Deep Learning for Joint Classification and Segmentation of Histopathology Image

Hyun-Cheol Park[1], Raman Ghimire[1], Sahadev Poudel[1], Sang-Woong Lee[2*]

[1] Department of IT Convergence Engineering, Gachon University, Republic of Korea
[2] Division of Software, School of Computing, Gachon University, Republic of Korea
h.fe.park@gmail.com, ghimirermn@gmail.com, sahadevp093@gmail.com, slee@gachon.ac.kr

Abstract

Liver cancer is one of the most prevalent cancer deaths worldwide. Thus, early detection and diagnosis of possible liver cancer help in reducing cancer death. Histopathological Image Analysis (HIA) used to be carried out traditionally, but these are time-consuming and require expert knowledge. We propose a patch-based deep learning method for liver cell classification and segmentation. In this work, a two-step approach for the classification and segmentation of whole-slide image (WSI) is proposed. Since WSIs are too large to be fed into convolutional neural networks (CNN) directly, we first extract patches from them. The patches are fed into a modified version of U-Net with its equivalent mask for precise segmentation. In classification tasks, the WSIs are scaled 4 times, 16 times, and 64 times respectively. Patches extracted from each scale are then fed into the convolutional network with its corresponding label. During inference, we perform majority voting on the result obtained from the convolutional network. The proposed method has demonstrated better results in both classification and segmentation of liver cancer cells.

Keywords: Histopathological image analysis, Whole-slide image, Segmentation, Classification, Patch-based method

1 Introduction

Hepatocellular Carcinoma (HCC) is the fourth most leading cause of cancer and is currently the biggest cause of liver-related death [1], constituting a major global health problem. Early detection and treatment can thus remarkably improve the likelihood of survival. Imaging techniques such as histopathological image analysis (HIA) [2] are considered ideal standard for cancer detection. Unfortunately, pathological analysis is a tedious and intense process that needs proficient knowledge. Thus, there is an increased need for processes that can automatically classify and segment the malignant cells in whole slide images.

Computer-aided automatic image analysis of liver cancer has become feasible with the advancement in computer vision and pattern recognition techniques. Compared to the traditional methods based on hand-engineered features, computer-assisted (CAD) helps lessen the burdens of pathologists. The current computerized approach, however, has a myriad of obstacles. First, the data extraction procedure is costly in the medical domain [3] and usually amounts to class imbalance problems. Second, generated whole slide image (WSI) is of very high resolution scale, generally up to 40 times magnification on a microscope. Third, the extracted information is difficult to generalize form medical point of view. And finally, the selection of features cannot be dynamically optimized with the change in dataset.

The development of CAD allows the automatic extraction of features from low-to-high levels with the usage of convolutional neural networks. Such a network performs at low cost and requires no manual feature design and can recognize objects with utmost accuracy. Over the years, various deep learning approaches have been proposed and are still being developed. For example, classification architectures [4-6], segmentation architectures [7-9], detection architectures [10] and so on. Applying such models, different types of histopathological image analysis have been carried out, such as melanoma detection [11], breast cancer classification [12-13], cancer cell segmentation [14-15]. However, HIA still faces some challenges. First of all, HIA fails to generalize complex clinical features. Secondly, the training data is insufficient. And finally, the WSI size is too large, in the range of 50,000x50,000 pixels, it is challenging to annotate and perform analysis on it accurately.

To mitigate the above-mentioned problems, various techniques have been proposed. Among them, patch-based techniques are very popular and still in use. A whole slide image is cropped into various patches, and such patches are fed independently into a deep learning model during training and inference. For the prediction and result during classification and segmentation, the patches are aggregated in a suitable order. [16] and [17] showed compelling results based on patch-based methods for classification and segmentation tasks, respectively. In [16], two different modes of classification are proposed - one patch in one decision (OPOD) and all patches in one decision (APOD).

In OPOD mode, for correct classification, all the extracted patches need to have same class label that matches with the ground truth. In APOD mode, final decision about class label is taken based on the majority voting among the classified patches. In [17] a U-net based approach to tissue segmentation in whole slide histopathological image is carried out. Patches of size 512x512 pixels are extracted, augmented, and fed to the U-Net network. Many approaches have been applied to solve the data imbalance and insufficiency issue. For instance, [18] proposed a classification of histopathology images using only global labels. Augmentation methods have also been

applied to increase the size and diversity of the training dataset. [19] uses efficient image augmentation methods to alleviate the demand for a large amount of data and improve the network's generalization capacity.

Given the complexities of liver cancer histopathological images and the insufficiency in training dataset, we propose a framework developed for analyzing whole slide images using patch-based methods. We have listed the advantages of proposed method as follows:

1. Efficient analysis of image – A typical whole slide image (WSI) can contain more than 50,000×50,000 pixels; utilizing patch-based methods, it can be divided into several patches and analyzed individually.
2. Training data insufficiency can be solved by utilizing the augmentation approach. We follow a method where each patch is further augmented into several other patches and thus, we have an increased number of training data.
3. We utilize a majority voting scheme that robustly estimates the cancerous or non-cancerous liver tissue.

