that generates the candidate. For a specific super pixel scale of Ik, the classification results of all super pixels are combined based on Fn(Ik) to form the image-level result of this super pixel scale. The detailed generation process is shown in Figure 3. We obtain s binary image-level results corresponding to s super pixel scale, which means, for each pixel, there are s prediction results [P1,P2,...,Ps]. min([P1,P2,...,Ps]) is the finally segmentation result of this pixel.

B. 3D FRACTAL RESIDUAL STRUCTURE

It is well known that liver tumor segmentation is considered as a challenging task. First, the morphology of liver tumor varies widely in patients, such as size, shape, location and numbers. Second, some tumor boundaries are fuzzy, which is similar to blood vessels and surrounding tissues in morphology. Therefore, we propose a 3D fractal residual (FR) structure inspired by the idea of the fractal structure. The original fractal network, due to the random discarding mechanism, has improved the generalization ability of the network, but has also led to the discarding of many effective features. In order to increase the generalization ability of the network as well as acquire more features of different resolutions, we add the shortcut connection in the FR structure. By means of building a deep network, the FR structure enlarges the width of the network, expands the dimension of the features extracted by the network, realizes the reuse of the features, and greatly improves the ability of the network to classify tumor candidates. FR structure can be expanded iteratively to an i-level structure, which is expressed as follows,

$$M_i = \begin{cases} res, & \text{if } i = 1 \\ 2M_{i-1}\ LM1\ Lz, & \text{otherwise} \end{cases} \quad (5)$$

where res represents residual structure. There are two types of residual structure as shown in Figure 4, We use residual structure (b) instead of residual structure (a) in order to reduce parameters. Mi represents the i-level FR structure, which is the adjacent structure of Mi−1. In order to illustrate the FR structure in detail, the structure of M3 is presented in Figure 5 as an example. z is the input and ⊕ is a join operation. The

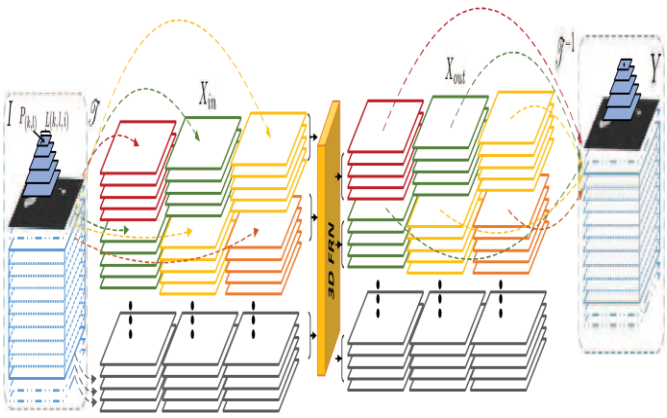

FIGURE 3. The detailed generate process T and multi-scale fusion process T 1.

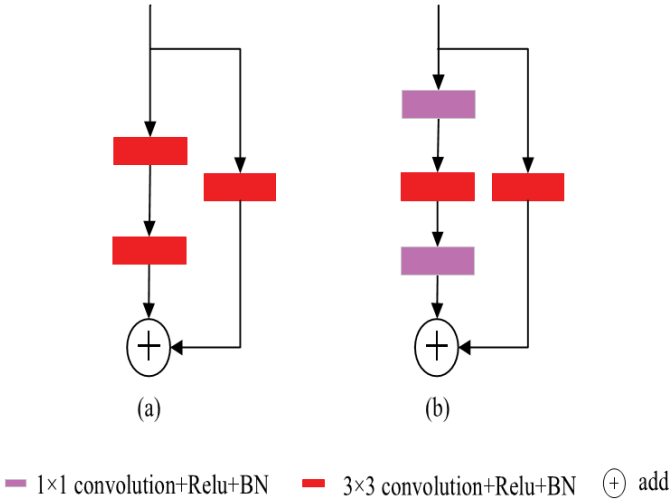

FIGURE 4. Two types of residual structure. (a) is the convolutional residual block with two-layer learning unit and (b) is the convolutional residual block with three-layer learning unit.

joint operation is to merge three globs into a single output glob. The three globs are obtained by residual block, Mi, and the original input. These globs represent three types of features in visual levels, which are extracted to increase the classification efficiency of network. In addition, because of the existence of random dropping paths in the join operation, the back propagation process randomly enhances each path, thereby, improving the overall performance of the network. Then, each path of the FR structure is a single resnet in the process of back propagation. resnet introduces the input of each layer of the network to the later layers by introducing a characteristic shortcut connection, thus making the gradient flow uninterrupted, allowing parameters to be updated in depth in the network, and making the convergence speed faster in the course of training. The FR structure, which combines the advantages of fractal structure and residual structure, reduces overfitting and improves the accuracy of classification with a continuous improvement in learning ability.

