AIMS Bioengineering, 9(3): 266–282.
DOI: 10.3934/bioeng.2022018
Received: 20 May 2022
Revised: 19 August 2022
Accepted: 22 August 2022
Published: 08 September 2022

http://www.aimspress.com/journal/Bioengineering

Research article

Gas-Net: A deep neural network for gastric tumor semantic segmentation

Lamia Fatiha KAZI TANI[1,*], Mohammed Yassine KAZI TANI[2] and Benamar KADRI[1]

[1] Biomedical Engineering dept. STIC Laboratory, University of Tlemcen, Tlemcen, Algeria
[2] LabRI-SBA Laboratory, The Higher School in Computer Science 08 May 1945 of Sidi Bel, Algeria

* **Correspondence:** Email: lamiakazitani@yahoo.fr; Tel: +213557189747.

Abstract: Currently, the gastric cancer is the source of the high mortality rate where it is diagnoses from the stomach and esophagus tests. To this end, the whole of studies in the analysis of cancer are built on AI (artificial intelligence) to develop the analysis accuracy and decrease the danger of death. Mostly, deep learning methods in images processing has made remarkable advancement. In this paper, we present a method for detection, recognition and segmentation of gastric cancer in endoscopic images. To this end, we propose a deep learning method named GAS-Net to detect and recognize gastric cancer from endoscopic images. Our method comprises at the beginning a preprocessing step for images to make all images in the same standard. After that, the GAS-Net method is based an entire architecture to form the network. A union between two loss functions is applied in order to adjust the pixel distribution of normal/abnormal areas. GAS-Net achieved excellent results in recognizing lesions on two datasets annotated by a team of expert from several disciplines (Dataset1, is a dataset of stomach cancer images of anonymous patients that was approved from a private medical-hospital clinic, Dataset2, is a publicly available and open dataset named HyperKvasir [1]). The final results were hopeful and proved the efficiency of the proposal. Moreover, the accuracy of classification in the test phase was 94.06%. This proposal offers a specific mode to detect, recognize and classify gastric tumors.

Keywords: endoscopic images; artificial intelligence; deep learning; gastric tumor, segmentation

1. Introduction

All around the world, cancer is the disease that is always scary. The gastric cancer is a type of malignant (cancer) cells that form in the lining of the stomach and must be treated by a team of experts from several disciplines. However, the diagnosis of gastric tumour mainly relies on clinical manifestation (abdominal pain and/or vomiting), pathological images and medical imaging [2]. Furthermore, medical imaging-based endoscopy captures, simplifies the interpretation of gastric cancer that is more available to be real, accessible and without effects for the patient [3]. The team of experts based on the diagnosis of the oncologists decide how to proceed and what is the best treatment for the patient. Here, new technologies play an important role.

The artificial intelligence (AI) assists systems in pathology to solve limitations of human in diagnostic errors. So, using AI via deep learning let the improvement of state-of-the-art approaches [4] in various medical fields. In imaging data analysis, the convolutional neural network (CNN) is the well adapted application including a deep learning architecture. Whatever, in current image processing, CNN has accomplished success in objects recognition and classification [5,6]. Sakai et al. have been attracted by gastric endoscopic images; they developed a patch-based classification algorithm based convolutional neural network (CNN) in order to identify gastric cancer in endoscopic images with an accuracy of 87.6% [7]. Nevertheless, as gastric tumor types vary and is more complex, the segmentation of the object that is the tumor creates a challenging task in computer-aided diagnosis, that has to correct performance and instant progress. Deep learning technics performance are certain to conduct such studies of medical images diagnosis in order to have the best results and satisfaction [8]. In 2018, study in [9] had used CNN to classify the genomics data of genes on Pan-Cancer Atlas. In [10], authors had combined the support vector machine (SVM) with random forest (RF) and CNN in classification of the high-grade gliomas, the study showed gain in time of 3.81 than other studies in literature. Furthermore, studies in this concern observed that deep learning technics detect and classify tumors based on histopathological images, that it was been more efficient than that of manual recognition.

Furthermore, [11] propose a method that controls the boundaries of the lungs and heart to measure the CTR (a parameter for cardiac and pulmonary diseases). They are based a segmentation network of two classes, X-RayNet-1 and X-RayNet-2, that aim to provide fine segmentation performance through a minimum of trainable parameters. The precision of segmentation is relative to the precision of the CTR value. So, the most famous technics are Normal distribution, Uniform Approach, K-means. Moreover, the fuzzy logic and the decision tree is developed by [12] for the detection of Colorectal Cancer including 1162 cases in the Masoud internal clinic (Tehran, Iran). This study can recognize a Colorectal Cancer helping healing sooner, and which could decrease the Colorectal Cancer mortality degree.

In this concern, several common segmentation techniques have been developed. An FCN [13] substitutes the fully connected layer of a CNN through an up-sampling layer. Each pixel is considered with the class of object. This is known as “semantic segmentation”. However, the semantic segmentation methods process the entire image, this can cause more processing time. Works such as OVIS [14], are based classical methods for large volumes of the video (a set of images) in databases. In addition, individual objects in an image cannot be recognized. For semantic segmentation, the application of Mask R-CNN [15] to the diagnosis of gastric cancer had a success

in several works [16–18] where, the detection of the tumor region can be surrounded by a bounding box and will be labeled in the mask layer.

In this work, an introduction to a new method of gastric cancer based Deep Learning technics is proposed. The second section displays the state of the art. The third section express problematic, and the fourth section describes the elementary proposal. The last section shows the different experiments and the obtained results confirming the effectiveness of the proposed method. At the end of the paper, a summary and future works are discussed.

2. Related works

The aim of image classification is to attribute one or more label to images according to predefined categories. New technologies based deep learning for object recognition and segmentation prove the best results actually. It is because of the idea to create hidden feature representations based on the Convolutional Neural Networks. This technology is able to analyze and compute more quickly than traditional one. In medical imaging, systems are developed to aid doctors in diagnosis of early cancers.