The organization of this paper is as follows. Section 2 consists of related work. Section 3 consists of methodology where dataset, network architecture, preprocessing, classification, and segmentation is discussed. Section 4 consists of performance measure of classification and segmentation models. Finally, the conclusion is presented in section 5.

2 Related Work

2.1 Deep Learning Methods for Histopathological Image Classification

The advent of convolutional neural networks has accelerated histopathological image analysis. With the usage of CNNs, the traditional hand-engineered features are no longer required, and features can be automatically easily extracted. However, CNN's require a lot of data to generalize on them, and the medical domain suffers from dataset insufficiency.

Especially in the medical domain, processes for tackling the improvement in classification performance and dataset insufficiency were required. [18] uses global labels instead of patch-level labels to solve the problem of insufficient training dataset. They utilize transfer learning techniques to mitigate the issue of insufficient histopathological images. [20] uses augmentation strategy to have enough diverse samples. Augmentation also helps the network to generalize on the dataset.

2.2 Deep Learning Methods for Histopathological Image Segmentation

Due to the complexity of shape, texture, overlapping of cells, scanning devices, automatic segmentation is still an arduous task in deep learning. However, compared to manual methods, the deep learning methods can handle images with increased complexity.

[21] uses a two-stage learning method based on U-Nets to solve the challenging issue on Nuclei segmentation. This two-stage method is used to solve the overlapping nuclei issue problem. Here, in the first stage, the nuclei image is segmented, and in the second stage, the overlapping nuclei image are segmented. Nucleus instance segmentation results from the first stage would be updated by addition of overlapped regions obtained from the second stage. [22] proposes an ensemble method where the sequence of techniques in the preprocessing-training-inference pipeline takes place. In preprocessing, they divide the WSI into patches to address the class imbalance issue; for training and inference, they use encoder-decoder architecture and follow it with a pixel-wise classification layer. [23] also uses the patch-based method; however, they allow 25% overlap during patch extraction. The extracted patches would then be passed to the convolutional neural network for prediction of possible tissue classes. After that, pixel-level segmentation is used to predict pixel-level activation maps before the extracted features of neighboring patches are averaged at overlapping areas. At the post processing stage, the segmented section is stitched together back to the slide level.

2.3 Deep Learning Methods for Histopathological Image Classification and Segmentation

The development of precise and efficient network for both classification and segmentation of tissue image is still a challenging problem because of the variations of tissue shapes and textures. Thus, algorithms must generalize on the tissue heterogeneity and learn to adapt to changes in tissue morphology.

[24] trains two convolutional networks, one to identify the nuclei blob and other for boundary detection. The CNNs consists of encoder-decoder architecture and extract nuclei pixels and nuclei boundary pixels and thus generating nuclei blob and border mask, respectively. The segmentation process is carried out by removal of nuclei boundaries and separation of clumped nuclei. They use ResNet-32 for training the deep neural network and use random forest regression model for classification purpose. In [25] authors have proposed a nuclear classification and segmentation network. Here, the nuclear pixels are first detected, and then with a post-processing method, nuclear instances are simultaneously segmented. With the segmented nuclear instances, corresponding nuclear type is also obtained. The overall network is based on horizontal and vertical distance maps.

3 Methodology

3.1 Dataset Description and Experimental Settings

The dataset of liver cancer WSI was taken from the PAIP 2019 challenge [26]. In the challenge, the participants were provided with 2 levels of dataset extracted from WSIs i.e., tumor with peritumoral reaction and tumor with minimal peritumoral or intratumoral reaction. The data and its segmented part were provided by Seoul National University

Hospital, Korea[1]. The training data contained 50 WSIs. Each image has an average dimension of 50,000x50,000 pixels.

The histopathological images were split into three datasets of ratios 0.6,02 and 0.2, corresponding to training, validation, and testing datasets respectively. Since the histopathological images were too big to be directly fed into the convolutional neural networks, they were transformed into patches of 224x224 pixels. Patches containing white background would be removed. In the classification stage, the whole slide images (WSI) were scaled three times. In the first step, 4 times, in the second step, 16 times, and in the last step, 64 times. 4 times scaled WSIs on average would have 50,000 images for each of positive and negative classes. The class imbalance among the dataset was mitigated through the augmentation process. Augmentation methods such as rotation, flipping, blurring, scaling and shifting was carried out. The 16 times scaled WSIs in average, would have 20,000 images for each of positive and negative classes, whereas the 64 times scaled WSIs in average would have 10,000 images for each of positive and negative classes.

In the case of the segmentation stage, unlike in the classification process, the histopathological images were not scaled and only patches each of 224x224 pixels were extracted.

3.2 Network Overview

An overview of the proposed network can be seen on Figure 2. It is a multi-step, multi-stage network. The whole slide images are transformed into individual patches of 224x224 pixels in the first step. Such patches are then fed into the encoder-decoder network. The individual segmented patches are then recombined and presented as result. In the next step, every whole slide image is scaled three times, i.e., 4 times, 16 times and 64 times. Individual patches are extracted from all three of the scaled images and fed into the classification model separately. The result from all 3 training models is then aggregated, and a majority voting scheme is applied after which we obtain the class label of the image.