C. 3D FRN WITH MULTI-CLASSIFIER

The 3D RFN is composed of b blocks which combine global information with detail information. Each block is a Mi structure. Then, the 3D FRN includes b × 4 × layers. In this paper, we build a 3D RFN with 112 layers (b = 4, i = 3) which includes 3D convolution layer, 3D maxpooling layer and predict layer. The structure of 3D FRN is shown in Figure 6. In [31], it is shown that for liver tumor analysis, the network stacking a large number of small convolutional kernels provides the same sensitivity field effect as the network with a large convolution kernel, but small convolution kernel reduces the computational cost of training network. To this end, the shape of all convolution kernels in 3D FRN are 3 × 3 × 3.

In order to increase the classification ability of the network, some auxiliary classifiers are added after the output of each block. This method can effectively alleviate the problem of gradient disappearance and assist the training process with direct supervision on the hidden layers.

Although the 3D FRN and the fractal network use the similar structure, there are major differences. First, 3D FRN combines with shortcut connection and randomly dropout paths and we extend the 3D FRN by replacing a single layer with a larger fractal residual block, so that the features of any layer can be transferred to any subsequent layer or discarded. In this way, the loss of a large amount of detailed information due to the high level of hierarchy (for example, fast R-CNN [32] and faster R-CNN [33] both use the features of the last convolution layer for target detection) can be avoided. Furthermore, this greatly increases the variety of features extracted by 3D FRN, either the detailed information of the lower level or the semantic features of the higher level. Second, the back propagation mode of the 3D FRN is different from the fractal network. The 3D FRN can reach a much deeper extent than the fractal network because of the existence of multiple shortcut connections during the back propagation process. Regardless if all paths are reserved or only one path is reserved, the shortcut connection can be used to ensure that the gradient can be passed back, ensuring the efficiency of back propagation and speeding up.

D. ACTIVE CONTOUR MODEL

The preliminary result can be experimentally obtained through aforesaid methods as shown in Figure 7. However, the data generated based on MCG leads to superfluous information. Consequently, in the boundary of the tumor, many candidate regions contain only part of tumors because the border of candidates overlap with tumor edges, but they tend to be classified as tumor regions due to the generalization

(a) (b) (c)

FIGURE 5. The preliminary result. (a). The green regions are the predict of liver tumor by 3D MCG-FRN. (b). The red regions are the ground truth and (c). The yellow regions denote correctly tumor segmentation we predict while the green ones for tumor false positive pixels and the red ones for tumor false negative pixels.

process of 3D FRN. As a result, the final segmentation result will be larger than the real tumor. In addition, even though multi-scale superpixel can narrow the burr boundaries to some extent, the smoothness of segmentation boundary still needs improving since the final result is a superpixel-level result. Therefore, the active contour model [27] is applied as a simple post-processing procedure to finetune the obtained boundary since 3D FRN could provide an excellent initial boundary. In this way, a secondary correction step will be performed on candidate regions. Meanwhile, the burr boundary problem can be effectively solved.

III. EXPLANATION

In this paper, a VGG16 Convolutional Neural Network has been proposed for liver tumor detection and segmentation. The system involves a training phase and a testing process for each neural network. The collected CT data has improved through some methods known as data augmentation during the training phase. In the neural network system, the enhanced knowledge, called input data, is then entered to obtain a qualified framework. In our feature extraction process, the testing of various layers of CNNs has tried to find a better feature extraction network. To address the limitations that are not entirely used by modern spatial 3d knowledge in the identification of neural networks. In the Proposal process for the field, the suggestions have been drawn from a pyramid structure to capture different sizes of lesions. The plan is called ROI. Whereas In this step, a texture classifier has been established to distinguish ROIs into normal and abnormal hepatic lesions. The abstract functionalities have been utilized for separating hepatocellular carcinoma (HCC), liver cysts, and hemangiomas irregular hepatic lesions at the classification detection stage. During the training phase for this project, to conducted several iterations to achieve a better model structure. In the testing stage, eventually tested the system based on the results using another batch of CT imaging. Figure 3 shows the pipeline for detecting liver lesions.

The study used CT images of the liver over three times (enhanced, arterial, and delayed, non-contrast agent) in a clinical retrospective study. The form of the lesion cannot be determined by a mass that only occurs in a given period. Therefore, a diagnosis must be validated in comparing the difference in contrast injection between times. In the extraction process, a range of traditions has been used, well-represented in the field of computer vision extraction functions. As the number in the deep-learning network has a significant impact on final classification and recognition performance, the usual way of designing the network is as deep as possible as a false positive reduction of lung CT nodule classification CNN-based classification system. A 2D CNN-based classification system has been developed for classifying the lung nodule candidates as positive lung nodules and

negative non-nodules to reduce the false positives of the initially identified lung nodule candidates.

