

DENSE-RES NET FOR ENDOSCOPIC IMAGE CLASSIFICATION

Quoc-Huy Trinh and Minh-Van Nguyen

Department of Computer Engineering, Ho Chi Minh University of Science,
Ho Chi Minh City, Vietnam

ABSTRACT

We propose a method that configures Fine-tuning to a combination of backbone DenseNet and ResNet to classify eight classes showing anatomical landmarks, pathological findings, to endoscopic procedures in the GI tract. Our Technique depends on Transfer Learning which combines two backbones, DenseNet 121 and ResNet 101, to improve the performance of Feature Extraction for classifying the target class. After experiment and evaluating our work, we get accuracy with an F1 score of approximately 0.93 while training 80000 and test 4000 images.

KEYWORD

Kvasir dataset, dense-res, medical image, classification, deep neural network.

1. INTRODUCTION

In recent years, the number of people that have been affected by colorectal cancer (CLC) is increasing. It is also on a third of the world for many years. However, can we diagnose and prevent CLC is a crucial issue for the health organization. Some studies illustrate that almost 95% of CLC is from the adenomatous polyp. The resection of Colorectal adenomatous polyps can reduce the CLC. On the other hand, the best way to deal with CLC is early diagnosis and have straight treatment.

Nowadays, the growing up of population is parallel with the increase of CLC also the number of people accepts for the CLC examination is getting higher. However, the detection technique for polyps in the past and in some country, current is dependent on almost all human tasks by doctors experienced and ability, which can easily be affected by environmental factor and can be inefficient.[1]

Recently, there are many approaches to classification these classes by using backbones, such as ResNet 50, Densnet 169, Efficientnet, etc. Almost all results of these approaches are high, but there is an issue that is the bias for some classes in the dataset. For this reason, we proposed Dense-Res Net for classification endoscopic images.

2. RELATED WORK

In later years, there are many kinds of research in the classification, detection of Gastrointestinal (GI) and endoscopic images. Almost all research uses Deep Learning with Deep Convolution architecture such as LeNet, AlexNet, ResNet, DenseNet and GoogleNet [7]. The results of those research are higher and cost low computational. Distinct from simple CNNs, the Deep

Convolution architectures extract feature better and more obviously [8]. Furthermore, these features are good to localize the area of symptoms on the images.[5]

In 2021, research creates a framework that uses a three-stage framework for diagnosing gastrointestinal diseases. They have three stages:

- First is the Image pre-processing step which is feature extraction by Deep architecture
- Next is Handcrafted Feature Extraction and Reduction.
- The last is Fusion Feature before being classified.

The result of this research achieves 94.75% accuracy, which is the highest test score on the Kvasir Dataset by using ResNet 50.[6]

Our project is inspired by studies of Transfer Learning and using Deep Neural Network to diagnose Endoscopic images [4]. Moreover, in later years, Dense-ResNet and Ensembling Dense Network- Residual Network are used to classify image on Imagenet.[2]

3. DATASET

To do this task, we use Kvasir Dataset. The Kvasir Dataset is collected using endoscopic equipment at VestreViken Health Trust (VV) in Norway. The VV consists of 4 hospitals and provides health care to 470.000 people. One of these hospitals (the Bærum Hospital) has a large gastroenterology department from where training data have been collected and will be provided, making the dataset larger in the future. Furthermore, the images are carefully annotated by one or more medical experts from VV and the Cancer Registry of Norway (CRN).[3]

The dataset consists of 80000 images in 10 folds for cross-validation in the training and evaluating process. 80000 images are split into eight classes: dyed-lifted-polyps, dyed-resection-margins, esophagitis, normal-cecum, normal-pylorus, normal-z-line, polyps and ulcerative-colitis.[3]

We use Kvasir dataset v2 consisting of 80000 images in 10 folds for the training process.

For evaluating the method, we propose the Kvasir dataset v1 containing 4000 images in 8 classes to be the test set to evaluate the metrics.