In Figure 1 we can see the portion of the scaled images. In 4 times scaled image, different shapes, textures, color distribution, are easy to quantify and read. On 16 times scaled image, the quality on scaled image gradually decreases compared with the prior scaled image. The texture, color distribution is slightly difficult to generalize. As we scale the WSIs up to 64 times, the image sharpness, color distribution loses its edge. The image is difficult to generalize. The texture, shape and color distribution are hazy.

3.3 Pre-Processing

As discussed earlier, the whole slide images (WSIs) are too large to be processed by convolution networks as a whole; thus, small patches need to be extracted from them. We also have to separate the tissue region from the background glass region of the whole slide image. For that, first, we extract patches without overlap using the patchify python package [27]. After the patches are extracted, we calculate the optimal threshold for each patch and remove the patches which contain white background. This white background denotes the background glass region. Segmenting them from the tissue region prevents unnecessary computational tasks. This step is repeated every time the WSIs are scaled for classification tasks. Finally, the patches which contain tissue regions are obtained and fed into the network for classification and segmentation tasks.

3.4 Segmentation

The proposed segmentation network can be seen in Figure 3. The segmentation network takes influence from UNet++ [28] network. It is a deeply-supervised encoder-decoder network where the encoder-decoder sub-networks are connected through a series of nested, dense skip connections. Unlike UNet++, we have used EfficientNet-B0 [29] as a feature extractor in the encoding path. Each node receives features as inputs from the previous node as well as the upsampled output of the lower dense block as seen on Figure 3 which is then concatenated and convolved. Similarly, each node also receives multiscale features at varying scale to aggregate diverse feature maps. Furthermore, we reduce the size of feature map by 2 at each level. On the decoding section, we have used transposed convolution to match the size of feature map with corresponding encoding section. 1*1 convolution followed with a sigmoid function is applied on output of each decoding block.   Finally, the output of the end layer and each decoding block with deep supervision [29] is averaged to obtain segmentation result.

| 4 times scaled | 16 times scaled | 64 times scaled |

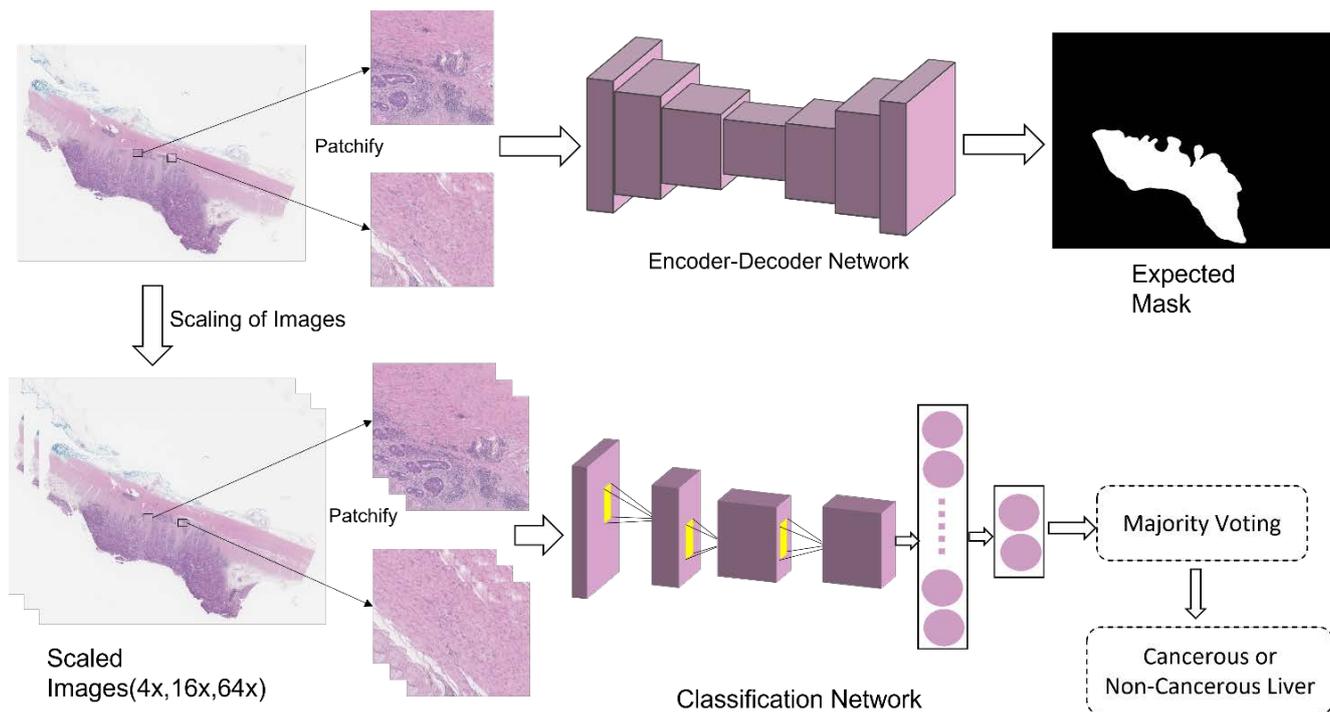**Figure 1.** Scaled WSI for classification. From left, 4 times, 16 times and 64 times scaled image