More applications and studies for classification, detection and segmentation in gastric cancer [19,20], had been advanced where more annotated datasets are available for free or limited access. Authors in [21] developed an automatic network able to learn features on two public datasets. But, most studies analysis of the evolution of gastric cancer progress [22–24]. In this area of research, detection and localization of tumors is very important, it represents an instance of a predefined object that is a set of pixels of an image. Models like YOLO-V2 [25], are used for multi-objects detector on the same image. Here in [21], authors, develop a detector for a small object increasing feature information from the basic network model. Work in [26] shows that bounding box can be annotated basing on label smoothing weighted loss enhancing the feature representation. Also, a new version of faster R-CNN had been presented in [27] called RetinaNet, they use pavement surface images obtained from 2400 m of dual road units.

Another study from [5] describes the sensitivity (SE) that is the detected number of correct gastric cancer lesion/actual number of gastric cancer lesions; And the Positive Predictive Value (PPV), is the detected number of correct gastric cancer lesions/number of lesions that were diagnosed as gastric cancer by a CNN. As a result of this study, the CNN takes 47 seconds to analyze the 2296 test images, it identifies 71 of 77 gastric cancer lesions with a PPV of 30.6% and a sensitivity of 92.2%. It diagnosed 232 total lesions as gastric cancer, 161 of which were non-cancerous and 70 of 71 lesions with diameter upper than 6 mm were correctly detected.

Also, in 2017, authors in [28] introduce an architecture called SegNet that serves to develop a model by [29] to detect real time polyps during colonoscopy as the same study in [30]. Furthermore, for automatic gastric cancer recognition, Hirasawa et al. [31] presented a system that transforms the output space of bounding boxes into a set of default boxes over various aspect ratios and scales per feature, with a sensitivity of 92.2% and a positive predictive value of 30.6%. Similarly, Sakai et al. [7] developed a patch-based classification system-based CNN to identify stomach cancer in endoscopic images, achieving an accuracy of 87.6%. Hence, 2020 has known a project for early detection of gastric cancer from endoscopic images [32], it is based mask R-CNN for instance segmentation areas, then, from different forms of morphologies missing detection of the early gastric cancer, they define 22.2% of the miss rate of gastric cancer.

Lesion based CNN was been developed in [6], it is a form of deep learning algorithm that uses endoscopic images to identify the entire lesion. Endoscopy aids doctors in detecting the GI tract more easily, making it an important aspect of the gastrointestinal (GI) tract examination; yet, the accuracy of its diagnostic results is limited by the practitioner's expertise and the GI system's complex environmental circumstances. Furthermore, when specialists utilize EGD (esophagogastroduodenoscopy), which is a standard procedure for diagnosing stomach cancer, the false negative rate for identifying gastric cancer is 4.6–25.8% [5]. They are based the cafe deep learning framework [4], which is one of the most general and usually used framework. Here, the CNN effectively recognized 71 of 77 gastric cancer lesions with a sensitivity of 92.2 percent and a positive prediction value of 30.6 percent. It diagnosed 232 total lesions as gastric cancer, 161 of which were non-cancerous and 70 of 71 lesions with diameter upper than 6 mm were correctly detected.

Recently, in 2019, Guitao and Zhenwei achieved 61.2 of average prediction gastric cancer classification. These results were very promising compared to previous work [32]. They can explore the improvement of the network structure of Mask R-CNN to make more suitable to detect medical images and using more data augmentation to perform their model. Also, other studies are based Deep Learning methods improved such studies in the areas of prediction and detection of gastric cancer [31,33,34].

3. Problematic

The objective of this work is to detect regions of lesions (abnormal tissue) that mean tumors in gastric organ. To this aim, we develop a method-based U-Net [26] architecture. We named it as GAS-Net. We know that this one is the usually adopted architecture in medical image segmentation. U-Net was initially developed for medical image understanding and segmentation, it was been the core architecture in medical imaging. This architecture is divided into two major paths: a contracting path and an expansive path, the first one comprises several steps of processes related to convolution, normalization, max pooling, activation, concatenation. It aims to extract the key features of the input images and outputs a feature vector of a specific length. The second path extracts information from the first one as a collection, and based on feature vector, it operates in up-convolutions, and by a sequential process, it produces an output segmentation map. The main idea here is the relationship between the first and the second path of the architecture, it permits to the network to achieve highly correct information from the first path, so producing the segmentation mask that converge to the intended output.

The main concept of the GAS-Net architecture consists of establishing a network that begins in input by a set of x images and x groundtruth corresponding masks. From these groundtruth masks and based on probability and generalized function, the network learns the area of tumor from the x learning images, then it proceeds to create a submersion relationship between x learning images and x groundtruth corresponding masks. This is mathematically known as: Ω

$$F^{\#} : T \rightarrow F^{\#} T = \Omega = T \circ F \tag{1}$$

F# is a continuous linear operator. T is the contracting function that determines the characteristic vectors of tumor areas to create a valid region. Similarly, F is the expansive function, it localizes the spatial features in order to detect the abnormal tissue region. Then, Ω corresponds to probabilities

extracted from the x input images. These probabilities are comprised between 0 and 1 and correspond to the final convolutional layer of classification by the SoftMax function. Here, the main motivation is the spatiotemporal representation of key bounding boxes representing the tumor region.

