electronics

MDPI

Article

A Novel Fuzzy DBNet for Medical Image Segmentation

Chiun-Li Chin [1], Jun-Cheng Lin [1], Chieh-Yu Li [1], Tzu-Yu Sun [1], Ting Chen [1], Yan-Ming Lai [1], Pei-Chen Huang [1], Sheng-Wen Chang [1,*] and Alok Kumar Sharma [2,*]

[1] Department of Medical Informatics, Chung Shan Medical University, Taichung 40201, Taiwan; ernestli@csmu.edu.tw (C.-L.C.); s0858031@gm.csmu.edu.tw (J.-C.L.); s0858021@gm.csmu.edu.tw (C.-Y.L.); s0858032@gm.csmu.edu.tw (T.-Y.S.); s0858003@gm.csmu.edu.tw (T.C.); s0858029@gm.csmu.edu.tw (Y.-M.L.); s0858009@gm.csmu.edu.tw (P.-C.H.)

[2] Department of Information Management, Chaoyang University of Technology, Taichung 413310, Taiwan

* Correspondence: s0858048@gm.csmu.edu.tw (S.-W.C.); s10814908@cyut.edu.tw (A.K.S.); Tel.: +886-965658928 (S.-W.C.)

Abstract: When doctors are fatigued, they often make diagnostic errors. Similarly, pharmacists may also make mistakes in dispensing medication. Therefore, object segmentation plays a vital role in many healthcare-related areas, such as symptom analysis in biomedical imaging and drug classification. However, many traditional deep-learning algorithms use a single view of an image for segmentation or classification. When the image is blurry or incomplete, these algorithms fail to segment the pathological area or the shape of the drugs accurately, which can then affect subsequent treatment plans. Consequently, we propose the Fuzzy DBNet, which combines the dual butterfly network and the fuzzy ASPP in a deep-learning network and processes images from both sides of an object simultaneously. Our experiments used multi-category pill and lung X-ray datasets for training. The average Dice coefficient of our proposed model reached 95.05% in multi-pill segmentation and 97.05% in lung segmentation. The results showed that our proposed model outperformed other state-of-the-art networks in both applications, demonstrating that our model can use multiple views of an image to obtain image segmentation or identification.

Keywords: Fuzzy DBNet; butterfly network; pill; lung X-ray; anteroposterior; posteroanterior

Citation: Chin, C.-L.; Lin, J.-C.; Li, C.-Y.; Sun, T.-Y.; Chen, T.; Lai, Y.-M.; Huang, P.-C.; Chang, S.-W.; Sharma, A.K. A Novel Fuzzy DBNet for Medical Image Segmentation. *Electronics* **2023**, *12*, 2658. https://doi.org/10.3390/electronics12122658

Academic Editor: Javid Taheri

Received: 3 May 2023
Revised: 2 June 2023
Accepted: 10 June 2023
Published: 13 June 2023

1. Introduction

There have been numerous cases where image segmentation in medicine has been implemented primarily because interpreting medical images from CT and MRI scans requires substantial medical knowledge and time. In addition, statistics reveal that medical errors contribute to the deaths of 7000–9000 people annually in the United States [1]. Hence, to address these issues, several technological solutions have been proposed. For instance, the U-Net developed by Ronneberger et al. [2] is used in the segmentation of biomedical images. The DoubleU-Net employed by Debesh Jha et al. [3] is used in procedures such as a colonoscopy. It is primarily built upon the U-Net and VGG-19, composed of two encoders and decoders. Similarly, Chin et al. [4] use the Mask R-CNN deep learning algorithm to segment the vocal cords and glottis regions from larynx videos, aiding doctors in diagnosis and treatment.

Among lung image recognition, Jakub et al. pointed out that the results of X-ray images of disease lesions are easily occluded by blood vessels. Though machine learning methods for lung image recognition can help reduce the burden on medical personnel, their accuracy is only 91% [5]. In addition, the deep-learning model for lung disease identification has proven to be superior than traditional machine learning methods [6,7]. Therefore, we used deep-learning for image recognition. In the case of pneumothorax thoracostomy, the doctor needs to look at the chest X-ray image to find the translucent pleural line that overlaps the ribs. Because this image is prone to blurring from tissue overlap, traditional image

recognition is not effective [8], especially because air accumulates on the front side of the body rather than at the top, making it difficult for doctors to interpreting pneumothorax areas. Due to the fact that initial symptoms of many lung lesions are not obvious and lungs have complex pathological characteristics [9], many scholars have proposed deep-learning methods to improve diagnosis accuracy and efficiency [10]. Laura et al. [11] proposed a complex network for lung image recognition, and their experiment showed a high accuracy for texture recognition. Accordingly, a complex network method can extract important features. Alhassan et al. [12] used the ensemble learning method to detect pneumonia in chest X-ray images with an improved accuracy rate of 93%. Based on the ensemble-learning characteristics, the method is used to improve model-identification accuracy. Mohammad et al. [13] proposed a deep-learning model to aid early detection of COVID-19, which helped medical staff reduce their workload. For lung image segmentation, Feidao et al. [14] proposed a three-terminal attention mechanism to highlight the target area automatically and improve the lung-segmentation performance. As a result, an attention-gate module is used to improve the training effect of the model. However, these models did not use chest X-ray images fully. There are two modes for taking lung X-ray images: anteroposterior and posteroanterior. Anteroposterior lung X-rays are taken from the chest cavity; posteroanterior lung X-rays are taken from the back. Existing deep-learning models can only input one image at a time for recognition [15]. Therefore, the model may not be able to identify the location of the lesion since only one side of the X-ray image is input [16]. Consequently, we used deep-learning methods for lung image identification.

Among pill image recognition, the majority of existing methods for classification rely only on single-sided information, but in some cases, many types of pills cannot be classified. For instance, the same type of pill may have different shapes when viewed from different angles; some pills may have lettering on only one side; and those with similar shapes and colors can also pose a challenge. In such cases, the model may not be able to classify a group of pills accurately. Because of the pandemic, large numbers of people flock to hospitals every day, leading to a substantial increase in pill usage and medical errors [17,18]. To address these problems, Ou et al. [19] proposed a two-stage deep-learning architecture for detecting and subsequently classifying 1000 types of pills. In addition, improving medication knowledge and providing patients with adequate information has become important issues for avoiding medication waste and harmful side effects [20]. However, appearance-based pill identification remains a daunting task for patients. Wang et al. [21] used the GoogLeNet Inception Network to train deep-learning architecture and image enhancement techniques by focusing on color, shape, and markings, but it can only identify a single type of pill, which is its big flaw. The simultaneous identification of multiple types of pills can meet the needs of the public better. In the field of pill image segmentation, Kwon et al. [22] used Mask R-CNN, and the training datasets they used contained only 27 types of pills, each of which had different shapes and colors. In reality, though, many more types of pills have a similar appearance. To suit real-world needs, 93 types of pills were used to train our model. They had similar shapes and colors and varied mainly in different imprints, making our model more applicable to clinical needs. In pill object detection, Lu et al. [23] proposed many object detection methods. Although these models detected the position of pills, the map is only 87%. Additionally, previous scholars have not effectively addressed issues in pill identification, including their random placement and the presence of multiple pills in an image. The rotation angle of the pill is also difficult to determine and standardize for each pill class.

