2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

# A Novel Adaptive Fuzzy Deep Learning Approach for Histopathologic Cancer Detection

Xiankun Yan[1], Jianrui Ding[2] and H. D. Cheng[1,*]

*Abstract*— **We proposed a novel model that integrates the fuzzy theory and group equivariant convolutional neural network for histopathologic cancer detection. The proposed fuzzy group equivariant convolutional neural network consists of the convolutional network, a novel fuzzy global pooling layer, and a fully connected network. In the fuzzy global pooling layer, the generated feature maps are transferred into the fuzzy domain by two different fuzzification methods. One of the fuzzy feature maps exploits the uncertainty information of histopathologic images, and the other keeps the original information. Furthermore, the fuzzy feature maps are processed by using Min-max operations. The experiments verified that the proposed method could always find the maximum fuzzy entropy and exploit and present the uncertainty of histopathologic images well. The experiments using the benchmark dataset demonstrate that the proposed model becomes more accurate and outperforms the existing models including the benchmark models. Compared to the benchmark model with 89.8% of accuracy, 96.3% of AUC, and 0.260 of negative log-likelihood loss, the proposed model obtained 91.7% of accuracy, 97.2% of AUC, and 0.214 of negative log-likelihood loss.**

## I. INTRODUCTION

Recently, high-resolution histopathology images that offer a more detailed overview of the disease have been employed for clinical cancer diagnosis. However, as the 'golden standard' in cancer diagnosing, histopathology image segmentation and classification cost a lot of time and labor, and it is difficult to annotate the images even for pathologists since there is no clear definition and criteria for annotating normal or cancerous regions boundaries. Therefore, to relieve the workload, computer-assisted diagnosis (CAD) was studied to segment and detect cancers in histopathology images [3, 10, 11]. Following the improvement of the deep learning model and computing hardware, convolutional neural networks (CNNs) models and the derivative models had been studied.

However, most of the traditional deep learning strategies did not handle the uncertainty other than random uncertainty well. The epistemic and aleatoric uncertainty in deep learning architecture and medical images was proved in the previous work [2]. To model the uncertainty in the medical images, people began to apply fuzzy logic [7]. Some image enhancement and learning models based on fuzzy logic have been raised for clinical datasets. In [6] and [4], the qualities of computerized tomography (CT) images, ultrasound images, and magnetic resonance imaging (MRI) images were improved by applying the fuzzy logic. And Funmilola *et al.* [1] proposed the fuzzy kc-means clustering algorithm for medical image segmentation and achieved a higher performance.

In this work, we describe a learning model that is trained on patch-level histopathologic images using fuzzy logic. The convolutional component will create feature maps as the inputs for classification tasks. Before using the fully connected network to classify them, the feature maps would be mapped into the fuzzy domain, where the uncertainty was exploited. Furthermore, we also explore the performance of several learning models for the classification tasks to show the improved performance the proposed model.

The main contributions of our work are as follows. To present and handle the uncertainty in the deep learning method and histopathologic images, we propose a novel approach to employ the fuzzy theory. Firstly, the group equivariant convolutional network [9] is applied for extracting the deep feature maps from histopathologic images. Then, the deep feature maps are transformed into a fuzzy domain by two different memberships. In the fuzzy domain, the feature maps keep the max original information and exploit the uncertainty information. To improve the performance of the classifier, combining the original information and the uncertainty of information as a novel input is an important step. Here, the Min-max fuzzy operation is applied to process the feature maps. Moreover, to obtain the maximum uncertainty information by tuning the parameters during fuzzification, the novel global pooling layer is proposed. We add the memberships and fuzzy operations in the global pooling layer. In the experiments, we validate that the proposed model can structure the uncertainty in the histopathologic images by using the fuzzy maximum entropy principle [5] and evaluate its performance on histopathologic cancer detection compared with other baselines.