4. Our proposal

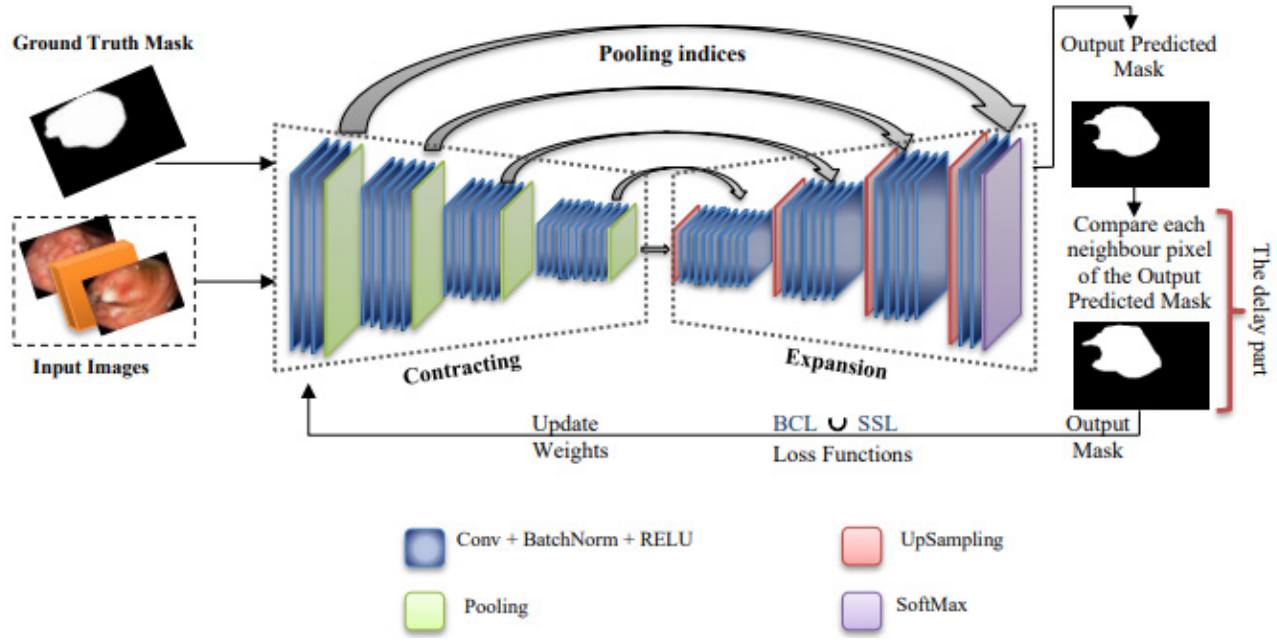

Figure 1. The General Network Architecture of GAS-Net for Semantic Segmentation of Gastric Tumor.

This section describes the GAS-Net method, it aims to detect and recognize gastric tumor area from endoscopic images. The GAS-Net method overview is shown in Figure 1. Typically, it is built on two major parts; the first one comprises the preprocessing step. While the second one, achieves the contracting and expansive processes in order to determine high level features from training set, and also, to detect the spatial information.

4.1. Global process

Our work discusses an efficient approach to overcome the need of classification of endoscopic gastric cancer images. However, to boost the development of image recognition in gastric cancer, we experiment latest technologies relevant to gastric cancer detection and recognition. So, the aim was to detect and recognize tumors regions. Our method is based on two consecutives phases; the first phase is for pre-processing, it is more important because it represents the basis step of our approach, where we apply techniques to the chosen dataset suitable with the training algorithm; the second phase concerns training where we define a new method based on contracting/expansion path. This leads us to add a pseudo-code to eliminate or enlarge the area of the segmentation. Based on

necessities, we adjust and add the hyper-parameters, after that, the combination of predictions and the application of the method to the predicted image and the mask, the outputs were enhanced. Furthermore, our experiments were based on Python language and Tensorflow libraries. This is satisfied by calculation done on GPU Nvidia GTX 1080. Additionally, Adam optimizer was used for training the network.

4.2. Preprocessing

At the beginning, the size of the images varied from 150 × 150 to 1280 × 1080 pixels; and were saved in JPEG format. Moreover, all the images were downsampled to 256 × 256 pixels as of the image matrix size in relationship with the different cases. Via this operation, we normalize diverse images into the different five types of gastric cancer: Adenocarcinomas images, Gastrointestinal stromal tumors (GISTs)images, Neuroendocrine tumors (including carcinoids) images, Lymphomas images, and other cancers images, removing noise, same images, and improved accuracy. "Table 1" resumes the total images for the five types of cancer collected for the dataset1.

4.3. GAS-Net architecture

In order to detect gastric tumor, from endoscopic images, we propose the GAS-Net method, based on the main architecture in medical images that is U-Net. The network of GAS-Net method is provided as input by a pre-processing step and a groundtruth masks. In order to locate feature map information, our method uses on a contracting path to extract all appropriate feature maps.

By increasing the depth of the network, *the contracting path* is generally composed of four residual blocks, each residual block involves two consecutives 3 × 3 convolutions complemented by a ReLU activation unit and a max-pooling layer, it approves understating vanishing gradient in order to save local information for enhancing precision and improving the conception of layers. These residual blocks form a feature map from an input vector to a residual unit's output. However, two inputs are introduced to *the expansion path*, the first one is from the output of the contracting path, the feature map is upsampled using 2 × 2 up-convolution, after that, the collected input from the parallel layer of the contracting path is linked to the upsampled feature map and complemented by two consecutives 3 × 3 convolutions and ReLU activation. Similarly, the second input is from the precedent layer of the expansion path. An applied 1 × 1 convolution is done to decrease the feature map, Softmax activation function is applied to predict the area of tumor (Function (2)).

$$\overrightarrow{\sigma(z)}_t = \frac{e^{z_t}}{\sum_{j=1}^{k} e^{z_j}} \tag{2}$$

After that, we propose to consider all neighbors pixels to the output contours in order to test them if they are able or not to be added to the area. Training the network was been satisfied, the common objects in the image were identified and classified correctly, however, these advances on the model output, leave a small margin to the objects with less quality of segmentation. To this end, with the spatial information of the input image, we propose to design prediction of original image and segmentation area around high overlap pixels. Algorithm1 describes all the process explained above.

ALGORITHME 1. GAS-Net Method

INPUT: E: EGD Images; L: Label; B: Batch-size; W0: initial weights; θ: threshold; A set of predicted pixels Pi; Sample s1…sn; ∇f(x): the stochastic gradient.

