Polyp Detection in Gastrointestinal Images using Faster Regional Convolutional Neural Network

Lourdes Duran-Lopez, Francisco Luna-Perejon, Isabel Amaya-Rodriguez, Javier Civit-Masot, Anton Civit-Balcells, Saturnino Vicente-Diaz and Alejandro Linares-Barranco
Robotics and Computer Technology Lab., University of Seville, Av. Reina Mercedes s/n, Seville, Spain

Keywords: Polyp, Colonoscopy, Deep Learning, Image Analysis, Faster Regional Convolutional Neural Network.

Abstract: Colorectal cancer is the third most frequently diagnosed malignancy in the world. To prevent this disease, polyps, the principal precursor, are removed during a colonoscopy. Automatic detection of polyps in this technique could play an important role to assist doctors for achieving an accurate diagnosis. In this work, we apply a state-of-the-art Deep Learning algorithm called Faster Regional Convolutional Neural Network to each colonoscopy frame in order to detect the presence of polyps. The proposed detection system contains two main stages: (1) processing of the colonoscopy frames for training and testing datasets generation, where artifacts are extracted and the number of images in the dataset is augmented; and (2) the Neural Network model, which performs feature extraction over the frames in order to detect polyps within the frames. After training the algorithm under different conditions, our result shows that the proposed system detection has a precision of 80.31%, a recall of 75.37%, an accuracy of 71.99% and a specificity of 65.70%.

1 INTRODUCTION

Colorectal cancer (CRC) is the third most frequently diagnosed malignancy and the fourth leading cause of cancer-related deaths in the world. Its burden is expected to increase by 60% to more than 2.2 million new cases and 1.1 million deaths from cancer by 2030 (Arnold et al., 2017). Certain types of polyps that grow in the colon are the precursors of more than 95% of cancer cases in this organ. Thus, the removal of polyps is usually practiced as a safety measure.

Polyps are detected and extracted during a colonoscopy, which is a medical analysis in which the colonoscopist examines the meter and a half length tube of the large intestine with the use of an colonoscope. This technique is a commonly used primary method to detect polyps (Lieberman et al., 2012). However, by using this technique, the polyp detection is not always effective (Corley et al., 2014), as there are some factors that make the correct evaluation difficult. These factors may be related to the experience of the physician, the conditions of the patient, the quality of the instruments, among others. Highly experienced colonoscopists can achieve an accuracy in the detection that is greater than 90%, while colonoscopists with less experience cannot perform better than an 80% in the recognition rate (Hewett et al., 2012) (Ignjatovic et al., 2009). On the other hand, there may be other external factors, such as the ambiguity in the characteristics of the polyp. Polyps could be difficult and confusing to diagnose depending on the characteristics shown when performing the colonoscopy (Hewett et al., 2012). An automated computerized diagnosis could provide a cross validation mechanism in order to assist colonoscopists to make a decision in this task and, thus, reducing the possibility of committing mistakes in the diagnosis and improving the quality of life of the affected patients. This technique could also be used for training purpose of new colonoscopists at the university or medical centers.

In the past few years, several researches investigated the problem of detecting polyps in colonoscopies (Bernal et al., 2017), (Zhang et al., 2017), (Ribeiro et al., 2016), (Axyonov et al., 2016), (Park and Sargent, 2016). Many of them used conventional Convolutional Neural Networks (CNNs) to classify frames obtained from videos, indicating the presence of a polyp. However, none of these solutions are able to locate the polyp inside the image, which could be very useful when making a colonoscopy to aid the experts in their task.