Figure 2. Proposed method for classification and segmentation of liver cancer histopathological images
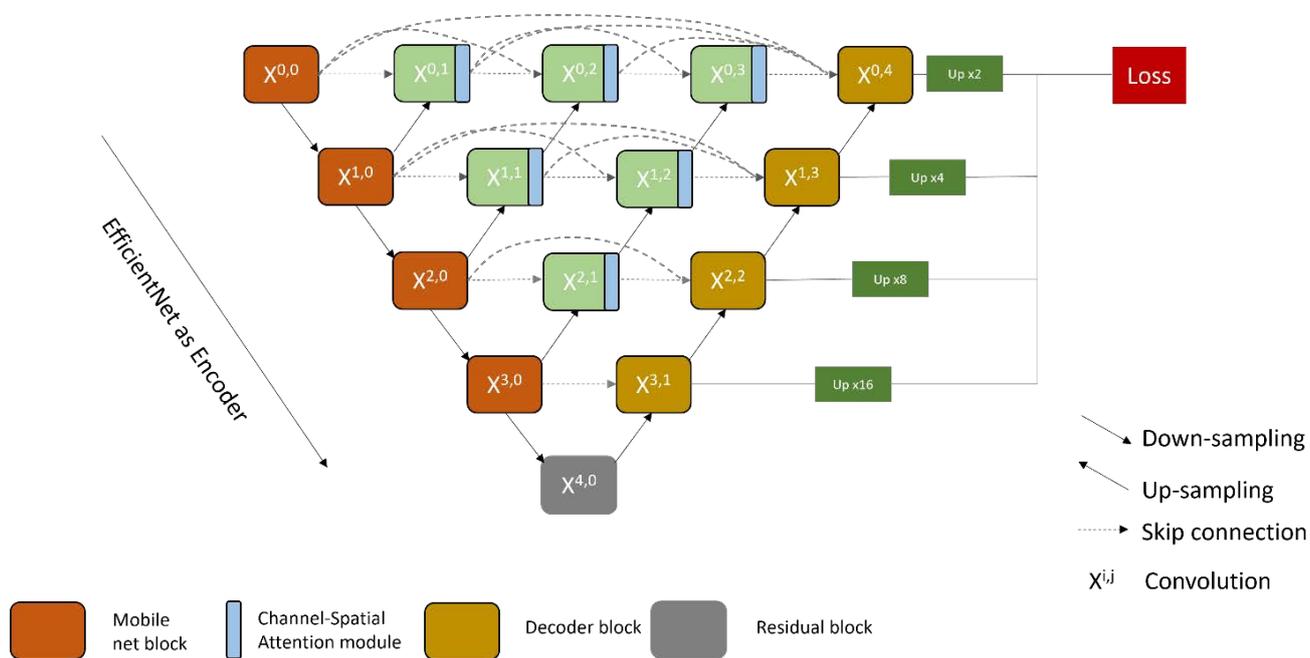

Figure 3. Proposed modified U-Net architecture for whole slide image segmentation

Loss function: The section to be segmented on the whole slide images is minuscule compared to the original image. This leads to class imbalance. So, we needed to choose a loss function that would have the potential to tackle this issue. The dice score coefficient is an overlap metric used to assess segmentation performance when ground truth is available. Dice loss [30] is a differentiable function that is defined using the predicted probability map and calculates a range from 0 to 1 where 1 denotes perfect and complete overlap. Dice loss approximates dice score coefficient and is defined in the equation as shown below:

$$D_1 = \frac{\sum_n^N p_n r_n}{\sum_n^N P_n^2 + \sum_n^N r_n^2} \tag{1}$$

where Dl represents dice loss, p_n and r_n denotes corresponding pixel value of predicted probability and ground truth.

We have also used Jaccard Index [31] as a metric, also known as Intersection over Union, is one of the most used metrics in image segmentation. It defines the area of overlap between predicted segmentation and ground truth over the section of union between predicted segmentation and ground truth. Jaccard Index signifies similarity between sample sets. It is represented in the range of 0 to 100% (0-1), 0 denotes no overlap and 1 denotes complete overlap between predicted segmentation and ground truth. It is expressed as follows:

$$ J\,(A,\,B) = \frac{|A \cap B|}{|A|+|B|-|A \cap B|} \tag{2} $$

where A and B represents area of segmented region and ground truth.

3.5 Classification

For the classification, we have compared the results on two different classification models. At first, we have used EfficientNet-B0 [29]. This network is pre-trained on ImageNet dataset, and we fine-tune its final layers. We have used stochastic gradient descent [32] as an optimization method. The network is trained with a batch size of 64 and the learning rate is set at 0.0001 which is multiplied by 0.1 every time the patience level crosses 8. We have also used weight decay which is set at 0.0001 and momentum which is set at 0.9. The second classification model we used is ResNet-50 [33]. It is also pre-trained on ImageNet. We use Adam Optimizer [34] with a batch size of 64. The learning rate starts from 0.000001 and when a metric stops improving it is reduced by a factor of 0.1. The patience is set at 25 with number of epochs as 200.