For computer vision, other cases use unmanned aerial vehicles (UAVs) combined with deep learning, Keiller et al. [24] studied plant classification from spatial and spectral perspectives using RGB and spectral UAV images based on 2D-CNN technology.

Current methods for image segmentation mostly use a single image as input. When a single low-quality image is input into the model, issues with imprecise object segmentation and classification arise. In the medical field, due to the complex characteristics of images, Akinobu et al. [25] proposed BtrflyNet to identify bone metastases, which can accept two

input images simultaneously. The experimental results indicated that it can enhance the success rate of model training. However, this model is only applicable to bone metastasis images. Based on this paper, we improved the BtrflyNet to propose Fuzzy DBNet, which has achieved excellent results in pill image and lung image recognition.

2. Materials and Methods

2.1. Datasets

In this paper, chest X-ray and pill datasets were used. Chest X-ray data were obtained from the NIH chest X-ray dataset [26] containing 112,120 images, from which we extracted 72,324 of both the anteroposterior (AP) and posteroanterior (PA) views of the same patients. We then paired the AP images with their corresponding PA images, resulting in a total of 267,105 AP–PA pairs after augmentation.

The pill dataset comprises 93 categories having a total of 1238 images: round, oval, rectangular, triangular, and of different colors. To capture both sides of each pill, images were taken after pills were scattered on a transparent board and photographed from directly above and below. The dataset was then divided into 80% for training, 10% for testing and 10% for validation. To improve the model's accuracy, we applied the AutoAugment scheme [27] to discretize each operation magnitude (M) from [0, 10] for data augmentation, resulting in an increased dataset of 2476 images.

The dataset was the most significant limitation to our model. First, our input images had to be dual-sided, not only general images but also images that penetrated objects, such as X-rays. Second, fixed correspondences of object positions in the images were required. To address this, our team made efforts to align the positions of the pills between their dual-sided images during the collection of the dataset.

It is important to note that the lung X-ray dataset used in this study was provided by the NIH, while the pill dataset was captured by our team's own photography.

2.2. Data Pre-Processing

To assist pharmacists in pill classification, 93 types of pills were annotated. We used VGG image Annotator [28] to label each one with its name in 93 different categories. We labeled the edge of the pill and converted the labeled data into a JSON file as the ground truth of the training data. The pre-processing algorithm can be found in Algorithm A1.

2.3. Fuzzy DBNet

This paper proposes a new deep-learning architecture, Fuzzy Double-Butterfly Network (DBNet), in which two images with complementary properties can be input to perform image segmentation. It is mainly composed of three parts: double-butterfly encoder–decoder architecture, Fuzzy Atrous Spatial Pyramid Pooling (ASPP) block and an attention gate, as shown in Figure 1.

A pretrained VGG 19 was used to encode of the first butterfly architecture to extract image features for saving training time and preventing overfitting. Between the two butterfly-shaped encoders and decoders, there were two concatenate blocks that connected two sets of Fuzzy ASPP blocks, thereby exchanging the different characteristics of the images. Then, the initial result output by the first butterfly-shaped architecture was multiplied with the original image as shown in the multiply block in Figure 1. This enhanced the specific gravity of the feature to achieve a more precise segmentation. The Fuzzy DBNet algorithm can be found in Algorithm A2.

Fuzzy ASPP combined ASPP with fuzzy theory and was placed between the encoder and decoder of the two butterfly-type networks. Figure 2 shows the structure of Fuzzy ASPP.

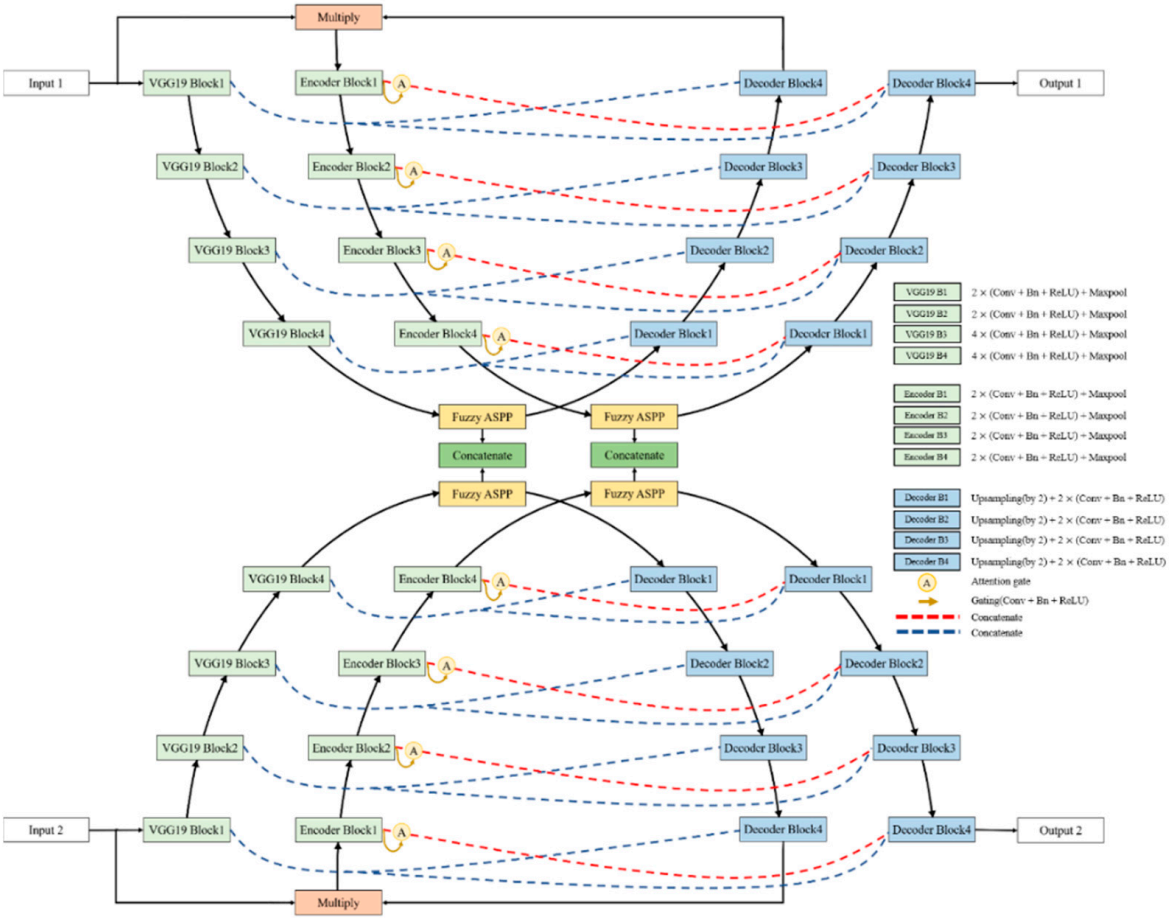

Figure 1. Fuzzy DBNet architecture.