In this paper, we report our study of an automated computerized system for detecting polyps in colonoscopy videos using a Deep Learning based al-

Duran-Lopez, L., Luna-Perejon, F., Amaya-Rodriguez, I., Civit-Masot, J., Civit-Balcells, A., Vicente-Diaz, S. and Linares-Barranco, A.
Polyp Detection in Gastrointestinal Images using Faster Regional Convolutional Neural Network.
DOI: 10.5220/0007698406260631
In *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2019)*, pages 626-631
ISBN: 978-989-758-354-4

gorithm called Regional Convolutional Neural Networks (R-CNN). These methods try to determine the presence of an object and also locate its position inside the image. One of the most recent algorithms based on the R-CNN approach are called Faster Regional Convolutional Neural Network (Faster R-CNN), which were developed in 2015 (Ren et al., 2015). Faster R-CNNs were designed with the purpose of reducing its predecessors computation time. In ILSVRC and COCO 2015 competitions, Faster R-CNN was the foundations of the 1st-place winning entries in several tracks.

Faster R-CNNs have been used for different purposes: face detection (Jiang and Learned-Miller, 2017), driver's cell-phone usage and hands on steering wheel detection (Hoang Ngan Le et al., 2016) are some application examples of this algorithm, which has proven to show good results.

In this paper, the authors present an application of Faster R-CNNs in order to build an automated recognition system to detect the presence of polyps in colonoscopy images.

The rest of the paper is structured as follows: section 2 presents the methods that are used in the approach that has been carried out in this paper. First, the dataset used is explained from its acquisition to its preprocessing and augmentation. Then, the Faster R-CNN algorithm is presented, along with the architecture, design, training and testing of the neural network that has been used in this work, which is based on those algorithms. Also, to evaluate the system, some tests to prove its performance and robustness are performed. Then, Section 3 presents the classification results, and Section 4 a discussion over the obtained values. Finally, the conclusions of this work are presented in Section 5.

2 MATERIALS AND METHODS

In this section the main materials and methods used in this work are presented, distinguishing between the dataset and the neural network. For the dataset, the applied stages are: image acquisition, pre-processing and data augmentation. For the neural network subsection, the architecture of the Faster R-CNN, the training and testing steps for polyp detection in colonoscopy images and the performance evaluation are presented.

A general view of our study can be seen in the block diagram that is shown in (Fig. 1).

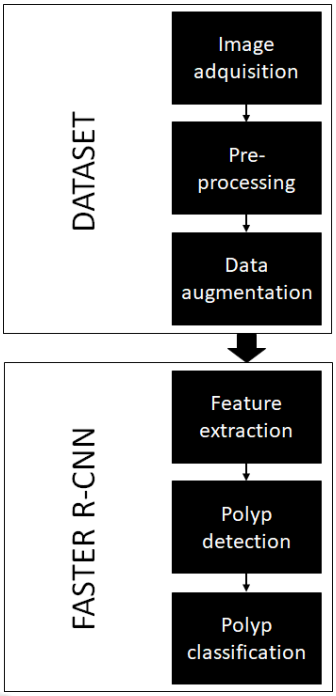

Figure 1: Block diagram of the implemented approach.

2.1 Dataset

2.1.1 Image Acquisition

We used the database provided by Endoscopic Vision's sub-challenge (MICCAI 2018[1]) called Gastrointestinal Image ANAlysis (GIANA[2]) (Angermann et al., 2017) (Bernal et al., 2018), which is composed of 18 videos collected from colonoscopies from different patients. These videos are already provided as a segmented set of frames. The dataset contains a total of 11954 images with their corresponding masks indicating where the polyp is located (only in the case that there was one in that specific image) (Fig. 2). All images have a resolution of 384x288 pixels.

2.1.2 Artifact Extraction

In order to enhance the performance of the system, we applied a pre-processing step to the images to make the network be able to extract descriptors more appropriately and, therefore, improve the results of the detection (Fig. 3). For this, we extracted the black edges of the endoscopy frames to remove the unwanted areas, reducing the images to a size of 284x265.