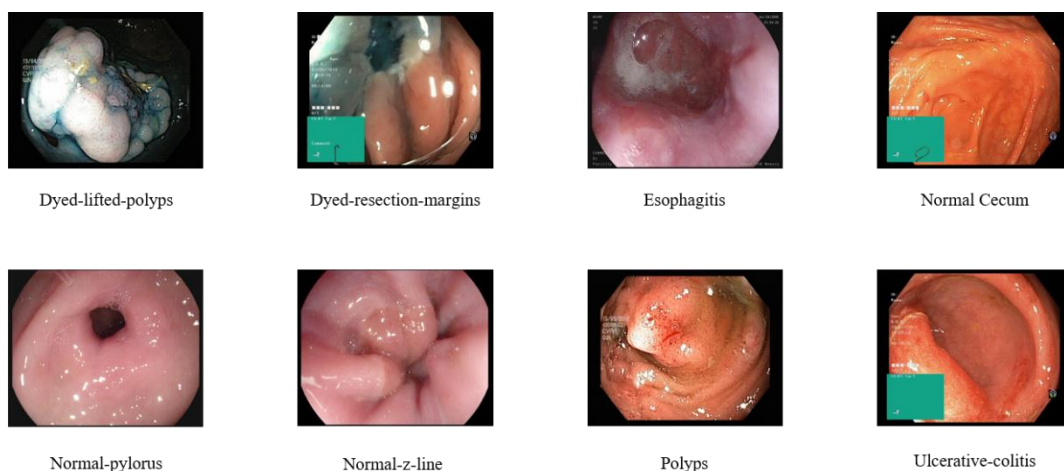


Figure 1. Eight classes in Kvasir dataset

4. METHODS

In the methods, we use 2 backbones are DenseNet 121 and ResNet 101 to build our model architecture. The dataset for training and evaluation of the architecture is from Kvasir Dataset and some parameter can be changed to get the better result.

4.1. Data Pre-Processing

After loading data, we resize all the images to the size (256,256), then we split the dataset into the training set and validation set in the ratio of 0.75:0.25. After resizing and splitting the validation set, we rescale the data pixel down to be in the range [-1,1] by divide by 127.5. Then we use the application of ResNet to preprocess input.

4.2. Data Augmentation

To reduce the Overfitting problem, we use augmentation to generate the data randomly by random flip images and random rotation with an index of 0.2.

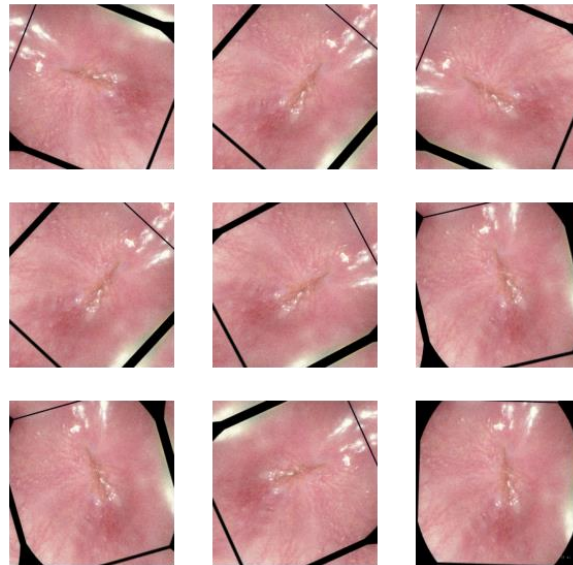


Figure 2. Data after Augmentation

4.3. Network Architecture

In our architecture, we propose to use ResNet 101 and DenseNet 121 backbones for the first layers. We will have two pipes: ResNet 101 and DenseNet. The output of ResNet 101 will be extracted, by a Conv2D, to have the same shape as DenseNet 121 output. After feature extraction, they are added to create the feature map before coming to global Average Pooling layers for being classified.