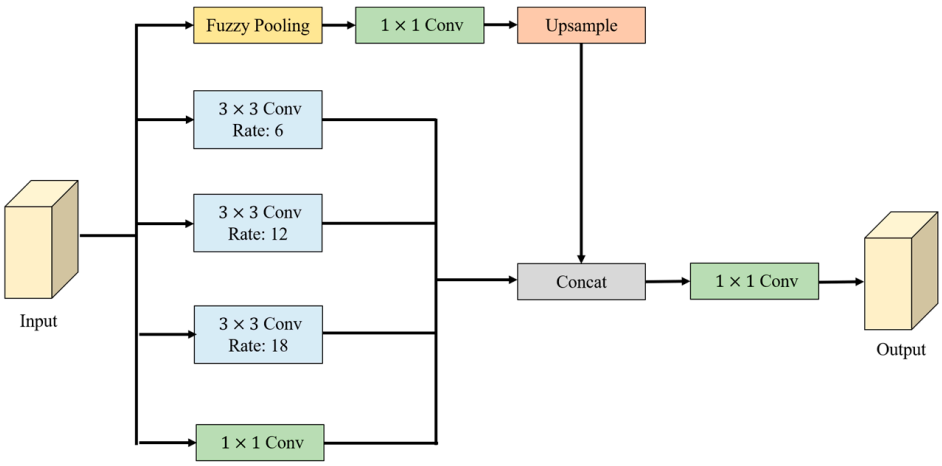

Figure 2. Fuzzy ASPP structure.

In the Fuzzy Pooling part, a bell-shaped membership function was used to complete the operation, and the results of each pooling patch were dynamically adjusted. Algorithm A3 shows the dynamic adjustment algorithm. The main purpose was to reduce the proportion of the feature influenced by uncertainty factors. The Fuzzy ASPP algorithm can be found in Algorithm A4.

In the skip connections, an attention mechanism was used to eliminate noisy and irrelevant responses by using features extracted from coarser feature maps. It effectively reduced noise and unnecessary features in the model and improved its performance and accuracy. Figure 3 shows the structure of the attention gate.

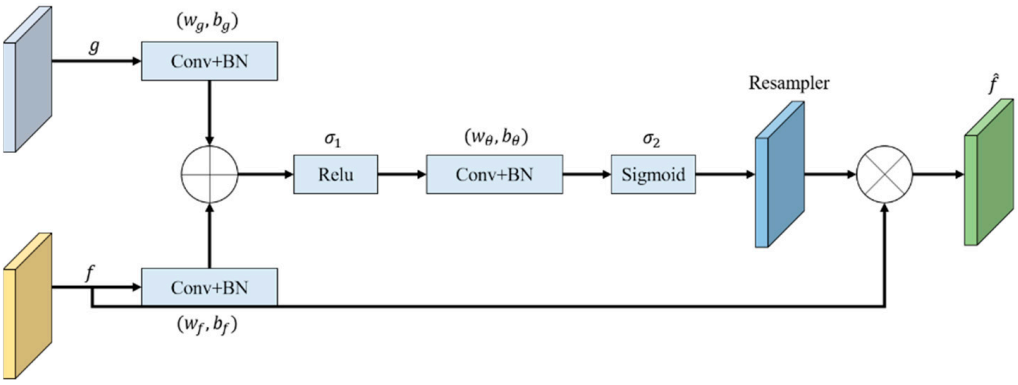

Figure 3. Attention gate structure.

3. Result

3.1. Experiment Setting

The training validation and testing of the proposed model was done on a computer with an 8-core CPU (Intel Xeon W-3223), 64 GB memory, a GPU (RTX 3090) with 24 GB graphics memory, and 10,496 CUDA cores. Implementation was done using the PyTorch framework. Table 1 shows the hyperparameters used for all experiments.

Table 1. Hyperparameters for training.

Hyperparameter	Selected Value
Loss function	L_{DC}
Optimizer	Adam
Learning rate	8×10^{-5}
Batch size	4
Epoch	500

The total loss function (L_{DC}) combines of the averaged Dice Loss ($L_{avgDice}$) with the Categorical Cross-Entropy Loss (L_{CCE}) and is calculated as follows:

$$L_{DC} = L_{avgDice} + L_{CCE} \tag{1}$$

$$L_{avgDice} = \frac{\sum_{i=1}^{N} 1 - \frac{2 \times \sum P_{N_{true}} \times P_{N_{pred}}}{\sum P_{N_{true}}^2 + \sum P_{N_{pred}}^2 + \epsilon}}{N} \tag{2}$$

$$L_{CCE} = -\frac{1}{N} \sum_{i=0}^{N} \sum_{c=0}^{C} P_{pred_c} \cdot log(P_{true_c}) + \left(1 - P_{pred_c}\right) \cdot log(1 - P_{true_c}) \tag{3}$$

where N is the number of samples and C is the number of classes. Since the output of the network was multi-classes images, we calculated the L_{CCE} using Equation (3). Then, we calculated the L_{Dice} for each class using Equation (2). It was possible to zero out all pixels in the P_{pred} that were not active in the P_{true}. For activated pixels, low-confidence predictions were mostly penalized, while higher prediction values obtained higher Dice coefficients. Therefore, the model learned objects of different classes and sizes through L_{Dice} and L_{CCE}.

3.2. Performance Evaluation Index

In this study, we used three metrics to evaluate model performance: pixel-wise accuracy, averaged *Dice* coefficient (*Dice*), and mean Intersection over Union (*mIoU*). These indicators are shown in the following formula.

$$Pixel - wise \ \ Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$Dice = \frac{1}{k+1} \sum_{i=0}^{k} \frac{2 * |X_i \cap Y_i|}{|X_i| + |Y_i|} \tag{5}$$

$$mIoU = \frac{1}{k+1} \sum_{i=0}^{k} \frac{|X_i \cap Y_i|}{|X_i \cup Y_i|} \tag{6}$$

where X_i denotes the ground truth values, and Y_i denotes the predicted values. The *TP*, *FP*, *TN*, and *FN* depict the case numbers of true positives, false positives, true negatives, and false negatives, respectively. Pixel-wise accuracy measured the percentage of correctly identified pixels in the image; the Dice score measured the overlap between the predicted segmentation and ground truth; and mIoU measured the predicted segmentation with the ground truth. These metrics were chosen because they provided a comprehensive view of the model's performance and allowed for meaningful comparisons with other models in the field. Higher values for these metrics indicated better model performance.

3.3. Segmentation of Lung X-ray Images

We conducted experiments to train the proposed method and validated the model by using a validation set. As shown in Figure 4, the training loss of Fuzzy DBNet on the lung X-ray dataset reached convergence at around the 100th epoch and achieved complete convergence at approximately the 300th.