## II. METHOD

### A. Fuzzification and Fuzzy operation

According to the fuzzy theory, set A can be mapped into the fuzzy set by a membership function, which can be written by *Eq. (1)*:

$$A = \{\mu_A(x_i), x_i | i = 1, 2, \cdots N\} \tag{1}$$

where $\mu_A(x_i)$ is the membership function that maps element $x_i$ to the fuzzy domain with a value between 0 and 1. $N$ is the number of elements in the set. The value indicates the degree of the element belonging to the fuzzy set. If $\mu_A(x_i)$

*Corresponding author: H. D. Cheng hengda.cheng@usu.edu
[1]Xiankun Yan and Heng-Da Cheng are with the Department of Computer Science, Utah State University, Logan, UT 84321, USA
[2]Jianrui Ding is with the School of Computer Science and Technology, Harbin Institute of Technology , Harbin, China

equals to 0, that means $x_i$ is not in the fuzzy set; if $\mu_A(x_i)$ is larger than 0 but lower than 1, $x_i$ is partially in the set; and $x_i$ is fully in the set when $\mu_A(x_i)$ equals to 1.

Because of the narrow histogram of the histopathologic image (Figure 1), the ordinary fuzzification approach is not suitable to apply to the original images. Thus, the membership function is applied to the feature maps that are generated by the convolutional layer. In the paper, the standard S-function is employed as the membership function as below,

$$\begin{aligned}\mu_{ij} &= S(z_{ij};a,b,c) \\ &= \begin{cases} 0 & z_{ij} < a \\ 2(\frac{z_{ij}-a}{c-a})^2 & a \le z_{ij} < b \\ 1-2(\frac{z_{ij}-c}{c-a})^2 & b \le z_{ij} < c \\ 1 & z_{ij} \ge c \end{cases}\end{aligned} \tag{2}$$

where $z_{ij}$ is the logit values at the location $(i,j)$ in the feature maps $Z$ and $\mu_{ij}$ is defined as the fuzzy values of the feature maps in the range $[0,1]$. $a$ and $c$ manipulate the shape of S-function, and $b = (a+c)/2$ is the cross-over point. In the neural network, $a$ and $c$ are a set of parameters learned during training in the range of 0 and 1.

On the other hand, the feature maps are generated by non-linear transformations in the convolutional networks. Here, we employ the normalization on feature maps to transform them into the range [0,1]. The process can be regarded as the fuzzification that keeps most of the original information of feature maps in the fuzzy domain. The normalization can be expressed as:

$$\mu\prime_{ij} = (z_{ij} - \min(z_{ij}))/(\max(z_{ij}) - \min(z_{ij})) \tag{3}$$

With $\mu_{ij}$ and $\mu\prime_{ij}$ obtained by fuzzification, we proposed the Max-min composition to mix them, which is expressed as:

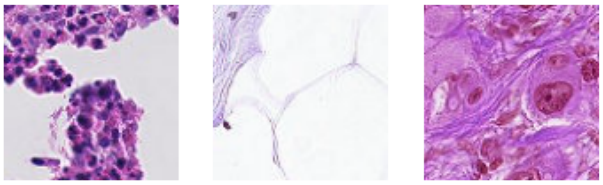
$$\mu\prime\prime_{ij} = \max(\min(\{\mu_i\}), \min(\{\mu\prime_j\})) \tag{4}$$

After the fuzzification of the feature maps, some transformations are used to exploit the uncertainty in the fuzzy domain and generate the input of a fully connected network as the pooling function. We define the fuzzy transformation $f_t(\cdot)$ as follows,

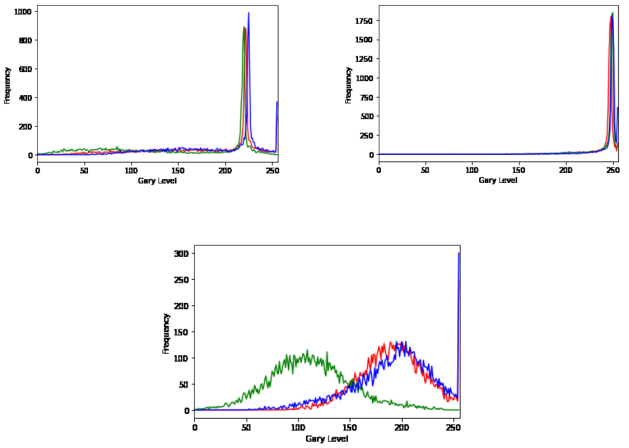
$$\begin{aligned} M_i = f_t(\mu\prime\prime_{ij}, \mu\prime\prime_{\overline{ij}}) &= \frac{\mu\prime\prime_{\overline{ij}} - \overline{\mu\prime\prime_{\overline{ij}}}}{\sigma_{\mu\prime\prime_{\overline{ij}}}} \\ \mu\prime\prime_{\overline{ij}} &= \frac{1}{|j|}\sum \mu\prime\prime_{ij} \\ \overline{\mu\prime\prime_{\overline{ij}}} &= \frac{1}{|n|}\sum \mu\prime\prime_{\overline{ij}} \\ \sigma_{\mu\prime\prime_{\overline{ij}}} &= \sqrt{\frac{\sum(\mu\prime\prime_{\overline{ij}} - \overline{\mu\prime\prime_{\overline{ij}}})^2}{|n|}} \end{aligned} \tag{5}$$