[1]https://www.miccai2018.org/en/
[2]https://giana.grand-challenge.org/

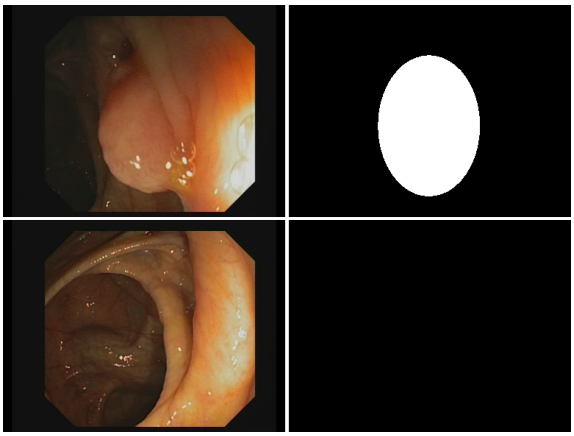

Figure 2: Samples of the database. Left: frames captured from the colonoscopy video recordings. Right: binary masks indicating if there is a polyp (marked as a white area) and its location inside the frame.

2.1.3 Data Augmentation

Deep learning models perform better when training with large datasets that include more variability in the samples. The most popular method to make datasets bigger is called data augmentation, which helps prevent overfitting when training on very little data (Wong et al., 2016). The simplest way to perform data augmentation is to add noise or apply geometric transformations to data to simulate other conditions of camera recordings with the same objects (polyps in our case). Hence, all the transformations would boost the models to learn better.

The dataset was augmented using a series of transformations so that the model would never train twice the exact same image. For each original pre-processed image, an horizontal flip, a vertical flip and a blur filter have been applied. Thus, we obtain three new images from each original sample (Fig. 3). After this data augmentation step, we obtain a dataset that consists of 47.816 images in total.

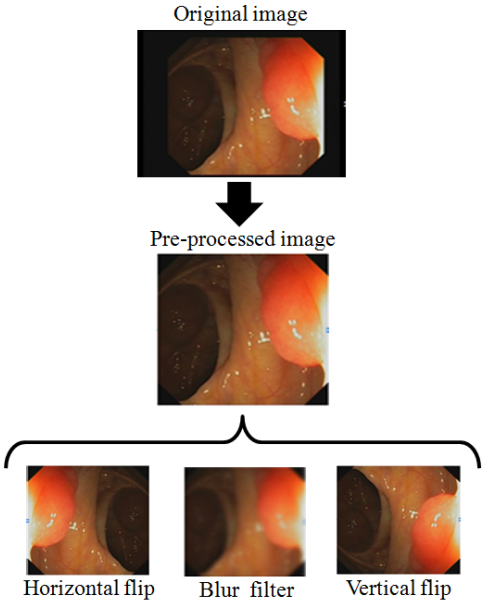

Figure 3: Processing applied to the original images. First, black edges are removed in a pre-processing step. Then data augmentation is applied, generating three different new images.

2.2 Faster R-CNN

2.2.1 Arquitecture

As it was mentioned in the introduction, a Faster RCNN was used in this work for the polyp detection task. This algorithm is divided in two modules (Fig. 4):

– First of all, a deep Fully Convolutional Network (FCN) (Ren et al., 2015) receives the images from the dataset that was presented in this section as input. Then, it extracts feature maps or descriptive characteristics and analyzes them to propose regions of interest. The novel step that this architecture introduces is the way to determine the regions of interest, by using a neural network that takes advantage of the mathematical operations made in the convolution layers. In our architecture we have used the ResNet50 model (He et al., 2016) as FCN. ResNet models try to solve the saturation of the accuracy caused by increasing the network depth.

– Secondly, the proposed regions are the input of the second module, called Fast RCNN detector, composed by two fully-connected layers, a regression layer and a classification layer (Ren et al., 2015).

2.2.2 Training and Testing

In order to reduce the number of false positives, a technique called hard-negative mining was used (Felzenszwalb et al., 2010). It consists on adding negative samples, which means, including examples of images that do not contain polyps in the training step, labeling them as background.

85% of the augmented dataset (15 videos) was used to train the network, while the remaining 15% (3 videos) for evaluating its accuracy and robustness in the detection. These training and testing dataset are obtained in such way that different patients are involve in each dataset.