All patches are aimed at classifying images into two classes based on a majority voting scheme on its predicted patch labels. Let P, a = {1,2, ...n} be the patch extracted from the whole slide image I and L represent the class label of patch P such that l = [0,1]. The result is counted as,

$$ C_i^y = \begin{cases} 1 \; if \; l = t \\ 0 \; if \; l \neq t \end{cases} \tag{3} $$

where t = [1,2] corresponds to accurate class labels. The total count can be expressed as:

$$ f_i = \sum_{j=0}^{n} C_i^y \tag{4} $$

The final label is one that majority of the individual classifiers agrees. The result can be simply expressed as:

$$ J = argmax|f_i| \tag{5} $$

4 Implementation and Experimental Results

We employed 29 whole slide images for the training set, 9 images for validation, and 9 images for testing purposes. From each of the whole slide image, patches and their corresponding mask is generated -. On training, taking the mask as a reference, we remove the white background region of the WSIs. For that, we extracted patches from the WSIs, compared the threshold of color with its corresponding mask, and removed it if it exceeded its threshold. On inference as well, the network first extracts the patches from the whole slide images. The extracted patches are fed through the network to obtain the segmentation maps. These individual maps are then aggregated to get the result and can be seen on Figure 4. Since patchification and un-patchification takes place, there are certain noises in the result. We can see that the patch-based method requires generality features in the patches for them to be equally efficient. And if such a case is not possible, then it might be better to use to convolutional networks that take entire image as input and produce result at a single forward pass. We obtained a Dice score of 0.69 and a Jaccard index of 0.68. The performance result can be seen in Figure 4. The predicted image is closer to the ground truth, however, there does exist certain noises and inconsistencies in the boundary.
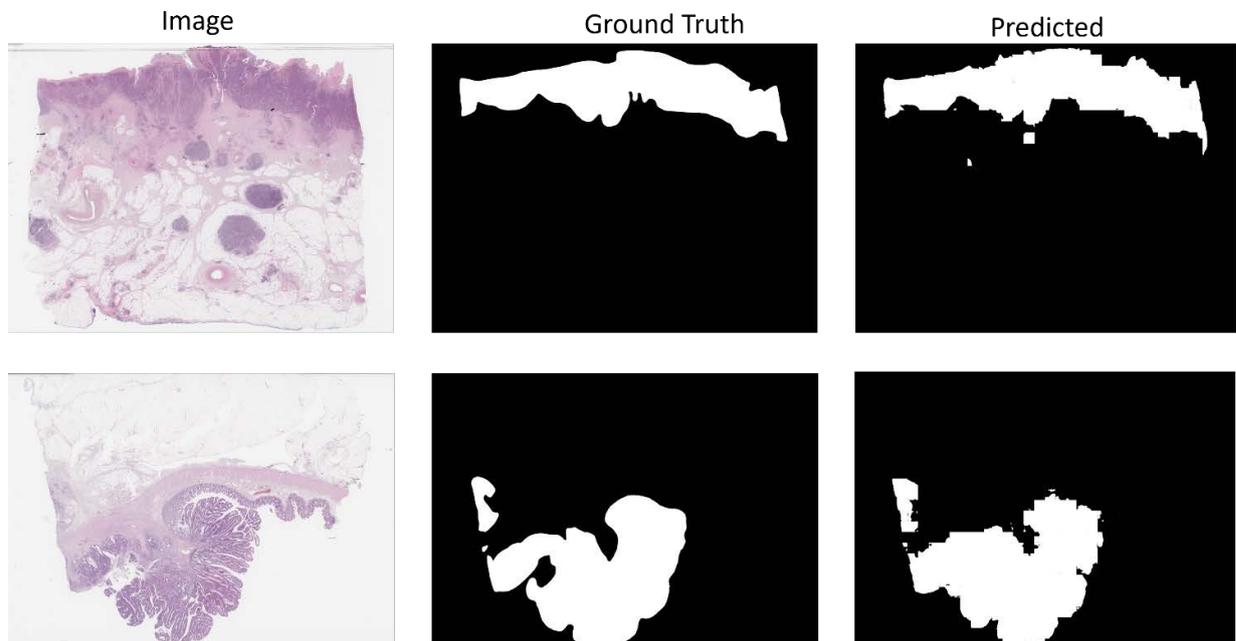