Dataset:

ImageNet is a dataset of over 15 million labeled high-resolution images belonging to roughly 22,000 categories. The images were collected from the web and labeled by human labelers using Amazon's Mechanical Turk crowd-sourcing tool. Starting in 2010, as part of the Pascal Visual Object Challenge, an annual competition called the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) has been held. ILSVRC uses a subset of ImageNet with roughly 1000 images in each of 1000 categories. At all, there are roughly 1.2 million training images, 50,000 validation images, and 150,000 testing images. ImageNet consists of variable-resolution images. Therefore, the images have been down-sampled to a fixed resolution of 256×256. Given a rectangular image, the image is rescaled and cropped out the central 256×256 patch from the resulting image.

IV. CONCLUSION

Computed Tomography (CT) scans are a widely used imaging modality in global healthcare systems owing to their fast acquisition time and high-resolution 3D rendering of the human body. These scans are read by radiologists who spend hours each day visually inspecting and interpreting the images, to provide a report that will be used to guide clinical care. Initially, lesion detectors will likely serve as an aid to the radiologist to highlight potential lesions that they may have missed. This will augment the radiologist's ability to process large volumes of images and scale to the volumes required for our growing healthcare system, without a simultaneous decrease in accuracy. In this work we explored three model architectures for automated liver lesion localization in CT scans. Our custom-built baseline model and transfer learning model with VGG feature extraction both were able to overfit to the training dataset, but failed to generalize on the test set. The Faster RCNN architecture performed comparably to reference literature when taking into account false positive rate. In the future we would like to explore the tradeoff between Faster RCNN sensitivity and higher false positive rate. We would also explore the effect of using different feature networks for Faster RCNN such as VGG-16. Finally, we would expand our analysis to include lesions from all parts of the human body, not just the liver.

V. REFERENCES

[1]. L.-H. Guo, D. Wang, Y.-Y. Qian, X. Zheng, C.-K. Zhao, X.-L. Li, X.-W. Bo, W.-W. Yue, Q. Zhang, J. Shi, and H.-X. Xu, "A two-stage multi-view learning framework based computer-aided diagnosis of liver tumors with contrast enhanced ultrasound images," Clin. Hemorheology Microcirculation, vol. 69, no. 3, pp. 343-354, Jun. 2018.

[2]. E. Vorontsov, M. Cerny, P. Régnier, L. Di Jorio, C. J. Pal, R. Lapointe, F. Vandenbroucke-Menu, S. Turcotte, S. Kadoury, and A. Tang, "Deep learning for automated segmentation of liver lesions at CT in patients with colorectal cancer liver metastases," Radiol. Artif. Intell., vol. 1, no. 2, Mar. 2019, Art. no. 180014.

[3]. P. F. Christ, M. E. A. Elshaer, F. Ettlinger, S. Tatavarty, M. Bickel, P. Bilic, andW. H. Sommer, "Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields," in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, Oct. 2016, pp. 415-423.

[4]. P. Hu, F.Wu, J. Peng, P. Liang, and D. Kong, "Automatic 3D liver segmentation based on deep learning and globally optimized surface evolution," Phys. Med. Biol., vol. 61, no. 24, pp. 8676-8698, Dec. 2016.

[5]. K.Wu, X. Chen, and M. Ding, "Deep learning based classification of focal liver lesions with contrast-enhanced ultrasound," Optik, vol. 125, no. 15, pp. 4057-4063, Aug. 2014.

[6]. M. Bellver, K.-K. Maninis, J. Pont-Tuset, X. Giro-i Nieto, J. Torres, ´ and L. Van Gool, "Detection-aided liver lesion segmentation using deep learning," arXiv preprint arXiv:1711.11069, 2017.

[7]. X. Han, "Automatic liver lesion segmentation using a deep convolutional neural network method," arXiv preprint arXiv:1704.07239, 2017.

[8]. X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P.-A. Heng, "H-denseunet: hybrid densely connected unet for liver and tumor segmentation from ct volumes," IEEE transactions on medical imaging, vol. 37, no. 12, pp. 2663–2674, 2018.

[9]. H. Jiang, T. Shi, Z. Bai, and L. Huang, "Ahcnet: An application of attention mechanism and hybrid connection for liver tumor segmentation in ct volumes," IEEE Access, vol. 7, pp. 24 898–24 909, 2019.

[10]. Y. Bengio, J. Louradour, R. Collobert, and J. Weston, "Curriculum learning," in Proceedings of the 26th annual international confer-ence on machine learning, 2009, pp. 41–