OUTPUT: S: mask;

BEGIN

Train, Test =train_test_split(E, L,test_size=0.25);

While epochs ≤ 200 **do**

  For k = 1 to n **do** //each mini batch sample

    z_1= f(En); // contracting

    z_2= f(z_1); // expansion

    Apply Softmax to predict z_2,

    M= minimize(loss_function);

    ∇f(x)=1/n $\sum_{i=1}^{n} \nabla fi(x)$; // compute the stochastic gradient

      Repeate

        w←w +α. Adam optimizer (w, ∇f(x));

        Update weights;

      Until convergenge;

        for each neighbour pixel Pn of S **do**

          if |P_existed- P_neighbor|<θ**then**

            add current pixel to S;

          endif;

        endfor;

        // feedback to the first layer.

  endfor;

endwhile;

END.

In fact, Algorithm 1 describes the process of Gas-Net Method as designed in Figure 1, where the input of the Softmax function is $\vec{z}(z_0, \dots, z_k)$, note that z_i is a real value. The function e^{zj} serves to a positive value of each portion of the input vector. Therefore, k represents the number of classes in the multiclass classifier to associate an effective probability distribution. Next, we calculate the stochastic gradient in order to minimize the loss function explained in paragraph.4.4. The weighted combination of (BCL and SSL) in the loss function allowed a good convergence of the gradient and relevant learning. Adam optimizer selects estimations of gradient to adapt a single learning rate for all weights updated and the learning rate does not change during training. So, here, the Adam optimizer is responsible for updating the weights based on the backpropagation method.

The last step of the process of Algorithm1calculates the color variation between the predicted pixels of segmentation and their neighbors' pixels. Hence, if the variation is less than a threshold θ, we replace the old point by the newest ones. Then, we use a delay followed by loss to delete noisy pixels in the predicted area. At the end, we remove detected area that have a low score.

4.4. *The loss functions*

In order to obtain precise mask and segmentation, we are based impacts of different loss functions in literature. The originality here is to combine two loss functions, the Binary Crossentropy Loss (BCL), and the Sensitivity Specificity Loss (SSL). In these loss functions, the aim is to measure the variance distributions in pixel level classification for segmentation. This is following by an

output value that passes by the Softmax activation function. Here, α means the label and β means the predicted probability by the predicted model.

The BCL function is defined as follow:

$$BCL(\alpha, \beta) = -(\alpha log(\beta) + (1 - \alpha)log(1 - \beta)) \quad (3)$$

$$SSL = \omega * sensitivity + (1 - \omega) * specificity \quad (4)$$

And the Sensitivity Specificity Loss (SSL) is defined as follow:

Where both sensitivity and specificity are based metrics to predict segmentations. This loss function, is added to the first one in order to decrease the inequality of classes via the ω parameter. "Table 2" shows results of loss functions. The advantage of the union between BCL and SSL improves performance compared to calculating BCL alone or SSL alone. To explore these improvements, we show results on different Average precisions (AP, AP50, AP75) and the Average Recall (AR), where different loss functions impact area segmentations. BCL loss runs significantly limited with precision but uncertain pixels for segmentations due to the inequality of classes. However, we enhance equalities weights but this can lead us to an imprecise outline drawing of the mask and consequently to the instance segmentations and consistent masks. On the other hand, SSL loss function also overcome to inequalities of classes but still also imprecise in outline drawing of the area. Thus, joining BCL loss to SSL loss permitted a good convergence of the gradient and relevant learning. We got high performance in the area of gastric tumor localization on the test phase.

5. Experiments

This section describes the whole experiments in order to consider the achievement of our proposed method. We define experimental results on two different datasets, the first one is from our basic collection (See Table 1) and the second one is a public available and open dataset named HyperKvasir [1]. The using of both two datasets is a primordial idea to confirm the robustness of the method. Also, we show the importance of using hyper-parameters in order to control the learning process. At the end, we give results of the different performance metrics in the aim to perform the effectiveness of our method.

5.1. The Dataset1 collection (Table 1)

For an automatic detection of gastric cancer, we had collected anonymously a complete dataset of EGD (Esophagogastroduodenoscopy) images of gastric cancer for our work from the cancer registry from 2011 to 2019 from a private medical-hospital clinic. These images served to establish automatic recognition of tumor based on GAS-Net method. Furthermore, a multidisciplinary team is organized to work together, it englobes gastroenterologist, surgeon, medical oncologist, pathologist, and dietitian. The pathologist observes some lesions from the stomach under a microscope, then, they assign a grade of 1 to 4 to stomach cancer. Low-grade cancer cells are well distinguished, whereas, High-grade cancer cells are poorly differentiated or undifferentiated. This grade plays a very important role for the team to decide on the results and to recognize the different cell kinds in

each image for the training phase. On the other hand, as the world health organization, it exists five types of the stomach cancer: Adenocarcinomas, Gastrointestinal stromal tumors (GISTs), Neuroendocrine tumors (including carcinoids), Lymphomas, and other cancers.

Table 1. Dataset1 collection.

Type of Tumor	Total patients	Grade	Total patients in grade
Adenocarcinomas	102	Grade 1	80
		Grade 3	22
GISTs	46	Grade 2	34
		Grade 3	12
Neuroendocrine	67	Grade 4	67
Lymphomas	75	Grade 3	75
Other tumors	31	-------	31

"Table 1" shows all the data collection where the Adenocarcinomas type comprises 102 patients divided into 2 grades: grade 1,3 which comprise respectively 80 and 22 patients; The GISTs type comprise 46 patients divided into 2 grades: grade 2 and grade 3 which comprise respectively 34 and 12 patients; The Neuroendocrine type comprise 67 patients with grade 4; The Lymphomas type comprise 75 patients with grade 3; And 31 patients are non-identified tumors.

So, in our experiments, we aim to satisfy detection and recognition of the five types of stomach cancer cases. We have selected randomly 75% of the images for training and 25% for validation and test. Figure 2 shows the obtained masks on Dataset1.

Table 2. Results for evaluating different loss functions and their combination.