Figure 4. Result of the segmentation process

For better results, we have used two different models for classification purposes. These models are ResNet-50 and EfficientNet-B0. We have adapted a majority-voting scheme to obtain the final results. For that reason, as discussed earlier, the whole slide image is scaled at 3 different scales, i.e., at 4 times scaled, at 16 times scale and 64 times scaled. Patches of 224x224 pixels would then be extracted from each scaled image and fed into two different models. At total each of the classification model had to be trained 3-3 times respectively. The result from 4 times, 16 times and 64 times scaled would then be sent through the majority voting scheme. Both models performed well. We used 9 WSI for testing purpose and all the images were accurately classified. The training loss plot can be seen on Figure 5. As seen from the figure, all of the classification models stop learning at around 180 epoch. The training beyond that is to just confirm the model saturation. The final result of each classification model can be seen on Table 1. We see better performance of both models on 4 times scaled images. And the learning accuracy seems to gradually decrease with the scale. Among the models, EfficientNet-B0 outperforms ResNet-50 at all scales. ResNet-50 is 50 layers deep with constant convolutional and max-pooling layers. The added skip connection helps in retaining previous information and provides an alternative path for the gradient. Contrary to ResNet-50, EfficientNet-B0 uses clever scaling depth, width, and resolution. This helps in improving accuracy, training time, and convergence as evident in Table 1.
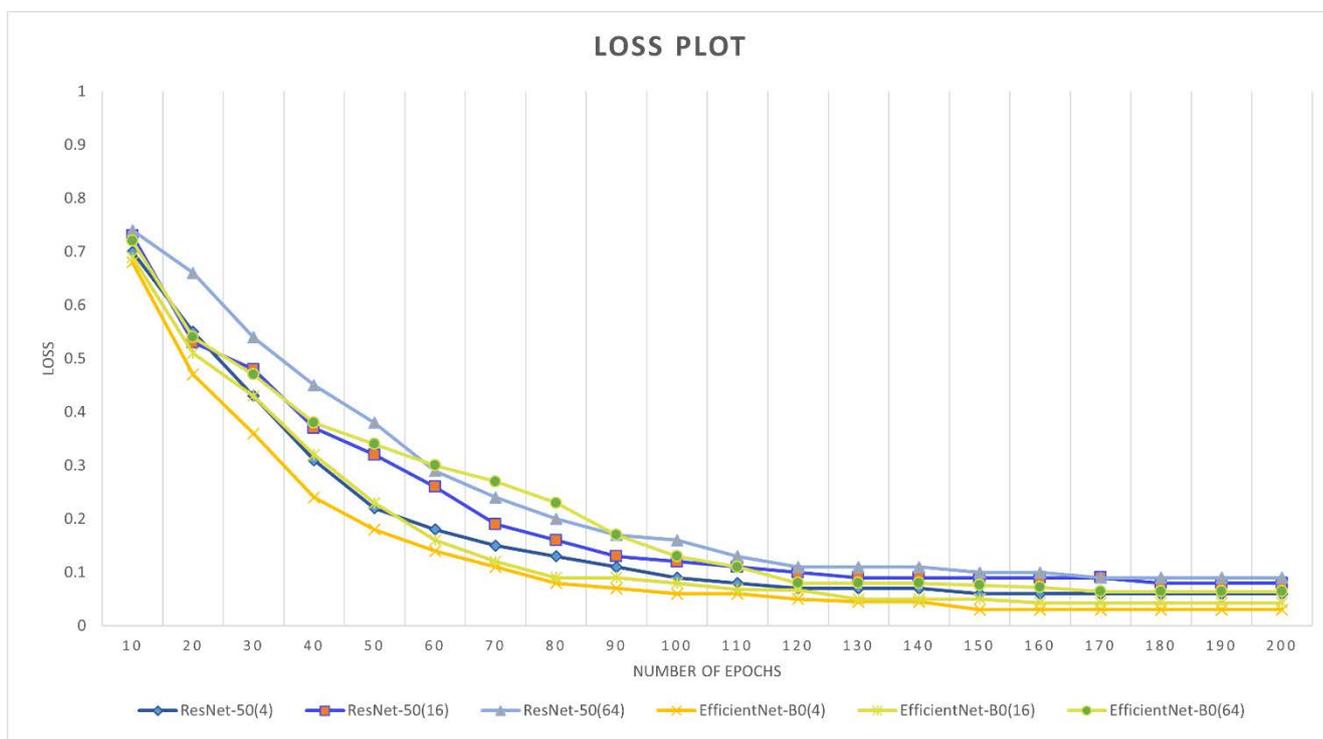

Figure 5. Training Loss plot of 6 different classification approach (Bracketed term means the scaled image parameter)

Table 1. Learning accuracy of ResNet and EfficientNet model

Scale ratio	ResNet-50	EfficientNet-B0
4	93.31%	96.2%
16	89.8%	92.8%
64	86.21%	89.21%

5 Discussion and Conclusion

In this paper, we have carried out both segmentation and classification of liver cancer in histopathological images. We have utilized patch-based methods for both segmentation and classification tasks and obtained compelling results. In the case of classification task, we have scaled the images 3 times and determined the result based on the majority score. The proposed method shows that the model can accurately classify the liver histopathological images into cancerous or non-cancerous type. For segmentation task, we have used modified version of traditional U-Net which produced exemplary results.