where $n$ is the channel number of feature maps, which is the hyperparameter in the last convolutional layer.

The novel global pooling is called adaptive fuzzy pooling. After the fuzzy global pooling, the fuzzy feature vector would be inputted into the fully connected neural network to conduct defuzzification and classification.



(a) Histopathologic images




(b)

Fig. 1. The histogram of RBG components of 1(a)

### B. Maximum entropy principle

Generally, to determine the parameters of S-function, it utilizes the maximum fuzzy entropy principle. Usually, the entropy is used to measure how much information in source $A$ which can be defined as:

$$H(A) = -\sum_{i=1}^{N} P(x_i) log P(x_i) \tag{6}$$

where $\sum_{i=1}^{N} P(x_i) = 1$ and $\{x_i | i = 1, \cdots N\}$ are the possible outputs from source $A$ with the probability $P(x_i)$. The larger entropy $H(A)$ is, there is more information in source $A$. Correspondingly, when source $A$ was transformed into the fuzzy domain, it uses the fuzzy entropy to measure how much information it can represent in the fuzzy domain. Here, the entropy of a fuzzy set [5] is written as below:

$$H(A) = -\frac{1}{NIn2}\sum S_n(\mu_A(x_i)) \tag{7}$$

where $A$ is a fuzzy set $\{(\mu_A(x_i), x_i) | i = 1, 2, \cdots, N\}$ and $S_n(*)$ is Shannon functions as following:

$$\begin{aligned} S_n(\mu_A(x_i)) = &-\mu_A(x_i) In(\mu_A(x_i)) \\ &-(1-\mu_A(x_i)) In(1-\mu_A(x_i)) \end{aligned} \tag{8}$$

In the paper, the maximum fuzzy entropy principle is used to validate whether the source can reach the maximum fuzzy entropy as well when the parameters of S-function are determined using the neural network.

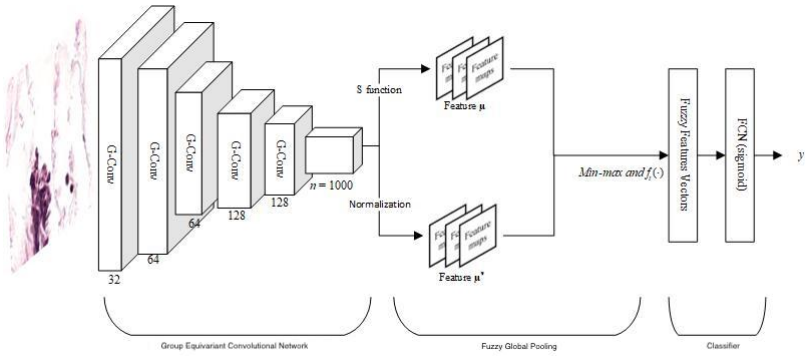



Fig. 2. The basic structure of the proposed model

### C. Proposed Model Structure and Training Details

For the architecture of the proposed network, there are three parts including convolutional network, fuzzy global pooling, and fully connected network (FCN). In the convolutional network, all convolutional layers and batch normalization layers are replaced with the corresponding group-equivariant versions. And the fuzzy global pooling is used to transform the features into the fuzzy domain and mix features as the novel fuzzy input. The fully connected network (FCN) is next to the global pooling layer to perform the defuzzification and classification task.