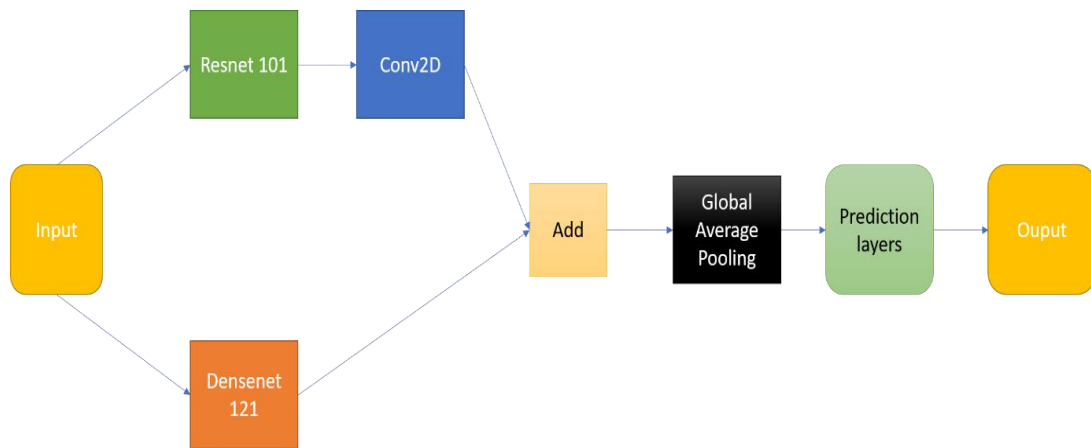


Figure 3. Visualize Dense-Res Net architecture

The figure below will illustrate our work and model that we design:

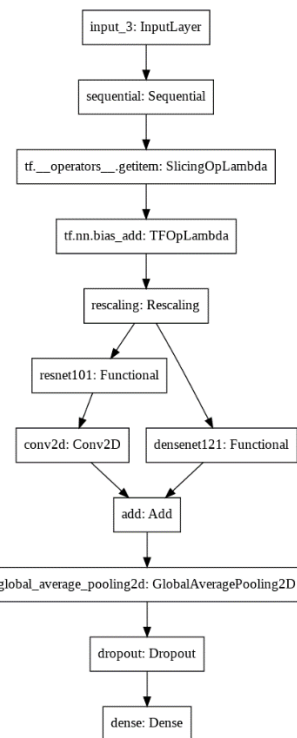


Figure 4. Visualize full model architecture include pre-processing layers

4.4. Training Model

Our models are initialized with pre-trained weight from TensorflowImagenet. We use a batch size of 32 for training data with an image's size of (256,256). We use RMSprop with a learning rate is 0.0001 for optimizer and evaluate the training process by accuracy and F1-score. For the loss function, we use Sparse Categorical Cross-entropy. We train the model with 20 epochs and get the checkpoint that has the highest validation loss.

Firstly, we freeze all the complicated layers of DenseNet and ResNet. Then, we start to train for the first time and get the result:



Figure 5. The learning curve for the first training

After the first training, we unfreeze all layers in both ResNet and DenseNet, and we define the model fine-tune from layer 100th. We have the result for the second training process:

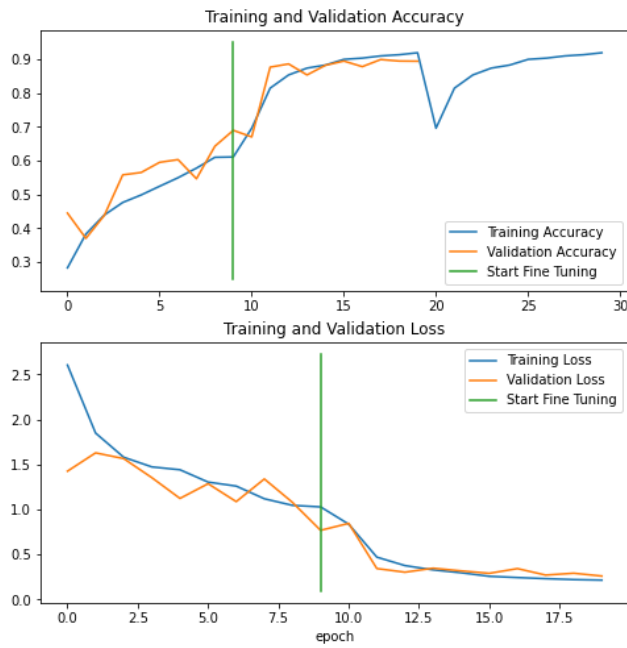


Figure 6. The learning curve for the second training

4.5. Evaluation

After training on Kvasir dataset V2, we approach testing with the data of 4000 images from Kvasir dataset v1. We have the result below:

Table 1. Evaluation of model on Kvasir dataset v1

Metrics	Value
Accuracy	0.9263
Precision	0.93375
Recall	0.92625
F1 Score	0.92625

The model has good performance with an accuracy of approximately 0.93. The other measurement scores have the same trend with that accuracy, demonstrating that this model has well-perform on this dataset.

The confusion matrix below evaluates the performance of each class:

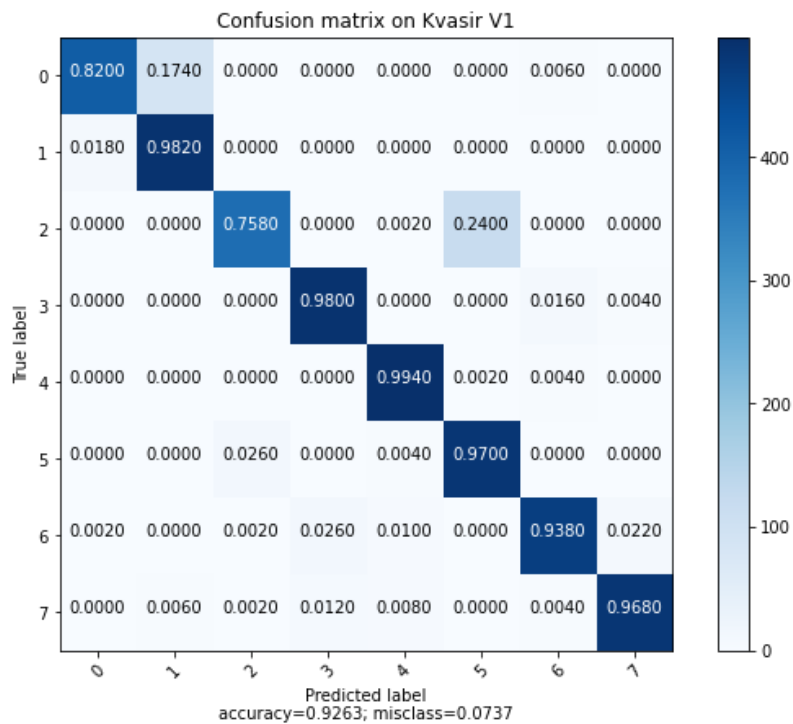


Figure 7. Confusion matrix evaluate performance of model on Kvasir Dataset V1

By the Confusion matrix, we can see some false prediction on label 1 and label 5, which belong to dyed-resection-margins and normal-z-line. Dyed-resection-margins and normal-z-line have predicted to dyed-lifted-polyps and esophagitis.

To deal with these problems, we propose methods to have better pre-processing data by reducing the noise, such as a green box on the endoscopic images.

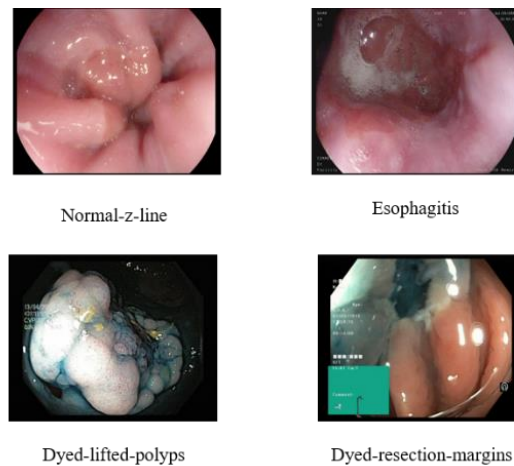


Figure 8. Comparison of False prediction

5. DISCUSSION

Although our method achieves a high score, there are some limitations in our experiment. Therefore, we need to improve. Initially, the model can probably get overfitting if we have little training data. Furthermore, it is necessary to do more experiments to choose the better parameter and the number of layers to freeze.