Our approach reduces the need for large training corpus annotated by expert pathologists. The usage of augmentation approach has helped in increasing the training dataset as well as adding regularization in the network. Prior methods have usually focused on either of segmentation or classification task, hence there exists a lack of relevant studies on a combined network that can accurately perform segmentation as well as classification. This also leads to lack of comparative results. To that end, our proposed method can accurately classify and segment histopathology images.

Acknowledgment

References

[1] A. Ally, M. Balasundaram, R. Carlsen, E. Chuah, A. Clarke, N. Dhalla, R. A. Holt, S. J. Jones, D. Lee, Y. Ma, M. A. Marra, Comprehensive and integrative genomic characterization of hepatocellular carcinoma, *Cell*, Vol. 169, No. 7, pp. 1327-1341, June, 2017.

[2] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, B. Yener, Histopathological image analysis: A review, *IEEE reviews in biomedical engineering*, Vol. 2, pp. 147-171, October, 2009.

[3] D. Shen, G. Wu, H. I. Suk, Deep learning in medical image analysis, *Annual review of biomedical engineering,* Vol. 19, pp. 221-248, June, 2017.

[4] W. Wang, D. Liang, Q. Chen, Y. Iwamoto, X. H. Han, Q. Zhang, H. Hu, L. Lin, Y. W. Chen, Medical image classification using deep learning, in: Y. W. Chen, L. Jain (Eds.), *Deep Learning in Healthcare*, Springer, Cham, 2020, pp. 33-51.

[5] K. Kowsari, R. Sali, L. Ehsan, W. Adorno, A. Ali, S. Moore, B. Amadi, P. Kelly, S. Syed, D. Brown, Hmic: Hierarchical medical image classification, a deep learning approach, *Information*, Vol. 11, No. 6, Article No. 318, June, 2020.

[6] S. Poudel, Y. J. Kim, D. M. Vo, S. W. Lee, Colorectal disease classification using efficiently scaled dilation in convolutional neural network, *IEEE Access*, Vol. 8, pp. 99227-99238, May, 2020.

[7] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, *International Conference on Medical image computing and computer-assisted intervention*, Munich, Germany, 2015, pp. 234-241.

[8] S. Poudel, S. W. Lee, Deep multi-scale attentional features for medical image segmentation, *Applied Soft Computing*, Vol. 109, Article No. 107445, September, 2021.

[9] H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y. W. Chen, J. Wu, Unet 3+: A full-scale connected unet for medical image segmentation, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, 2020, pp. 1055-1059.

[10] Q. Dou, H. Chen, L. Yu, L. Zhao, J. Qin, D. Wang, V. C. Mok, L. Shi, P. A. Heng, Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks, *IEEE transactions on medical imaging*, Vol. 35, No. 5, pp. 1182-1195, May, 2016.

[11] S. H. Kassani, P. H. Kassani, A comparative study of deep learning architectures on melanoma detection, *Tissue and Cell*, Vol. 58, pp. 76-83, June, 2019.

[12] Z. Han, B. Wei, Y. Zheng, Y. Yin, K. Li, S. Li, Breast cancer multi-classification from histopathological images with structured deep learning model, *Scientific reports*, Vol. 7, No. 1, pp. 1-10, June, 2017.

[13] A. Golatkar, D. Anand, A. Sethi, Classification of breast cancer histology using deep learning, *International Conference Image Analysis and Recognition*, Póvoa de Varzim, Portugal, 2018, pp. 837-844.

[14] N. Hatipoglu, G. Bilgin, Cell segmentation in histopathological images with deep learning algorithms by utilizing spatial relationships, *Medical & biological engineering & computing*, Vol. 55, No. 10, pp. 1829-1848, October, 2017.

[15] Kurnianingsih, K. H. Allehaibi, L. E. Nugroho, Widyawan, L. Lazuardi, A. S. Prabuwono, T. Mantoro, Segmentation and classification of cervical cells using deep learning, *IEEE Access*, Vol. 19, No. 7, pp. 116925-116941, August, 2019.

[16] K. Roy, D. Banik, D. Bhattacharjee, M. Nasipuri, Patch-based system for classification of breast histology images using deep learning, *Computerized Medical Imaging and Graphics*, Vol. 71, pp. 90-103, January, 2019.

[17] K. R. Oskal, M. Risdal, E. A. Janssen, E. S. Undersrud, T. O. Gulsrud, A U-net based approach to epidermal tissue segmentation in whole slide histopathological images, *SN Applied Sciences*, Vol. 1, No. 7, pp. 1-12, June, 2019.

[18] C. Sun, A. Xu, D. Liu, Z. Xiong, F. Zhao, W. Ding, Deep learning-based classification of liver cancer histopathology images using only global labels, *IEEE journal of biomedical and health informatics*, Vol. 24, No. 6, pp. 1643-1651, June, 2020.

[19] Y. Xiao, E. Decencière, S. Velasco-Forero, H. Burdin, T. Bornschlögl, F. Bernerd, E. Warrick, T. Baldeweck, A new color augmentation method for deep learning segmentation of histological images, *IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, Venice, Italy, 2019, pp. 886-890.