The propose architecture includes the following layers: $g\_conv1\_3 \times 3 \times 32$, $g\_conv2\_3 \times 3 \times 64$, $pool1\_3 \times 3$, $g\_conv3\_5 \times 5 \times 64$, $pool2\_3 \times 3$, $g\_conv4\_3 \times 3 \times 128$, $pool3\_3 \times 3$, $g\_conv5\_3 \times 3 \times 128$, $pool4\_3 \times 3$, $g\_conv6\_1 \times 1 \times 1000$, $g\_pool1\_3 \times 3$, $fuzzy\_pool$, and $FCN\_N_{class}$ (Figure 2). Here, the naming rule follows the format: "$layer\ type\_kernel\ size \times kernel\ size \times channel\ number$". "$g\_conv$" donates group equivariant convolution; "$pool$" is the traditional pooling layer; "$g\_pool$" is the group equivariant pooling layer; "$fuzzy\_pool$" is the adaptive fuzzy gloabl pooing layer and "$FCN\_N_{class}$" is the fully connected network with $N$ classes outputs. For instance, "$g\_conv1\_3 \times 3 \times 32$" means that the first convolutional layer is a group equivariant convolutional layer, the kernel size is $3 \times 3$ and the number of channels is 32.

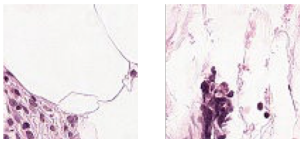
To train the proposed model, it uses the Adam optimization algorithm [8] with with initial learning rate $1e-3$ and the batch size 32. It And it uses the ReLu activation for the convolutional layers and the 10% dropout layers have been applied after the pooling layers. Furthermore, it uses a sigmoidal activation function in the fully connected output layers to make the binary decision. The symmetry group $G = p4$ is used in the proposed model. It also selects the weights with the lowest validation losses to evaluate the test set in the dataset. The models have been trained for about 25 minutes of each epoch using two GeForce GTX 1080 Ti with the batch size 32.

## III. Experimental Result

The experimental procedures do not involve human subjects and animal models. The PatchCamelyson dataset is public and open for research.

### A. Dataset and Validation of Maximum Fuzzy Entropy

PatchCamelyson (PCam) dataset [9] is a new large-scale path level dataset for histopathologic cancer detection. There are 327,680 color images whose size is $96 \times 96px$ (Figure 3). Each image is randomly sampled from histopathologic scans of lymph node sections and is binarily labeled to indicate the cancerous tissue. The dataset is divided into three subsets: a training set (262,144 patches), a validation set (32,768 patches), and a test set (32,768 patches), respectively.



(a) Negative (b) Positive

Fig. 3. Examples of histopathologic patches in dataset

Here, we get the weight of each epoch during training and calculate the average of fuzzy entropy of the images in each epoch (Figure 4). From Figure 4, it indicated that the fuzzy entropy sharply increased between the beginning epoch and twentieth epoch. In the end, it fluctuated in the range of 93 after the fortieth epoch, which means that the information of the image in the fuzzy domain has reached the maximum level. Figure 4 demonstrated that the uncertainty of histopathologic image has been exploited in the fuzzy domain with the increase of the fuzzy entropy.

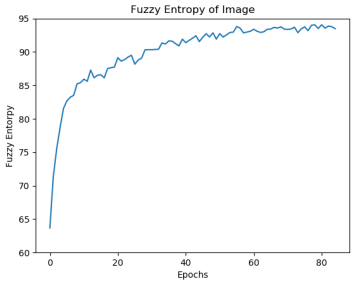



Fig. 4. The fuzzy entropy during training

| Models | NLL | Acc (%) | AUC (%) | F1 Score |
|---|---|---|---|---|
| **Proposed Model** | **0.214** | **91.7** | **97.2** | **0.912** |
| *p4m*-Densenet (Benchmark) | 0.260 | 89.8 | 96.3 | N/A |
| G-CNN (Average Pooling) | 0.346 | 86.3 | 94.0 | 0.849 |
| G-CNN (Max Pooling) | 0.414 | 82.0 | 92.1 | 0.793 |
| DenseNet(Fuzzy Pooling) | 0.295 | 88.5 | 96.2 | 0.889 |
| DenseNet | 0.397 | 82.0 | 90.3 | 0.817 |
| InceptionV3 | 0.640 | 78.8 | 85.7 | 0.778 |
| Resnet | 0.450 | 77.9 | 88.4 | 0.757 |