In this work, both TensorFlow[3] and Keras[4], which

[3]https://www.tensorflow.org/?hl=es

[4]https://keras.io/

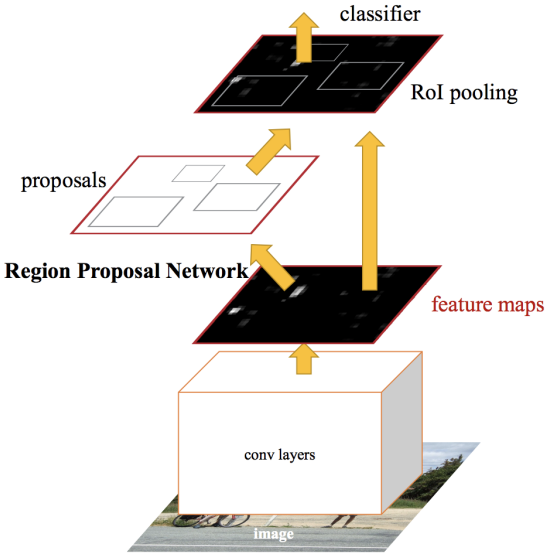

Figure 4: Faster R-CNN architecture. Image taken from (Ren et al., 2015).

are two widely known Deep Learning frameworks, are used to design, train and test the network.

2.2.3 Performance Evaluation Measures

To quantitatively present the capabilities of the proposed computerized-aided diagnosis method, performance metrics were used to show how the implemented technique was able to detect and classify the polyps. These are precision (Equation 1), recall (Equation 2), accuracy (Equation 3), specificity (Equation 4), F1-score (Equation 5) and F2-score (Equation 6).

$$Precision = 100 \times \frac{TP}{TP+FP} \quad (1)$$

$$Recall = 100 \times \frac{TP}{TP+FN} \quad (2)$$

$$Accuracy = 100 \times \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$Specificity = 100 \times \frac{TN}{TN+FP} \quad (4)$$

$$F1-score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (5)$$

$$F2-score = 5 \times \frac{Precision \times Recall}{4 \times Precision+Recall} \quad (6)$$

where, TP and FP denote true positive cases (the system detects a polyp in a frame that contains a polyp) and false positive cases (the system detects a polyp in a frame that does not have a polyp), respectively. TN and FN denote true negative cases (the system does not detect a polyp in a frame that does not have a polyp) and false negative cases (the system does not detect a polyp in a frame that contains a polyp), respectively.

These metrics was applied to the classification layer of this method. Since this layer returns as output a confidence value, a minimal threshold was established to apply these metrics. On this way, if the confidence value exceeds the threshold then the polyp is detected.

3 RESULTS

Different experiments were carried out to determine if polyps were detected correctly or not. Tests were performed every 50 epochs, selecting different confidence thresholds in order to obtain the best results. Polyps detection performance is reported in Table 1. The results show the robustness of the proposed Faster R-CNN architecture on detecting the polyp position in colonoscopy images with a precision of 80.31%, a recall of 75.37%, an accuracy of 71.99% and a specificity of 65.70%. The minimal threshold was established at 0.80.

In Fig. 5 the results of our recognition system can be seen by showing the precision when detecting polyps inside samples from the dataset, and their corresponding mask images (as a ground truth) indicating where the polyps are located.

4 DISCUSSION

In this work, we have performed a polyp detection system using Faster R-CNN. This kind of networks have had a very high impact in recognition systems in recent years thanks to the fact that they are able to both classify objects or items inside images and also provide the specific location in which that item is. The application of this kind of networks in medical image analysis is becoming more and more usual. In this case, this work presents, what we believe it is, the first application of Faster R-CNNs to polyp detection in videos obtained from colonoscopies. As it can be seen, we have focused on the polyp detection task but, as we mentioned earlier, Faster R-CNNs can also provide the location of polyps inside images. The results obtained show that there is still room to improve and we will continue to pursue the goal of implementing an automatic polyp detection system that

Table 1: Polyps detection performance. TP: True Positive, FP: False Positive, TN: True Negative, FN: False Negative.