[20] M. H. Motlagh, M. Jannesari, H. Aboulkheyr, P. Khosravi, O. Elemento, M. Totonchi, I. Hajirasouliha, Breast cancer histopathological image classification: A deep learning approach, *BioRxiv*, Article No. 242818, January, 2018.

[21] Y. Kong, G. Z. Genchev, X. Wang, H. Zhao, H. Lu, Nuclear Segmentation in Histopathological Images Using Two-Stage Stacked U-Nets With Attention Mechanism, *Frontiers in Bioengineering and Biotechnology*, Vol. 8, Article No. 573866, October, 2020.

[22] M. Khened, A. Kori, H. Rajkumar, G. Krishnamurthi, B. Srinivasan, A generalized deep learning framework for whole-slide image segmentation and analysis, *Scientific reports*, Vol. 11, No. 1, pp. 1-4, June, 2021.

[23] L. Chan, M. S. Hosseini, C. Rowsell, K. N. Plataniotis, S. Damaskinos, Histosegnet: Semantic segmentation of histological tissue type in whole slide images, *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Seoul, South Korea, 2019, pp. 10662-10671.

[24] Q. D. Vu, S. Graham, T. Kurc, M. N. N. To, M. Shaban, T. Qaiser, N. A. Koohbanani, S. A. Khurram, J. Kalpathy-Cramer, T. Zhao, R. Gupta, J. T. Kwak, N. Rajpoot, J. Saltz, K. Farahani, Methods for segmentation and classification of digital microscopy tissue images, *Frontiers in bioengineering and biotechnology*, Vol. 7, Article No. 53, April, 2019

[25] S. Graham, Q. D. Vu, S. E. Raza, A. Azam, Y. W. Tsang, J. T. Kwak, N. Rajpoot, Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images, *Medical Image Analysis*, Vol. 58, Article No. 101563, December, 2019.

[26] PAIP, *Pathology AI Platform*, URL: http://www.wisepaip.org, 2019.

[27] W. Weiyuan, D. Verma, W. Yang, *Patchify Github Repository*, GitHub n.d, https://pypi.org/project/patchif.

[28] Z. Zhou, M. M. Siddiquee, N. Tajbakhsh, J. Liang, Unet++: A nested u-net architecture for medical image segmentation, *Deep learning in medical image analysis and multimodal learning for clinical decision support*, Granada, Spain, 2018, pp. 3-11.

[29] M. Tan, Q. Le, Efficientnet: Rethinking model scaling for convolutional neural networks, *International Conference on Machine Learning*, Long Beach, California, USA, 2019, pp. 6105-6114.

[30] W. Wang, J. Shen, Deep visual attention prediction, *IEEE Transactions on Image Processing*, Vol. 27, No. 5, pp. 2368-2378, May, 2018.

[31] R. Shi, K. N. Ngan, S. Li, Jaccard index compensation for object segmentation evaluation, *IEEE international conference on image processing*, Paris, France, 2014, pp. 4457-4461.

[32] S. Ruder, *An overview of gradient descent optimization algorithms*, September, 2016. https://arxiv.org/abs/1609.04747

[33] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, *Proceedings of the IEEE conference on computer vision and pattern recognition*, Las Vegas, NV, USA, 2016, pp. 770-778.

[34] D. P. Kingma, J. Ba, *Adam: A method for stochastic optimization*, December, 2014. https://arxiv.org/abs/1412.6980

Biographies

Hyun-Cheol Park received the B.S and M.S degrees in computer engineering from Chosun University, Gwangju, South Korea, and the Ph.D. degree in IT convergence engineering from Gachon University, Seongnam, South Korea, in 2015, 2017, and 2022, respectively. He is currently working as a Research Professor with Department of AI Software, Gachon University, Seongnam, South Korea. His research interests include machine learning, aritificial intelligence, computer vision, and medical imaging analysis.

Raman Ghimire received the B.S. degree in Electrical Engineering from Tribhuwan University, Nepal, in 2018 and M.E degree in I.T Convergence Engineering from Gachon University, South Korea, in 2022. Currently he is working as an AI Research Engineer at Cybermed Inc., South Korea. His research interests include computer vision, image processing and medical imaging analysis.

Sahadev Poudel received the B.S. degree in Information Technology from Purbanchal University, Nepal, in 2016, and the M.E. degree in IT convergence engineering from Gachon University, South Korea, in 2020. Currently, he is pursuing his PhD degree in Gachon University. His current research fields include Medical Image Generalization and Analysis, Federated Learning and Optimization in deep learning.

Sang-Woong Lee received the B.S. degree in electronics and computer engineering and the M.S. and Ph.D. degrees in computer science and engineering from Korea University, Seoul, South Korea, in 1996, 2001, and 2006, respectively. From June 2006 to May 2007, he was a Visiting Scholar with the Robotics Institute, Carnegie Mellon University. From September 2007 to February 2017, he worked as a Professor with the Department of Computer Engineering, Chosun University, Gwangju, South Korea. He is currently a Professor with the Department of Software, Gachon University. His present research interests include face recognition, computational aesthetics, machine learning, and medical imaging analysis.