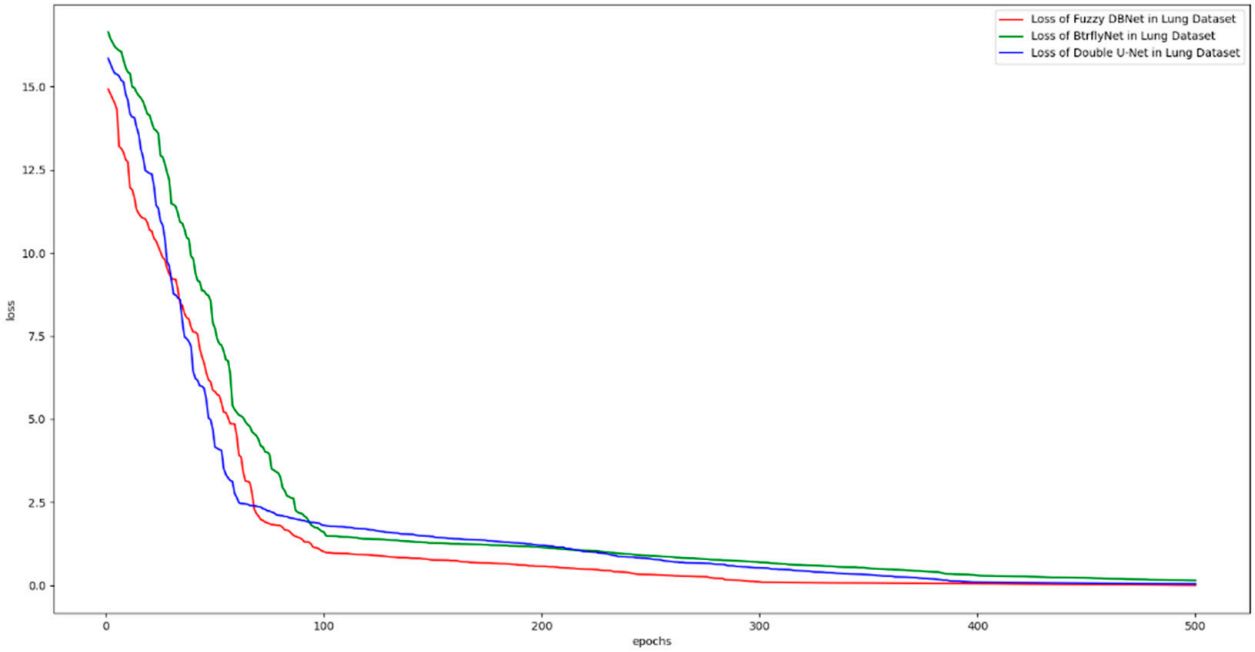

Figure 4. Validation loss of Fuzzy DBNet, BtrflyNet, and Double U-Net models on the Lung Dataset.

On the lung X-ray dataset, we selected a set of images as examples for model testing. It comprised six images: anterior–posterior and posteroanterior raw images, the corresponding ground truth images and the segmented results. These images are displayed in Figure 5.

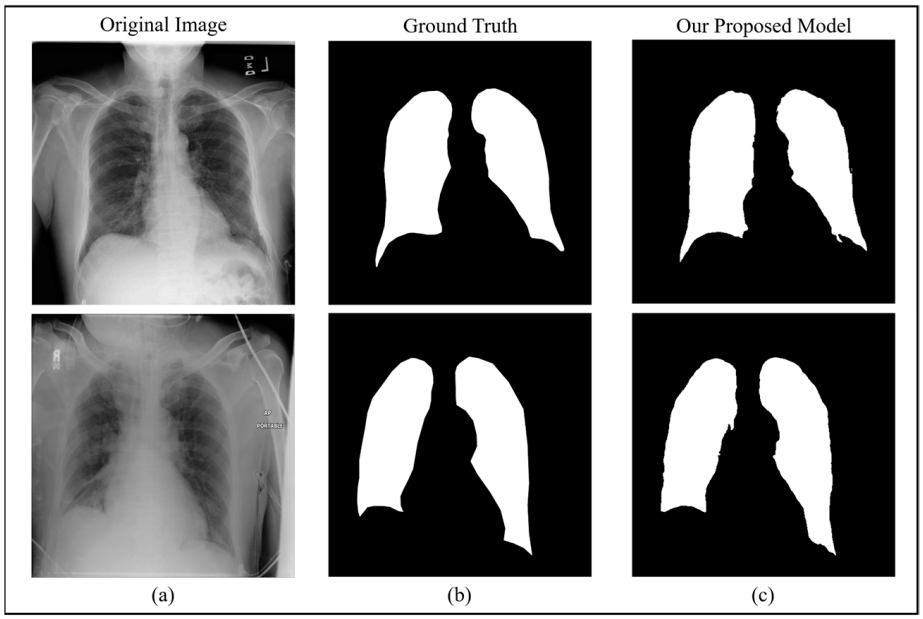

Figure 5. A patient's lung: (**a**) anteroposterior and posteroanterior lung X-rays images, (**b**) their corresponding ground truth images and (**c**) segmentation images.

We compared the outputs of Fuzzy DBNet and its ground truth on the testing data using average Dice coefficient, mIoU and pixel-wise accuracy to measure the performance of the model. The results are shown in Table 2.

Table 2. Averaged Dice coefficient, mIoU, and pixel-wise accuracy in lung testing dataset.

Evaluation Metrics	Front Image	Back Image
averaged Dice coefficient	0.963 ± 0.012	0.978 ± 0.019
mIoU	0.930 ± 0.011	0.957 ± 0.021
pixel-wise accuracy	0.961 ± 0.007	0.977 ± 0.009

We selected two sets of lung X-rays from the segmentation results of our test set as examples. In Figure 6, the segmentation completeness of our model was much better than that for BtrflyNet. In Figure 7, when the original lung images were blurry, the segmentation performance of DoubleU-Net was poor, while our model accurately segmented the lungs.

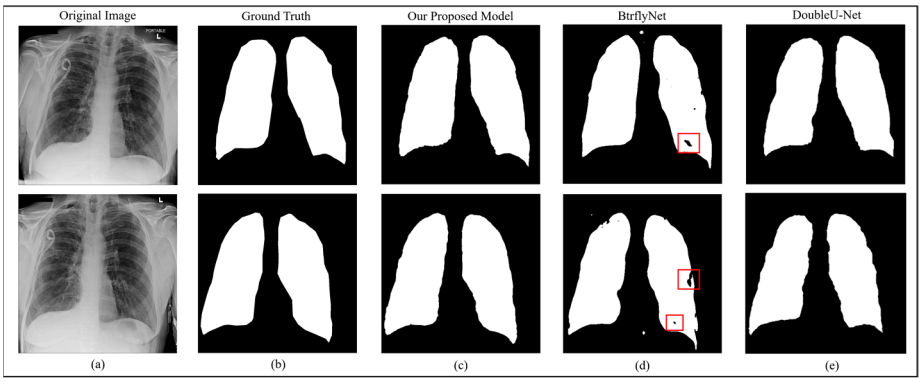

Figure 6. A pair of lung images (**a**). Ground truth images (**b**). Our Proposed Model (Fuzzy DB-Net) results (**c**). BtrflyNet results (The red square in the caption highlights the missing output of BtrflyNet.) (**d**). DoubleU-Net results (**e**).