Average Precision	BCL	SSL	BCL/SSL combination
AP	43.2	43.1	44.6
AP_{50}	45.9	45.0	51.1
AP_{75}	38.3	38.2	42.2
AR	56.7	54.7	66.8

5.2. HyperKvasir dataset

This dataset is publicly available for research and educational purposes. It is divided into four distinct parts; Labeled image data, which contains 23 different classes of results, Unlabeled image data, that contains 99,417 unlabeled images, Segmented image data, that contains 2 folders, one for images and the other is for masks, each of them contains 1,000 JPEG-compressed images, and Annotated video data, that contains approximately 11.62 hours of videos and 1,059,519 video frames, it represents a total of 171 annotated results. Figure 3 shows both in the left original images as well as in the middle their corresponding ground truth masks and in the right their obtained masks.

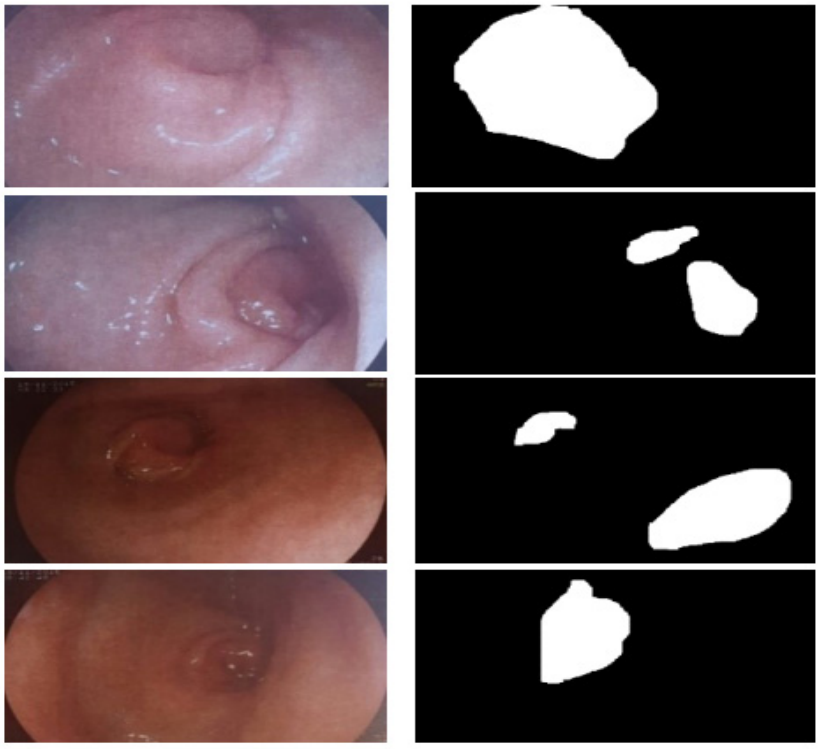

Figure 2. Obtained Masks by GAS-Net on Dataset-1.

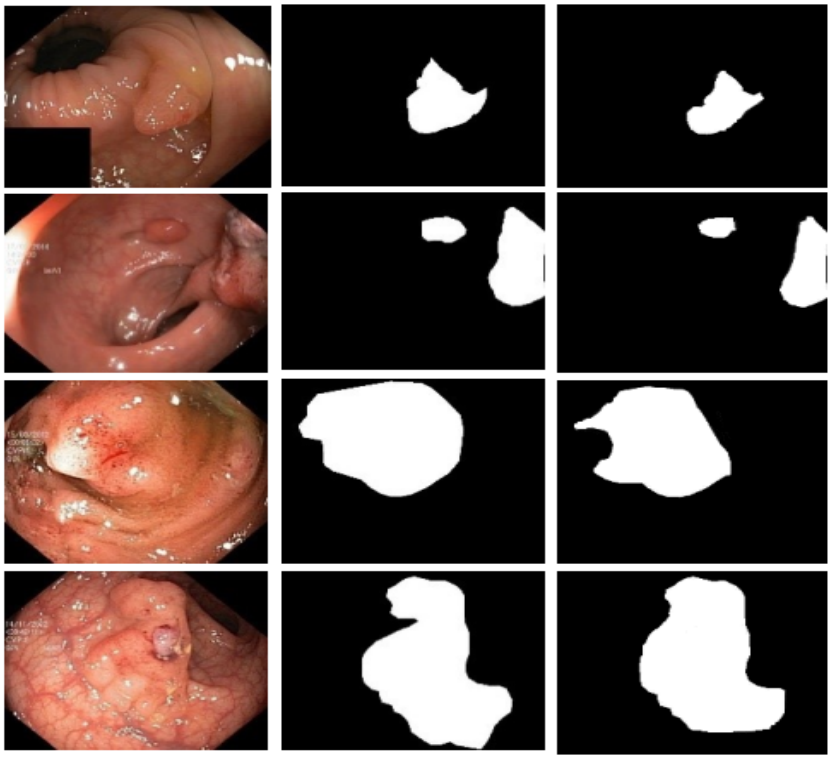

Figure 3. Obtained Masks by GAS-Net on Dataset-2 compared to the Ground-Truth.

5.3. Performances of the network

All the implementations of the network are conducted on a workstation equipped with an Intel i7-6850K CPU with a 64 GB Ram and a single NVIDIA GTX Geforce 1080 Ti GPU and the operating system is Ubuntu 18.04.2., the batch is 64 and learning rate is initialized to 0.002 for all layers and reduced to 0.0002 every 10 epochs from the 200. Also, the aim of all systems in literature is to improve time in training and test phases to be an ergonomic system. In this respect, our study required time in execution at the training and test phases as a variation in accordance to the number of regions of interest to be processed per image, it was between 480 and 890 ms for an image sized 256 × 256.

Moreover, the performance of the network is based on using hyper-parameters. Therefore, we realized diverse experiments in order to select best values of the hyper-parameters and verify achievements of our method.