TABLE I

DIFFERENT MODELS PERFORMANCES ON THE PCAM DATASET IN TERMS OF NEGATIVE LOG-LIKELIHOOD LOSS, ACCURACY, AUC, AND F1 SCORE (THE BOLD FONT INDICATED THE PROPOSED MODEL)

### B. Evaluation results

We evaluate the proposed fuzzy model and prove that the proposed fuzzy model is more robust and reliable. In experiments, the proposed model is compared with some baseline models as well. And the GCNNs with different global pooling functions including the global average pooling function and global max pooling function also are employed to verify the performance of the proposed model. To measure the learning models, four metrics are utilized including the negative log-likelihood loss (NLL), accuracy (Acc), area under the receiver operating characteristic curve (AUC) and F1 score.

The proposed model achieves 91.7% of accuracy, 97.2% of AUC, 0.214 of negative log-likelihood loss and 0.912 of F1 score. It outperforms the benchmark model (*p*4*m*-Densenet model) [9]. The benchmark model got 89.8% of accuracy, 96.3% of AUC, and 0.260 of negative log-likelihood loss. Table I clearly indicates that the proposed model with adaptive fuzzy global function improves the performance a lot on the histopathologic cancer detection compared with other learning models.

## IV. CONCLUSION

In this paper, a novel fuzzy deep learning approach is proposed to detect cancer in histopathological images. The fuzzy group equivariant convolutional neural network is proposed, which consists of the convolutional network as feature extractor, fuzzy global pooling layer as fuzzy processing, and fully connected network as classifier. In the fuzzy global pooling layer, the generated feature maps are transferred into the fuzzy domain by two different fuzzification methods. One of the fuzzy feature maps exploits the uncertainty information of histopathologic images, and the other keeps the original information. Furthermore, the fuzzy feature maps are mixed by using Min-max operations.

Using the proposed model, the maximum fuzzy entropy can be obtained after the fortieth epoch, and the narrow histogram issue of histopathologic images is handled well. The proposed model obtained state-of-the-art performance with 91.7% of accuracy, 97.2% of AUC, 0.214 of negative log-likelihood loss and 0.912 of F1 score.

## REFERENCES


[1] A Ajala Funmilola et al. "Fuzzy kc-means clustering algorithm for medical image segmentation". In: *Journal of Information Engineering and Applications* 2.6 (2012), pp. 21–32.

[2] Christian F Baumgartner et al. "Phiseg: Capturing uncertainty in medical image segmentation". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2019, pp. 119–127.

[3] Juan C Caicedo, Angel Cruz, and Fabio A Gonzalez. "Histopathology image classification using bag of features and kernel functions". In: *Conference on Artificial Intelligence in Medicine in Europe*. Springer. 2009, pp. 126–135.

[4] Heng-Da Cheng and Huijuan Xu. "A novel fuzzy logic approach to contrast enhancement". In: *Pattern recognition* 33.5 (2000), pp. 809–819.

[5] Aldo De Luca and Settimo Termini. "A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory". In: *Information and control* 20.4 (1972), pp. 301–312.

[6] M Hanmandlu, SN Tandon, and AH Mir. "A new fuzzy logic based image enhancement." In: *Biomedical sciences instrumentation* 33 (1997), p. 590.

[7] Robert Ivor John and Peter R Innocent. "Modeling uncertainty in clinical diagnosis using fuzzy logic". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 35.6 (2005), pp. 1340–1350.

[8] Diederik P Kingma and Jimmy Ba. "Adam: A method for stochastic optimization". In: *arXiv preprint arXiv:1412.6980* (2014).

[9] Bastiaan S Veeling et al. "Rotation equivariant CNNs for digital pathology". In: *International Conference on Medical image computing and computer-assisted intervention*. Springer. 2018, pp. 210–218.

[10] Mitko Veta et al. "Breast cancer histopathology image analysis: A review". In: *IEEE Transactions on Biomedical Engineering* 61.5 (2014), pp. 1400–1411.

[11] Yan Xu et al. "Weakly supervised histopathology cancer image segmentation and classification". In: *Medical image analysis* 18.3 (2014), pp. 591–604.