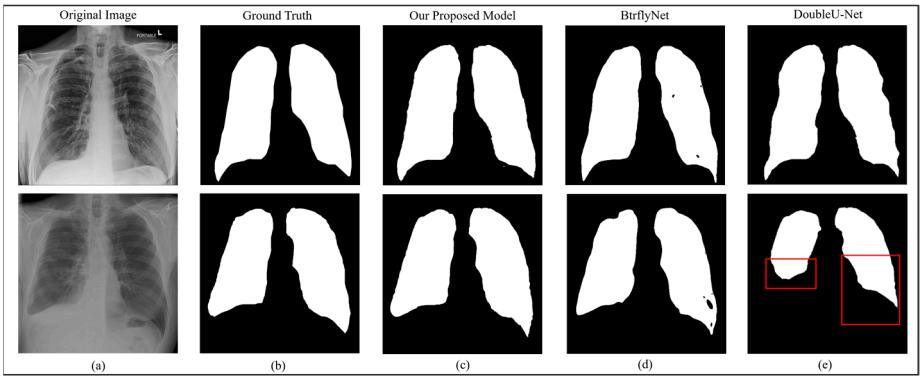

Figure 7. A pair of lung images (**a**). Ground truth images (**b**). Our Proposed Model (Fuzzy DBNet) results (**c**). BtrflyNet results (**d**). DoubleU-Net results (The red square in the caption highlights the area where DoubleU-Net demonstrates a blurry omission in the lung region.) (**e**).

3.4. Segmentation of Pill Images

Fuzzy DBNet performed well among the pill datasets. To verify the generalization of the model further, we also conducted experiments on the lung dataset. Figure 8 shows the validation loss.

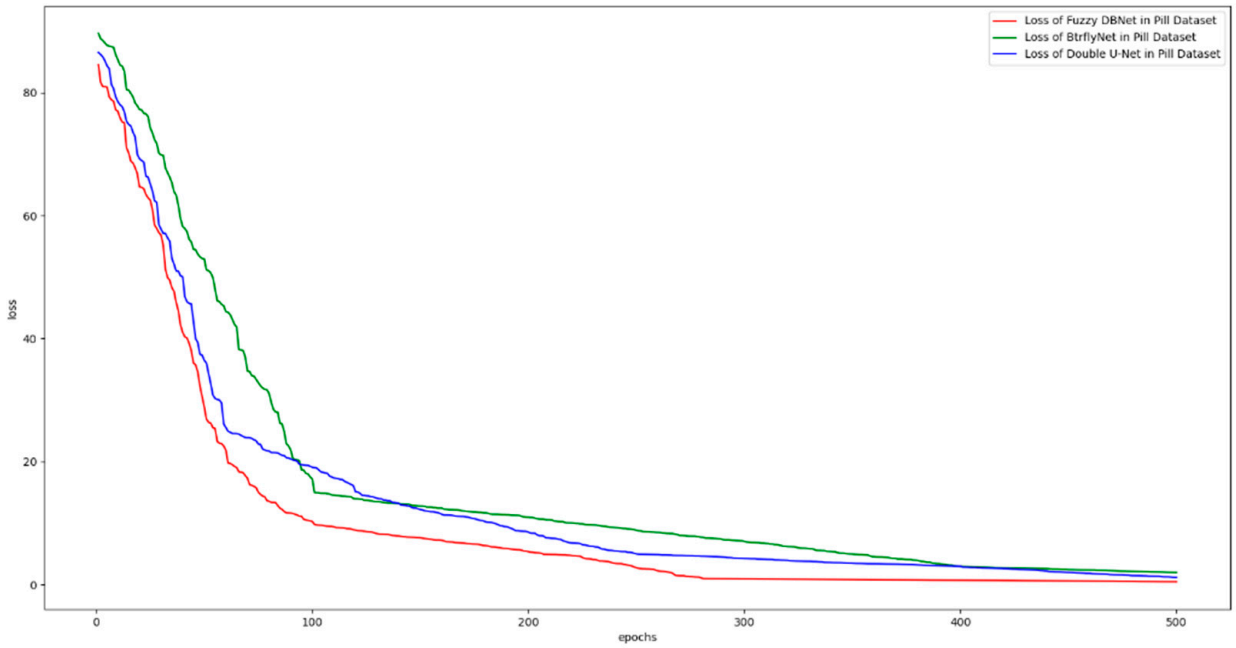

Figure 8. Validation loss of Fuzzy DBNet, BtrflyNet, and Double U-Net models on the Pill Dataset.

On the pill dataset, we chose a group of images to test the model. This set contained six images: raw images from top to bottom and bottom to top and the corresponding ground truth images and segmented results. These images are shown in Figure 9.

In order to evaluate the performance of our proposed model, we compared the output of Fuzzy DBNet and its ground truth on the testing dataset using averaged Dice coefficient, mIoU and pixel-wise accuracy. The results are shown in Table 3.

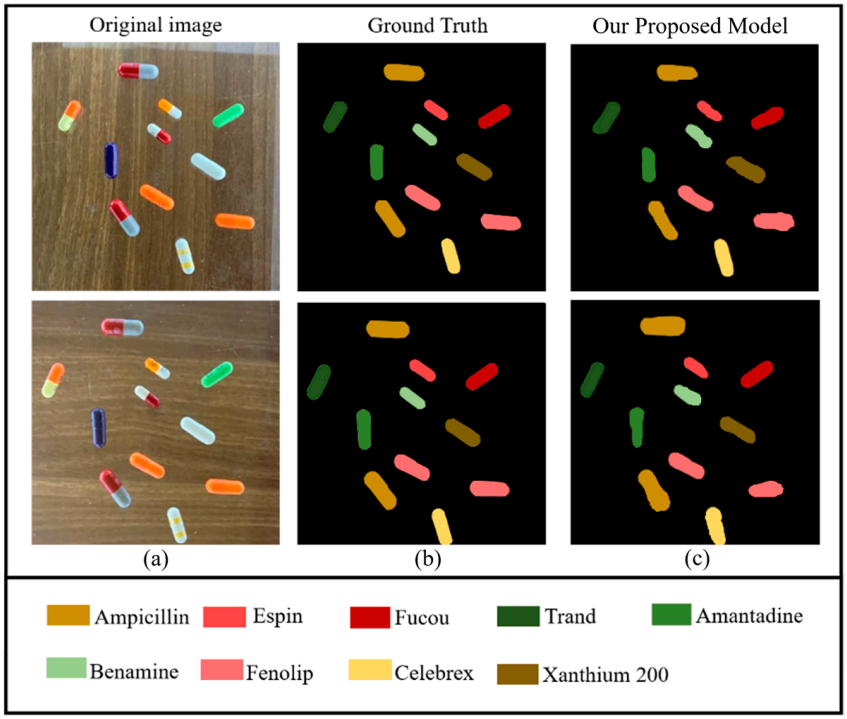

Figure 9. (**a**) a randomly selected pair of pill images, (**b**) corresponding ground truth images and (**c**) segmentation images.

Table 3. Averaged Dice coefficient, mIoU, and pixel-wise accuracy evaluation results in pill testing dataset.

Evaluation Metrics	Front Image	Back Image
averaged Dice coefficient	0.933 ± 0.012	0.968 ± 0.019
mIoU	0.874 ± 0.011	0.939 ± 0.021
pixel-wise accuracy	0.903 ± 0.008	0.953 ± 0.010

We selected two sets of images from the pill test dataset segmentation results as examples. In Figure 10, the segmentation performance of BtrflyNet was inferior. On the other hand, when DoubleU-Net recognized pills without text, it was prone to classification error. Figure 11 shows test cases with intentionally incorrect image inputs, where images of pills were captured at different angles instead of the front and back views. This led not only to the poor classification performance of DoubleU-Net but also a decrease in classification accuracy for our model and BtrflyNet because they were affected by the misplaced drug features.