5.4. Performance metrics

This step of study is fundamental to approve the behaviour of the model and whether the model is achieving its purposes. Diverse performance metrics are used in literature for object detection and image classification such as Accuracy, Precision, Sensitivity, Specificity, and other metrics. The formulas for evaluating these performance measures are represented respectively as below in Eqs 5–10.

$$\text{Accuracy} = (TP+TN)/(TP+TN+FP+FN) \quad (5)$$

$$\text{Precision} = TP/(TP+FP) \quad (6)$$

$$\text{Recall (Sensitivity)} = TP/(TP+FN) \quad (7)$$

$$\text{Specificity} = TN/(TN+FP) \quad (8)$$

$$\text{F-Measure} = 2*(\text{Precision}*\text{Recall})/(\text{Precision}+\text{Recall}) \quad (9)$$

$$\text{Dice} = 2*TP/(TP+FN+FP) \quad (10)$$

Note that TP means the True Positive, TN means the True Negative, FP means the False Positive and FN means the False Negative. The parameters are estimated based on the confusion matrix, offering details about true and false classification of images from all categories. Furthermore, the sensitivity had an impact when an overlap was perceived between the obtained area of tumor and the ground truth created by the team discussed above.

6. Results and discussion

The goal of our segmentation method is to simplify and/or transform the representation of the input image into a set of segments that are more expressive and easier to analyse.

6.1. Comparison study on different backbones

In our study, we achieved several experiments using different architectures as a backbone (VGG16, VGG19, Resnet-50, Resnet-101 and ResNeXt-101), where the top outcomes were gained with the ResNeXt-101 architecture and the Adam optimizer. "Table 3" compares results on the different backbones on the proposed method. Precision, recall and F-measure, determine that our method with ResNeXt-101 as a backbone architecture gives the best result in precision and F-measure, in parallel for the recall, the architecture with ResNet-50 gives the best results compared to the others. "Table 4" resumes a comparison on the diverse training and test performances in order to enrich the study and improve the detection, the recognition and the classification of gastric tumor in images. In order to control the learning process, the impact of hyper-parameters is well seen, they were been used on the model's performance, and primarily at the test phase. It is evident to randomly select images from test dataset to achieve evaluation and consider performance of instance segmentation from the group of experts and the adopted method. This study has satisfied detection and recognition of objects with diverse sizes. Furthermore, this study is based on an improvement in the U-Net model. This improvement is able to draw contours around detected and recognized objects that have the same properties. Hence, experiments support the proposed method achievements and approve success of deep learning approaches for the detection and recognition of tumors.

Table 3. experiments on different backbones.

Backbones	Precision	Recall	F-measure
VGG-16	0.690	0.613	0.649
VGG-19	0.681	0.565	0.617
ResNet-50	0.799	0.721	0.757
ResNet-101	0.590	0.546	0.567
ResNeXt-101	0.875	0.702	0.779

Table 4. Performance's comparison in training and test phases.

Performances	Training	Test
Classification	0.9654	0.9498
Detection/ Recognition	0.9783	0.9634
Segmentation	0.9645	0.9406

Note that predictions with low score (less than 85%) have been eliminated. This is realized when a threshold must be fine-tuned. So, we use a larger threshold to confirm contours around the detected and recognized object. Moreover, identify all parts of images by considering the combination of the same number of each type, which induce an inequality in the dataset, and consequently a bad generalization. It can also be easily seen that predicted contours are clearer and conform to their adequate pixels. Our proposal proves the rapidity of deep learning approaches, in the detection and recognition of gastric tumors at all part of gastric system. Therefore, our proposal demonstrates that it is able to detect all relevant objects, thus, it is more efficient than the natives' models, which can be benefit for gastric diagnosis problems. Also, our proposal includes object detection in the segmentation, it consists originally of locating the tumor, after that, achieving the segmentation of each

recognized object delimited by contours, which permits diverse parts according to the whole features instead of diverse data of each pixel.

Consequently, this study proves results realized through deep learning approaches. Our proposal is robust according to this kind of dataset, it offers a good vision and adoption of its pertinence. Figures 4,5 show the effectiveness of our proposal, where a trigger interrupts the training where the performance measure starts to decrease. Based on our discussions with team of experts at the unit of oncology, these obtained results based on our proposal will allow experts to have accurate results in record time and thus alleviate tedious tasks for oncologists and gastrologist in their daily practice. The main benefit of our proposal is the rapidity of reply compared to natives' methods. This approach can be added into a web application installed on a server in order to be used by several experts in parallel.

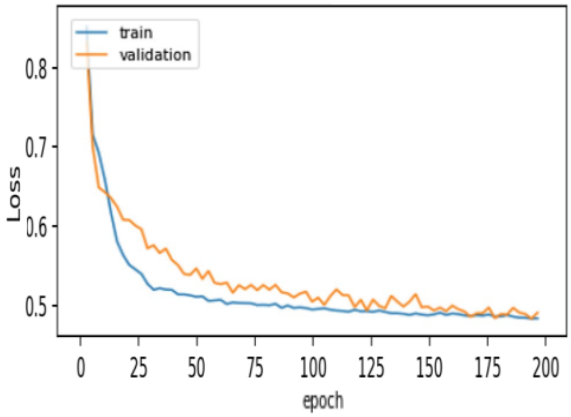

Figure 4. Training and Validation loss.

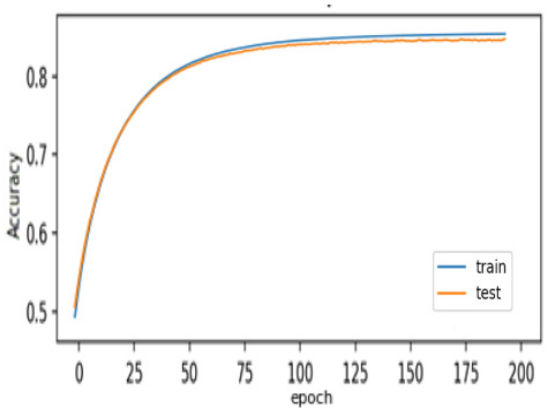

Figure 5. Model Accuracy of GAS-Net.