TP	FP	TN	FN	Precision	Recall	Accuracy	Specificity	F1	F2
3533	866	1659	1154	80.31	75.37	71.99	65.70	77.76	76.30

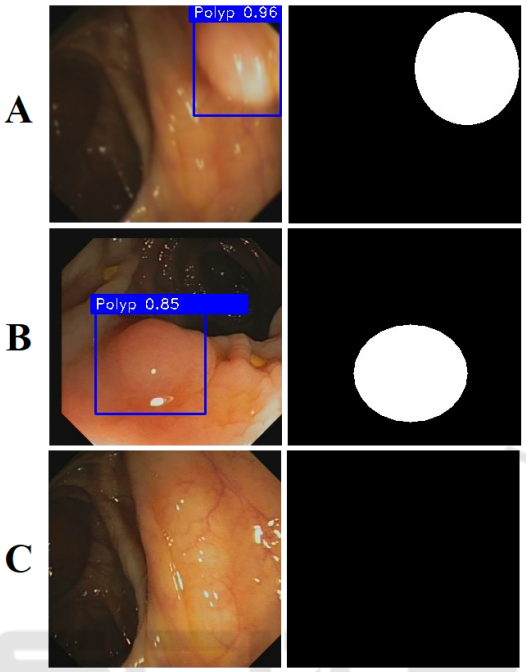

Figure 5: Polyp detection results. Left: polyps detected by Faster R-CNN. Confidence values are represented in blue. Right: their corresponding ground truth. A and B show the performance in case a polyp appears, while C shows the performance in case there is no polyp.

could be used in hospitals as an aid for colonoscopists in their task. This future work will include the polyp location task on top of the already performed detection system by adding a new analysis layer and a set of metrics that would improve the overall recognition system.

For now, we have achieved first results using this aforementioned Deep Learning technique. Designing a new and more appropriate CNN model instead of using the well-known ResNet50, along with fine-tuning the hyperparameters of it could improve the classification results not only in terms of accuracy but also in terms of specificity and sensitivity. Also, other methods for pre-processing the data and increasing the variability of the dataset will make the recognition system more robust and plausible to use in a real case scenario. These two approaches will be studied in future works in which we will try to improve state-of-the-art solutions and also the accuracy achieved by colonoscopist in order to build a system that could work together with them and aid them in their recognition task.

5 CONCLUSIONS

In this paper, we present the use of Faster R-CNN as a mechanism for a computarized-aided diagnosis system that automatically detects polyps in colonoscopies in order to assist doctors in this task. The images were obtained from GIANA Endoscopic Vision's subchallenge to create the dataset. These samples from the dataset were first pre-processed, removing the unwanted areas and, after that, increased in number by using data augmentation techniques. 85% of dataset was used to train the network, which, on a first step, extracts features over the colonoscopy frames; then, it proposes regions of interests where polyps could be located; and finally, it classifies the regions most likely to contain a polyp. The remaining 15% of dataset was used to test the network in order to evaluate its performance. The results show that the proposed recognition system has a precision of 80.31%, a recall of 75.37%, an accuracy of 71.99% and a specificity of 65.70%. The system is able to detect the presence of polyps in the images studied in this paper and the results prove that the system could be used as an aid for doctors in this task. In future works, the authors will study and evaluate different custom CNN models in order to improve the performance, classification results and also the polyp location task, along with new pre-processing algorithms to improve the variability of the dataset and the feature extraction step.

ACKNOWLEDGEMENTS

This work was supported by the Spanish grant (with support from the European Regional Development Fund) COFNET (TEC2016-77785-P).

REFERENCES

Angermann, Q., Bernal, J., Sánchez-Montes, C., Hammami, M., Fernández-Esparrach, G., Dray, X., Romain, O., Sánchez, F. J., and Histace, A. (2017). Towards real-time polyp detection in colonoscopy videos: Adapting still frame-based methodologies for video sequences analysis. In *Computer Assisted and Robotic Endoscopy and Clinical Image-Based Procedures*, pages 29–41. Springer.

Arnold, M., Sierra, M. S., Laversanne, M., Soerjomataram, I., Jemal, A., and Bray, F. (2017). Global patterns

and trends in colorectal cancer incidence and mortality. *Gut*, 66(4):683–691.