In the future, we can optimize the parameter and add some Batch Normalization layers to optimize computational cost and improve the score of the model.

6. CONCLUSION

We demonstrated the proposal of using Res-Dense Net with Fine-tuning technique to classify endoscopic images. The result of our research is positive, which are 0.9263 for accuracy and 0.92625 for F1 score. However, there are some drawbacks that we have to do to improve the performance of the model, such as pre-processing data, reduce noise, change the size of the image to train.

Furthermore, we can apply ResNet101 V2 or DenseNet 169 backbone to have better feature extraction and better performance of the model.

7. ACKNOWLEDGMENT

This research is partially supported by the research funding from the Faculty of Information Technology, University of Science, Ho Chi Minh city, Vietnam.

REFERENCES

- [1] Chien-Hsiang Huang, Hung-Yu Wu, and Youn-Long Lin: HarDNet-MSEG: A Simple Encoder-Decoder Polyp Segmentation Neural Network that Achieves over 0.9 Mean Dice and 86 FPS.
- [2] Victor Cheung: DenResNet: Ensembling Dense Networks and Residual Networks
- [3] Konstantin Pogorelov, Kristin Ranheim Randel, Carsten Griwodz, Sigrun Losada Eskeland, Thomas de Lange, Dag Johansen, Concetto Spampinato, Duc-Tien Dang-Nguyen, Mathias Lux, Peter Thelin Schmidt, Michael Riegler, Pål Halvorsen, Kvasir: A Multi-Class Image Dataset for Computer Aided

- Gastrointestinal Disease Detection, In MMSys'17 Proceedings of the 8th ACM on Multimedia Systems Conference (MMSYS), Pages 164-169 Taipei, Taiwan, June 20-23, 2017.
- [4] Trung-Hieu Hoang, Hai-Dang Nguyen, Viet-Anh Nguyen, Thanh-An Nguyen, Vinh-Tiep Nguyen, Minh-Triet Tran: Enhancing Endoscopic Image Classification with Symptom Localization and Data Augmentation, in Proceedings of the 27th ACM International Conference on Multimedia.
- [5] Nini Rao, Hongxiu Jiang, Chengsi Luo: Review on the Applications of Deep Learning in the Analysis of Gastrointestinal Endoscopy Images., Article in IEEE Access - September 2019.
- [6] Omneya Attalah and Maha Sharkas: GASTRO-CADx: a three stages framework for diagnosing gastrointestinal diseases, PeerJ Computer Science.
- [7] Konstantin Pogorelov, Kristin Ranheim Randel, Carsten Gri-wodz, Sigrun Losada Eskeland, Thomas de Lange, Dag Johansen, Concetto Spampinato, Duc-Tien Dang-Nguyen, Mathias Lux, Peter Thelin Schmidt, Michael Riegler, and Pål Halvorsen. Kvasir: A multi-class image dataset for computer aided gastrointestinal disease detection. In Proceedings of the 8th ACM on Multimedia Systems Conference, MMSys'17, page 164–169, New York, NY, USA, 2017. Association for Computing Machinery.
- [8] Dinh Viet Sang, Tran Quang Chung, Phan Ngoc Lan, Dao Viet Hang, Dao Van Long, Nguyen Thi Thuy: AG-CUResNeSt: A Novel Method for Colon Polyp Segmentation, Artificial Intelligence in Medicine, May 6 2021.

AUTHORS

Quoc-Huy Trinh and Minh-Van Nguyen Study Bachelor of Computer Science at Ho Chi Minh University of Science Ho Chi Minh city, Vietnam