6.2. Comparison between based models and our proposal on the two datasets

In this sub-section, Table 5 presents obtained results of evaluation on dataset 1 and dataset 2, of our method of endoscopic images compared to based models such as UNET [35], Attention

U-Net [36], DenseUNet [37], Mask R-CNN [15], and UNet++ [38]. Dice, Sensitivity, and Accuracy metrics achieve 89.3%, 92.8% and 95.2% on dataset-1, and 89.3% 89% 94.7% on dataset-2.

Table 5. Performance metrics Comparison between based models and our proposal on the two datasets.

Methods		U-Net	Attention U-Net	DenseUNet	U-Net++	Mask R-CNN	GAS-Net
Sensitivity	*Dataset 1*	0.718	0.723	0.617	0.915	0.798	0.928
	Dataset 2	0.625	0.772	0.594	0.802	0.810	0.890
Accuracy	*Dataset 1*	0.865	0.851	0.732	0.912	0.889	0.952
	Dataset 2	0.747	0.795	0.890	0.906	0.888	0.947
Dice	*Dataset 1*	0.708	0.711	0.571	0.819	0.801	0.893
	Dataset 2	0.682	0.761	0.621	0.821	0.892	0.874

7. Conclusions

In this work, we introduced a smart proposal from new technologies in concern with deep learning. These approaches aim to automatically overcome to problems of detection, recognition and classification of diverse tumors in gastric medical images. A method was developed and adopted and different architectures as a backbone were tested such as VGG-16, VGG-19, Resnet-50, Resnet-101 and ResNeXt-101 where this last one proved the best results. This proposal is able to recognize abnormal tissue inline to the stomach. From the five types of tumors in different grades, a binary mask is calculated to achieve the segmentation. The time of training is minimized for a best generalization. We had used two functions for loss in order to measure the variance distributions and also to decrease the inequality of classes. We had collected a first dataset based on a set of images form a private clinic and a second dataset was introduced to confirm the proposal regarding to the ground truth. Results in experiments demonstrate that the proposal can distinguish lesions from any input image. Also, high precision in segmentation and classification of tumor proves its effectiveness. Our future works consists on developing an extension to the adopted backbone in order to ensure a best instance segmentation of the area of the tumor.

Acknowledgments

This research was partially supported by the Ministry of higher education and scientific research who provided insight and expertise that greatly assisted the research.

Conflict of interest

Authors declare that they have no conflict of interest.

Author contributions

Lamia Fatiha KAZI TANI, Mohammed Yassine KAZI TANI and Benamar KADRI contribute to

realize the presented idea. Lamia Fatiha KAZI TANI developed the theory, achieved the programs and verified the critical methods. Mohammed Yassine KAZI TANI and Benamar KADRI supervised the results of this work. All authors discussed the results and contributed to the final manuscript.

References

1. Borgli H, de Lange T, Eskeland SL, et al. (2020) HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy. *Sci Data* 7: 283. https://doi.org/10.1038/s41597-020-00622-y
2. Cutler J, Grannell A, Rae R (2021) Size-susceptibility of Cornu aspersum exposed to the malacopathogenic nematodes Phasmarhabditis hermaphrodita and P. californica. *Biocontrol Sci Technol* 31: 1149–1160. https://doi.org/10.1080/09583157.2021.1929072
3. Kitagawa Y and Wada N (2021) Application of AI in Endoscopic Surgical Operations, In: Seiichi Takenoshita, Hiroshi Yasuhara, *Surgery and Operating Room Innovation*, Singapore: Springer, 71–77. https://doi.org/10.1007/978-981-15-8979-9_8
4. Jia Y, Shelhamer E, Donahue J, et al. (2014) Caffe: Convolutional architecture for fast feature embedding. *Proceedings of the 22nd ACM international conference on Multimedia*. MM14: 675–678. https://doi.org/10.1145/2647868.2654889
5. Hirasawa T, Aoyama K, Tanimoto T, et al. (2018) Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images. *Gastric Cancer.* 21: 653–660. https://doi.org/10.1007/s10120-018-0793-2
6. Kim JH and Yoon HJ (2020) Lesion-based convolutional neural network in diagnosis of early gastric cancer. *Clin Endosc* 53: 127–131. https://doi.org/10.5946/ce.2020.046
7. Sakai Y, Takemoto S, Hori K, et al. (2018) Automatic detection of early gastric cancer in endoscopic images using a transferring convolutional neural network. *In 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE: 4138–4141. https://doi.org/10.1109/EMBC.2018.8513274
8. de Groof AJ, Struyvenberg MR, van der Putten J, et al. (2020) Deep-learning system detects neoplasia in patients with barrett's esophagus with higher accuracy than endoscopists in a multistep training and validation study with benchmarking. *Gastroenterology* 158: 915–929. e4. https://doi.org/10.1053/j.gastro.2019.11.030
9. Haque A and Lyu B (2018) Deep learning based tumor type classification using gene expression data. *Proceedings of the 2018 ACM international conference on bioinformatics, computational biology, and health informatics.* 18: 89–96. https://doi.org/10.1145/3233547.3233588
10. Wang P, Xiao X, Glissen Brown JR, et al. (2018) Development and validation of a deep-learning algorithm for the detection of polyps during colonoscopy. *Nat Biomed Eng* 2: 741–748. https://doi.org/10.1038/s41551-018-0301-3
11. Arsalan M, Owais M, Mahmood T, et al. (2020) Artificial intelligence-based diagnosis of cardiac and related diseases. *J Clin Med* 9: 871. https://www.mdpi.com/2077-0383/9/3/871
12. Nopour R, Shanbehzadeh M, Kazemi-Arpanahi H (2021) Developing a clinical decision support system based on the fuzzy logic and decision tree to predict colorectal cancer. *Med J Islam Repub Iran* 35: 44. https://doi.org/10.47176/mjiri.35.44