Axyonov, S., Liang, J., Kostin, K., Zamyatin, A. V., et al. (2016). Advanced pattern recognition and deep learning for colon polyp detection.

Bernal, J., Histace, A., Masana, M., Angermann, Q., Sánchez-Montes, C., Rodriguez, C., Hammami, M., Garcia-Rodriguez, A., Córdova, H., Romain, O., et al. (2018). Polyp detection benchmark in colonoscopy videos using gtcreator: A novel fully configurable tool for easy and fast annotation of image databases. In *Proceedings of 32nd CARS conference*.

Bernal, J., Tajkbaksh, N., Sánchez, F. J., Matuszewski, B. J., Chen, H., Yu, L., Angermann, Q., Romain, O., Rustad, B., Balasingham, I., et al. (2017). Comparative validation of polyp detection methods in video colonoscopy: results from the miccai 2015 endoscopic vision challenge. *IEEE transactions on medical imaging*, 36(6):1231–1249.

Corley, D. A., Jensen, C. D., Marks, A. R., Zhao, W. K., Lee, J. K., Doubeni, C. A., Zauber, A. G., de Boer, J., Fireman, B. H., Schottinger, J. E., et al. (2014). Adenoma detection rate and risk of colorectal cancer and death. *New england journal of medicine*, 370(14):1298–1306.

Felzenszwalb, P. F., Girshick, R. B., McAllester, D., and Ramanan, D. (2010). Object detection with discriminatively trained part-based models. *IEEE transactions on pattern analysis and machine intelligence*, 32(9):1627–1645.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.

Hewett, D. G., Kaltenbach, T., Sano, Y., Tanaka, S., Saunders, B. P., Ponchon, T., Soetikno, R., and Rex, D. K. (2012). Validation of a simple classification system for endoscopic diagnosis of small colorectal polyps using narrow-band imaging. *Gastroenterology*, 143(3):599–607.

Hoang Ngan Le, T., Zheng, Y., Zhu, C., Luu, K., and Savvides, M. (2016). Multiple scale faster-rcnn approach to driver's cell-phone usage and hands on steering wheel detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 46–53.

Ignjatovic, A., East, J. E., Suzuki, N., Vance, M., Guenther, T., and Saunders, B. P. (2009). Optical diagnosis of small colorectal polyps at routine colonoscopy (detect inspect characterise resect and discard; discard trial): a prospective cohort study. *The lancet oncology*, 10(12):1171–1178.

Jiang, H. and Learned-Miller, E. (2017). Face detection with the faster r-cnn. In *Automatic Face & Gesture Recognition (FG 2017), 2017 12th IEEE International Conference on*, pages 650–657. IEEE.

Lieberman, D. A., Rex, D. K., Winawer, S. J., Giardiello, F. M., Johnson, D. A., and Levin, T. R. (2012). Guidelines for colonoscopy surveillance after screening and polypectomy: a consensus update by the us multi-society task force on colorectal cancer. *Gastroenterology*, 143(3):844–857.

Park, S. Y. and Sargent, D. (2016). Colonoscopic polyp detection using convolutional neural networks. In *Medical Imaging 2016: Computer-Aided Diagnosis*, volume 9785, page 978528. International Society for Optics and Photonics.

Ren, S., He, K., Girshick, R., and Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems*, pages 91–99.

Ribeiro, E., Uhl, A., Wimmer, G., and Häfner, M. (2016). Exploring deep learning and transfer learning for colonic polyp classification. *Computational and mathematical methods in medicine*, 2016.

Wong, S. C., Gatt, A., Stamatescu, V., and McDonnell, M. D. (2016). Understanding data augmentation for classification: when to warp? *arXiv preprint arXiv:1609.08764*.

Zhang, R., Zheng, Y., Mak, T. W. C., Yu, R., Wong, S. H., Lau, J. Y., and Poon, C. C. (2017). Automatic detection and classification of colorectal polyps by transferring low-level cnn features from nonmedical domain. *IEEE J. Biomedical and Health Informatics*, 21(1):41–47.