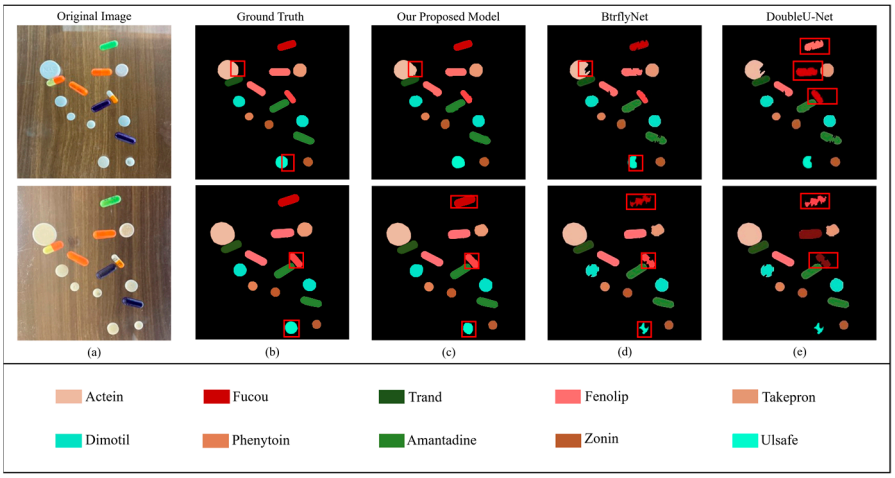

Figure 10. (**a**) A pair of pill images. (**b**) Ground truth images. (**c**) Our Proposed Model (Fuzzy DBNet) results. (**d**) BtrflyNet results. (**e**) DoubleU-Net results. (The red square in the caption highlights the differences in segmentation between each model and the ground truth.)

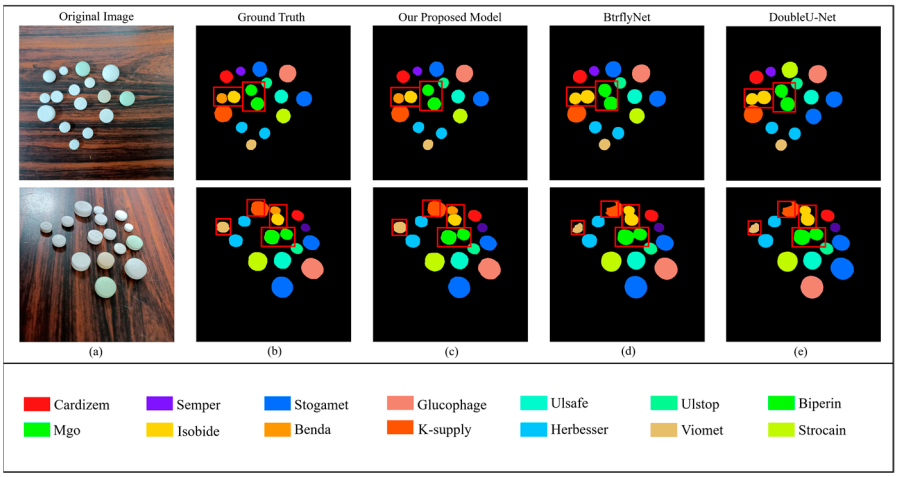

Figure 11. (**a**) An incorrectly Angled Image Set. (**b**) Ground truth images. (**c**) Our Proposed Model (Fuzzy DBNet) results. (**d**) BtrflyNet results. (**e**) DoubleU-Net results. (The red square highlights the differences among the models in terms of classification.)

4. Discussion

We compared the performance of our proposed method with that of BtrflyNet on the pill and lung datasets and demonstrated that our approach achieved better segmentation results.

Referring to Figures 6, 7 and 10, it can be observed that our model trained by integrating dual-sided images. Our model achieved more a complete segmentation and had higher classification accuracy compared to other models. Conversely, our model also had limitations. For instance, as shown in Figure 11, it lacked advantages in non-dual-sided images or dual-sided images with misaligned positions, and it was prone to misclassify segmented objects. From these two cases, it was inferred that the model combined important features from both sides of the input images during training. Therefore, two input images of the same target object should overlap as much as possible to leverage the advantages of our model fully.

In the lung X-ray dataset, our model outperformed BtrflyNet in segmentation accuracy. Due to the incorporation of fuzzy ASPP and the use of dual-sided images as input, our model produced more complete lung segmentations even for blurry images, compared to DoubleU-Net.

In the pill dataset, our model outperformed both BtrflyNet and DoubleU-Net in segmentation accuracy and classification precision, particularly when the text on the back of pills was present. By leveraging features from the backside images, our proposed method achieved accurate classification.

After consulting with medical centers, we learned that many pathological conditions require information from dual-sided medical imaging for an accurate diagnosis. For instance, physicians typically use AP and PA views from radionuclide bone scanning to diagnose metastatic lesions, as well as images from different perspectives in knee X-ray data to diagnose arthritis. Concurrently, based on the aforementioned analysis, it can be inferred that our model could achieve superior results when applied to datasets containing this type of dual-sided imaging.

Overall, as shown in Table 4, our proposed method achieved a higher mIoU score, dice coefficient, and pixel-wise accuracy compared to BtrflyNet and DoubleU-Net. These results demonstrated that our proposed method improved image segmentation accuracy from dual-sided images.

Table 4. Comparison of various indicators between Fuzzy DBNet and BtrflyNet in the pill dataset and the lung X-ray dataset.

Dataset	Model	Averaged Dice Coefficient	mIoU	Pixel-Wise Accuracy
Pill Dataset	Fuzzy DBNet	0.951	0.907	0.928
	BtrflyNet	0.887	0.843	0.859
	Double U-Net	0.949	0.881	0.824
Lung X-ray Dataset	Fuzzy DBNet	0.971	0.944	0.969
	BtrflyNet	0.909	0.883	0.897
	Double U-Net	0.965	0.922	0.941

5. Conclusions

In this paper, we proposed the Fuzzy DBNet, which takes two input images to solve the issue of one side of an object being blurry or incomplete. Our proposed model achieved a pixel-wise accuracy of 92.8% on a drug dataset, which was 10.4% more accurate than Double U-Net and 6.9% compared to BtryflyNet, effectively solving the problem of inconsistent text on both sides of a drug that cannot be accurately classified based on a single image. On a chest X-ray dataset, pixel-wise accuracy reached 96.9%, which was 2.8 and 7.2% more accurate than for Double U-Net and BtryflyNet, respectively. This improved the issue of obscured or noisy regions in image segmentation.

Future work mainly consists of two parts. First, we aim to apply our model to various disease lesion recognitions that require dual-sided medical images, such as bone metastasis and knees. Following this, we plan to develop a network architecture that can integrate multi-angle images that go beyond the current scope of double-sided images. This would allow for the creation of a multi-view model to simulate stereoscopic vision. By leveraging depth information captured from different angles, the effectiveness of the model can be further enhanced.

This advancement would facilitate the collection and preparation of more varied datasets and increase the applicability of the proposed approach to a broader range of fields.