13. Darrell T, Long J, Shelhamer E (2016) Fully convolutional networks for semantic segmentation. *IEEE transactions on pattern analysis and machine intelligence.* 39: 640–651. https://doi.org/10.1109/TPAMI.2016.2572683
14. Ghomari A, Kazi Tani MY, Lablack A, et al. (2017) OVIS: ontology video surveillance indexing and retrieval system. *Int J Multimed Info Retr* 6: 295–316. https://doi.org/10.1007/s13735-017-0133-z
15. He K, Gkioxari G, Dollár P, et al. (2017) Mask r-cnn. *In Proceedings of the IEEE international conference on computer vision*. IEEE: 2961–2969. https://doi.org/10.48550/arXiv.1703.06870
16. Tani LFK, Ghomari A, Tani MYK (2019) A semi-automatic soccer video annotation system based on Ontology paradigm. *2019 10th International Conference on Information and Communication Systems. ICICS:* 88–93. https://doi.org/0.1109/IACS.2019.8809161
17. KAZI TANI LF, Conception and Implementation of a Semi-Automatic Tool Dedicated To The Analysis And Research Of Web Video Content. Doctoral Thesis, Oran, Algeria, 2020. Available from: https://theses.univ-oran1.dz/document/TH5141.pdf.
18. Tani LFK, Ghomari A, Tani MYK (2019) Events recognition for a semi-automatic annotation of soccer videos: a study based deep learning. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences.* 42: 135–141. https://doi.org/10.5194/isprs-archives-XLII-2-W16-135-2019
19. Sun JY, Lee SW, Kang MC, et al. (2018) A novel gastric ulcer differentiation system using convolutional neural networks. *2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE: 351–356. https://doi.org/10.1109/CBMS.2018.00068
20. Martin DR, Hanson JA, Gullapalli RR, et al. (2020) A deep learning convolutional neural network can recognize common patterns of injury in gastric pathology. *Arch Pathol Lab Med* 144: 370–378. https://doi.org/10.5858/arpa.2019-0004-OA
21. Gammulle H, Denman S, Sridharan S, et al. (2020) Two-stream deep feature modelling for automated video endoscopy data analysis. *International Conference on Medical Image Computing and Computer-Assisted Intervention.* Springer, Cham, 2020: 742–751. https://doi.org/10.1007/978-3-030-59716-0_71.
22. Li Z, Togo R, Ogawa T, et al. (2019) Semi-supervised learning based on tri-training for gastritis classification using gastric X-ray images. *2019 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE: 1–5. https://doi.org/10.1109/ISCAS.2019.8702261
23. Kanai M, Togo R, Ogawa T, et al. (2019) Gastritis detection from gastric X-ray images via fine-tuning of patch-based deep convolutional neural network. *2019 IEEE International Conference on Image Processing (ICIP).* IEEE: 1371–1375. https://doi.org/0.1109/ICIP.2019.8803705
24. Cho BJ, Bang CS, Park SW, et al. (2019) Automated classification of gastric neoplasms in endoscopic images using a convolutional neural network. *Endoscopy* 51: 1121–1129. https://doi.org/10.1055/a-0981-6133
25. Redmon J, Divvala S, Girshick R, et al. (2015) You look only once: unified real-time object detection. *arXiv preprint arXiv*:1506.02640. https://doi.org/10.48550/arXiv.1506.02640
26. Islam M, Seenivasan L, Ming LC, et al. (2020) Learning and reasoning with the graph structure representation in robotic surgery. *International Conference on Medical Image Computing and Computer-Assisted Intervention.* Cham, Springer: 627–636. https://doi.org/10.1007/978-3-030-59716-0_60.

27. Tran VP, Tran TS, Lee HJ, et al. (2021) One stage detector (RetinaNet)-based crack detection for asphalt pavements considering pavement distresses and surface objects. *J Civil Struct Health Monit* 11: 205–222. https://doi.org/0.1007/s13349-020-00447-8.
28. Badrinarayanan V, Cipolla R, Kendall A (2017) Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*. 39: 2481–2495. https://doi.org/10.1109/TPAMI.2016.2644615
29. Urbanos G, Martín A, Vázquez G, et al. (2021) Supervised machine learning methods and hyperspectral imaging techniques jointly applied for brain cancer classification. *Sensors* 21: 3827. https://doi.org/10.3390/s21113827
30. Urban G, Tripathi P, Alkayali T, et al. (2018) Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. *Gastroenterology* 155: 1069–1078.e8. https//doi.org/10.1053/j.gastro.2018.06.037
31. Li L, Chen Y, Shen Z, et al. (2020) Convolutional neural network for the diagnosis of early gastric cancer based on magnifying narrow band imaging. *Gastric Cancer* 23: 126–132. https://doi.org/10.1007/s10120-019-00992-2
32. Ishihara K, Ogawa T, Haseyama M (2017) Detection of gastric cancer risk from X-ray images via patch-based convolutional neural network. *2017 IEEE International Conference on Image Processing (ICIP).* IEEE: 2055–2059. https://doi.org/10.1109/ICIP.2017.8296643
33. Song Z, Zou S, Zhou W, et al. (2020) Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning. *Nat Commun* 11: 1–9. https://doi.org/10.1038/s41467-020-18147-8
34. Ikenoyama Y, Hirasawa T, Ishioka M, et al. (2021) Detecting early gastric cancer: Comparison between the diagnostic ability of convolutional neural networks and endoscopists. *Digest Endosc* 33: 141–150. https://doi.org/10.1111/den.13688
35. Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention.* Cham, Springer: 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
36. Folgoc LL, Heinrich M, Lee M, et al. (2018) Attention u-net: Learning where to look for the pancreas. Available from: https://arxiv.org/pdf/1804.03999.pdf%E4%BB%A3%E7%A0%81%E5%9C%B0%E5%9D%80%EF%BC%9Ahttps://github.com/ozan-oktay/Attention-Gated-Networks
37. Li X, Chen H, Qi X, et al. (2018) H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes. *IEEE transactions on medical imaging.* 37: 2663–2674. https://doi.org/10.1109/TMI.2018.2845918
38. Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, et al. (2018) Unet++: A nested u-net architecture for medical image segmentation. *Deep learning in medical image analysis and multimodal learning for clinical decision support.* Cham, Springer:: 3–11. https://doi.org/10.1007/978-3-030-00889-5_1