Author Contributions: Methodology, C.-L.C., C.-Y.L., Y.-M.L., S.-W.C. and A.K.S.; Software, J.-C.L., C.-Y.L., Y.-M.L. and S.-W.C.; Validation, J.-C.L., C.-Y.L. and A.K.S.; Formal analysis, Y.-M.L.; Writing—original draft, T.-Y.S., T.C. and P.-C.H.; Supervision, T.-Y.S., T.C. and P.-C.H.; Project administration, C.-L.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All datasets used in this paper are publicly available.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Algorithm A1 Data pre-processing Algorithm

Require: annotations, original images, the label of images
Ensure: training dataset, testing dataset, validation dataset
Loading images, labels, annotations;
for label, annotation, image **in** images **do**
  mask $\leftarrow$ AnnotationToMask($image$, $annotation$);
  newData $\leftarrow$ OneHotEncode($image$, $mask$, $label$);
end for
newData $\leftarrow$ Random($newData$);
counts $\leftarrow$ Count($newData$);
for Iter, data **in** newData **do**
 if Iter $<$ (counts $\times$ 0.8):
  trainval dataset **append** data;
 else :
  testing dataset **append** data;
end for
trainvalCounts $\leftarrow$ Count($trainval\ dataset$);
for Iter, data **in** trainval dataset **do**
 if Iter $<$ (trainvalCounts $\times$ 0.8):
  training dataset **append** data;
 else :
  validation dataset **append** data;
end for
return training dataset, testing dataset, validation dataset;

Algorithm A2 Fuzzy DBNet Algorithm

Require: annotations, images, labels, batch size N, epoch k
Ensure: Fuzzy DBNet model
Loading images, labels, annotations;
trainData, testingData, valData $\leftarrow$ pre-processing(*annotations, images, labels*);
n $\in$ Count($train\ data$) $\div$ N + 1;
Define Fuzzy DBNet ($frontInput$, $backInput$):
 VGG19 $\leftarrow$ LoadVGG19Block();
 Encoder $\leftarrow$ LoadEncoderBlock();
 VGG19Decoder $\leftarrow$ LoadVGG19DecoderBlock();
 FixDecoder $\leftarrow$ LoadDecoderBlock();
 FuzzyASPP $\leftarrow$ LoadFuzzyASPP();
ModelF, ModelB **add** VGG19, FuzzyASPP
Concatenate(*ModelF,ModelB*)
ModelF, ModelB **add** VGG19Decoder
ModelF, ModelB **add** Encoder, FuzzyASPP;
Concatenate(*ModelF,ModelB*)
 ModelF, ModelB **add** FixDecoder ;
 frontOnput $\leftarrow$ ModelF($frontInput$);
 backOnput $\leftarrow$ ModelB($backInput$);
return frontOnput, backOnput;
for Iter = 1 to k **do**
 for step = 1 to n **do**
  frontInput, backInput $\leftarrow$ ImageSplit($trainData$);
  $\hat{y}_f$, $\hat{y}_b$ $\leftarrow$ Fuzzy DBNet ($frontInput$, $backInput$);

Algorithm A2 *Cont.*

23 : Update the Fuzzy DBNet by Adaptive Moment Estimation :

$$\frac{\sum_{i=1}^{N} 1-\frac{2\times\sum y_{N_f}\times\hat{y}_{N_f}}{\sum y_{N_f}^2+\sum\hat{y}_{N_f}^2+\epsilon}-\frac{2\times\sum y_{N_b}\times\hat{y}_{N_b}}{\sum y_{N_b}^2+\sum\hat{y}_{N_b}^2+\epsilon}}{N}$$

$$-\frac{1}{N}\sum_{i=0}^{N}\sum_{j=0}^{J}\hat{y}_{f_j}\cdot\log\left(y_{f_j}\right)+\left(1-\hat{y}_{f_j}\right)\cdot\log\left(1-y_{f_j}\right)+$$

$$\hat{y}_{b_j}\cdot\log\left(y_{b_j}\right)+\left(1-\hat{y}_{b_j}\right)\cdot\log\left(1-y_{b_j}\right)$$

24 : **end for**
25 : frontVal, backVal ← ImageSplit($valData$);
26 : $\hat{y}_f$, $\hat{y}_b$ ← Fuzzy DBNet ($frontVal$, $backVal$)
27 : $Pixel-wise\ Accuracy$ ← Evaluate($\hat{y}_f$, $\hat{y}_b$)
28 : **save** checkpoint_Iter
29 : **end for**

Algorithm A3 Genetic Algorithm

Require: Fuzzy Set
Ensure: new Fuzzy Set
1 : **Define** Genetic Algorithm ($FuzzySet$, $epochs$, r_mutate, set_size, $r\_elitism$):
2 : **for** epoch **in** epochs **do**:
3 : elitism ← $set\_size \times r\_elitism$
4 : newSet1 ← select the best elitism in FuzzySet
5 : crossovers ← $\frac{FuzzySet-\text{elitism}}{2}$
6 : **for** co **in** crossover **do**:
7 : X_A , X_B ← Randomly select from
8 : X_C , X_D ← generate by one − point crossover to X_A , X_B
9 : newSet2 append X_C , X_D
10 : **end for**
11 : **for** co **in** crossover **do**:
12 : X_{co} ← Randomly select from newSet2
13 : X'_{co} ← mutate each bit of X_{co} under the rate r_mutate
14 : newSet2 update X_{co} , X'_{co}
15 : **end for**
16 : **return** newSet1 + newSet2

Algorithm A4 Fuzzy ASPP Algorithm

Require: featureMap
Ensure: newFeatureMap
1 : **Define** FuzzyPooling ($featureMap$):
2 : **for** patch **in** featureMap **do**:
3 : **for** n, pixel **in** patch **do** :
4 : Calculate π_v^n by bell shaped:

$$\pi_{pixel}^n = bell(piexl^n) = \frac{1}{\sigma_{pixel}\sqrt{2\pi}}e^{-\frac{(x-\mu_{pixel})^2}{2\sigma_{pixel}{}^2}}$$

5 : **end for**
6 : **for** k, pixel **in** patch **do**:
7 : Calculate the scores $s^k_{\pi_{pixel}}$:

$$s^k_{\pi_{pixel}} = \sum_{i=1}^{k}\sum_{j=1}^{k}\pi^k_{pixel_{i,j}}$$

Algorithm A4 *Cont.*

```
    end for
    Calculate π′ then Calculate patch′;
    FuzzyPoolingFM ← Concate(FuzzyPoolingFM, patch′);
  end for
return FuzzyPoolingFM
FuzzyPoolingFeatureMap ← FuzzyPooling(featureMap);
covFeatureMap ← Convolution(FuzzyPoolingFeatureMap);
atrousFeatureMap ← AtrousConvolution(covFeatureMap)
newFeatureMap ← Concate(featureMap, atrousFeatureMap)
return newFeatureMap
```

References

1. Tariq, R.A.; Vashisht, R.; Sinha, A.; Scherbak, Y. *Medication Dispensing Errors and Prevention*; StatPearls: Treasure Island, FL, USA, 2020.
2. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the 18th International Conference, MICCAI, Munich, Germany, 5–9 October 2015; pp. 234–241.
3. Jha, D.; Riegler, M.A.; Johansen, D.; Halvorsen, P.; Johansen, H.D. Doubleu-net: A deep convolutional neural network for medical image segmentation. In Proceedings of the IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS), Rochester, MN, USA, 28–30 July 2020; pp. 558–564.
4. Chin, C.L.; Chang, C.L.; Liu, Y.C.; Lin, Y.L. Automatic segmentation and indicators measurement of the vocal folds and glottal in laryngeal endoscopy images using Mask R-CNN. *Biomed. Eng. Appl. Basis Commun.* **2021**, *33*, 2150027.
5. Płudowski, J.; Mulawka, J. Machine Learning in Recognition of Basic Pulmonary Pathologies. *Appl. Sci.* **2022**, *12*, 8086. [CrossRef]
6. Rajaraman, S.; Guo, P.; Xue, Z.; Antani, S.K. A Deep Modality-Specific Ensemble for Improving Pneumonia Detection in Chest X-rays. *Diagnostics* **2022**, *12*, 1442. [CrossRef] [PubMed]
7. Granata, V.; Fusco, R.; Costa, M.; Picone, C.; Cozzi, D.; Moroni, C.; Petrillo, A. Preliminary report on computed tomography radiomics features as biomarkers to immunotherapy selection in lung adenocarcinoma patients. *Cancers* **2021**, *13*, 3992. [CrossRef] [PubMed]
8. Zarogoulidis, P.; Kioumis, I.; Pitsiou, G.; Porpodis, K.; Lampaki, S.; Papaiwannou, A.; Katsikogiannis, N.; Zaric, B.; Branislav, P.; Secen, N.; et al. Pneumothorax: From definition to diagnosis and treatment. *J. Thorac. Dis.* **2014**, *6* (Suppl. S4), S372–S376. [PubMed]
9. Zhang, R.; Yang, F.; Luo, Y.; Liu, J.; Li, J.; Wang, C. Part-Aware Mask-Guided Attention for Thorax Disease Classification. *Entropy* **2021**, *23*, 653. [CrossRef] [PubMed]
10. Yang, F.; Lu, P.X.; Deng, M.; Wáng, Y.X.J.; Rajaraman, S.; Xue, Z.; Jaeger, S. Annotations of Lung Abnormalities in the Shenzhen Chest X-rays Dataset for Computer-Aided Screening of Pulmonary Diseases. *Data* **2022**, *7*, 95. [CrossRef] [PubMed]
11. Broască, L.; Truşculescu, A.A.; Ancuşa, V.M.; Ciocârlie, H.; Oancea, C.I.; Stoicescu, E.R.; Manolescu, D.L. A Novel Method for Lung Image Processing Using Complex Networks. *Tomography* **2022**, *8*, 1928–1946. [CrossRef] [PubMed]
12. Mabrouk, A.; Díaz Redondo, R.P.; Dahou, A.; Abd Elaziz, M.; Kayed, M. Pneumonia Detection on Chest X-rays Images Using Ensemble of Deep Convolutional Neural Networks. *Appl. Sci.* **2022**, *12*, 6448. [CrossRef]
13. Khishe, M.; Caraffini, F.; Kuhn, S. Evolving deep learning convolutional neural networks for early COVID-19 detection in chest X-rays images. *Mathematics* **2021**, *9*, 1002. [CrossRef]
14. Cao, F.; Zhao, H. Automatic lung segmentation algorithm on chest X-rays images based on fusion variational auto-encoder and three-terminal attention mechanism. *Symmetry* **2021**, *13*, 814. [CrossRef]
15. Alharithi, F.; Almulihi, A.; Bourouis, S.; Alroobaea, R.; Bouguila, N. Discriminative learning approach based on flexible mixture model for medical data categorization and recognition. *Sensors* **2021**, *21*, 2450. [CrossRef] [PubMed]
16. Rajaraman, S.; Folio, L.R.; Dimperio, J.; Alderson, P.O.; Antani, S.K. Improved semantic segmentation of tuberculosis—Consistent findings in chest X-rays using augmented training of modality-specific u-net models with weak localizations. *Diagnostics* **2021**, *11*, 616. [CrossRef] [PubMed]
17. Shakeri, A.; Konstantelos, N.; Chu, C.; Antoniou, T.; Feld, J.; Suda, K.J.; Tadrous, M. Global Utilization Trends of Direct Acting Antivirals (DAAs) during the COVID-19 Pandemic: A Time Series Analysis. *Viruses* **2021**, *13*, 1314. [CrossRef] [PubMed]
18. Ellis, R.; Hay-David, A.G.C.; Brennan, P.A. Operating during the COVID-19 pandemic: How to reduce medical error. *Br. J. Oral Maxillofac. Surg.* **2020**, *58*, 577–580. [CrossRef] [PubMed]
19. Ou, Y.Y.; Tsai, A.C.; Zhou, X.P.; Wang, J.F. Automatic drug pills detection based on enhanced feature pyramid network and convolution neural networks. *IET Comput. Vis.* **2020**, *14*, 9–17. [CrossRef]
20. Wondmieneh, A.; Alemu, W.; Tadele, N.; Demis, A. Medication administration errors and contributing factors among nurses: A cross sectional study in tertiary hospitals, Addis Ababa, Ethiopia. *BMC Nurs.* **2020**, *19*, 4. [CrossRef] [PubMed]
21. Wang, Y.; Ribera, J.; Liu, C.; Yarlagadda, S.; Zhu, F. Pill recognition using minimal labeled data. In Proceedings of the 2017 IEEE Third International Conference on Multimedia Big Data (BigMM), Laguna Hills, CA, USA, 19–21 April 2017; pp. 346–353.

22. Kwon, H.J.; Kim, H.G.; Lee, S.H. Pill Detection Model for Medicine Inspection Based on Deep Learning. *Chemosensors* **2022**, *10*, 4. [CrossRef]
23. Tan, L.; Huangfu, T.; Wu, L.; Chen, W. Comparison of YOLO v3, Faster R-CNN, and SSD for Real-Time Pill Identification. Available online: https://doi.org/10.21203/rs.3.rs-668895/v1 (accessed on 30 July 2021).
24. Nogueira, K.; dos Santos, J.A.; Menini, N.; Silva, T.S.F.; Morellato, L.P.C.; Torres, R.D.S. Spatio-Temporal Vegetation Pixel Classification by Using Convolutional Networks. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 1665–1669. [CrossRef]
25. Shimizu, A.; Wakabayashi, H.; Kanamori, T.; Saito, A.; Nishikawa, K.; Daisaki, H.; Higashiyama, S.; Kawabe, J. Automated measurement of bone scan index from a whole-body bone scintigram. *Int. J. Comput. Assist. Radiol. Surg.* **2020**, *15*, 389–400. [CrossRef] [PubMed]
26. NIH Chest X-ray Dataset. Available online: https://scidm.nchc.org.tw/dataset/nih-chest-x-ray-dataset (accessed on 2 May 2023).
27. Cubuk, E.D.; Zoph, B.; Shlens, J.; Le, Q.V. Randaugment: Practical automated data augmentation with a reduced search space. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 702–703.
28. VGG Image Annotator. Available online: https://www.robots.ox.ac.uk/vgg/software/via/ (accessed on 26 September 2022).