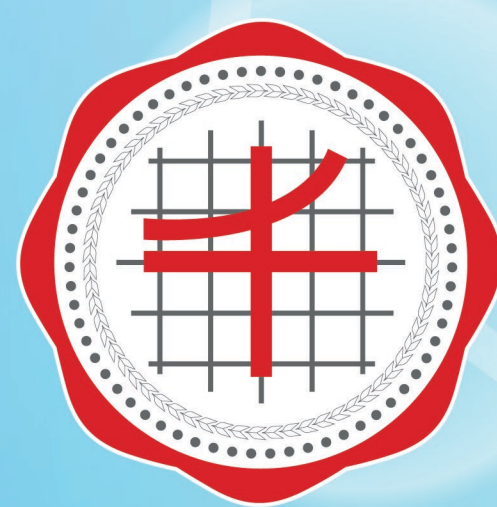




Digestive organ classification from wireless capsule endoscopy images with convolutional neural network

Supaksagee Plukaram, Sippapart Subsanong, Supawee Soythongcharoen, Amporn Atsawarungrangkit, Kawee Numpacharoen and Nuwee Wiwatwattana

Bachelor of Science Program in Computer Science

Abstract

Lesion location is a critical information that must be reported by gastroenterologists when interpreting capsule endoscopy images. There are few studies on how deep learning models can be applied to classify the locations of gastrointestinal tract. Therefore, we aim to create a deep learning model for classifying the locations of gastrointestinal tract (i.e., esophagus, stomach, small bowel, and colon) based on capsule endoscopy images. Dataset of capsule endoscopy images (n = 90,171 images from 123 patients) was divided into training (n = 40,000) and testing (n = 50,171). The images were labeled into 4 organ locations (esophagus, stomach, small bowel, and colon). We applied the deep learning architecture (InceptionResnet V2) to the training dataset. Then, the performance of the trained model was externally validated using the testing dataset. The confusion matrix of the deep learning model, visualizing the performance of our algorithm. The model classified capsule endoscopy images into 4 classes with accuracy 93.22%, precision 97.78%, recall (sensitivity) 93.22%, and F1 95.25%. The average prediction time was 20 milliseconds per image. The deep learning model demonstrated an excellent classification performance, which can be used as a building block for the ultimate goal of creating a fully automated model for interpreting capsule endoscopy images.

Methods and Experiments

1.Data Preparation

A wireless capsule endoscopy dataset of 90,171 images from 123 patients from Brown University and Rhode Island Hospital was divided into 40,000 training images and 50,171 testing images. Images were firstly resized into 256x256 pixels, and were manually selected for each of the organ classes (Esophagus, Stomach, Small Bowel, and Colon). This is to improve the efficiency of the model later on and avoid abnormal images.

2.Data Modeling

In this project, we used several deep learning neural network models such as InceptionV3, Resnet50, and InceptionResnetV2. However, InceptionResnetV2 was reported here as it outperforms others. The optimizer used is SGD. The learning rate is set to 0.01. The momentum is set to 0.9.

Related Knowledge

Capsule Endoscopy [1]

Capsule endoscopy is a procedure that uses a tiny wireless camera to take pictures of digestive tract. A capsule endoscopy camera sits inside a vitamin-size capsule. As the capsule travels through your digestive tract, the camera takes thousands of pictures that help doctors see inside the small intestine — an area that isn't easily reached with more-traditional endoscopy procedures.

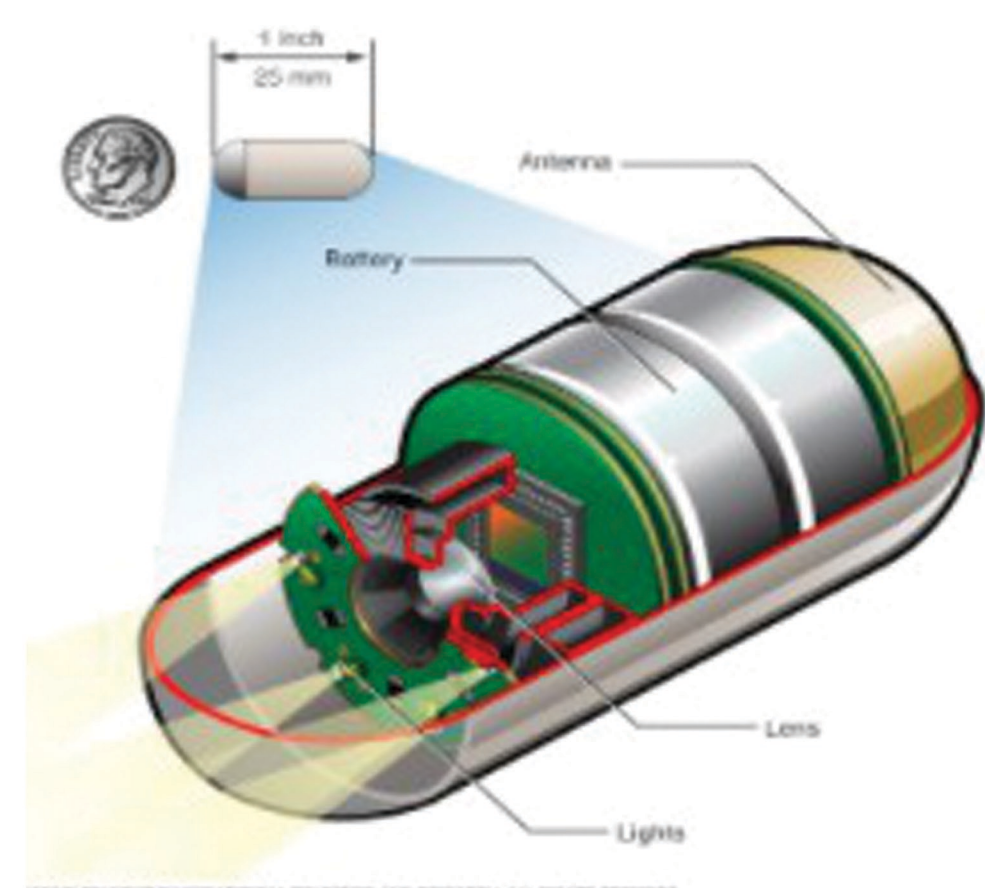


Figure 1 : Capsule Endoscopy

Convolutional Neural Network (CNN) [2]

Convolutional Neural Network (CNN) is a deep learning algorithm that is commonly used for computer vision applications.. CNN is a multi-layered neural network that analyzes image inputs and performs tasks such as imagerecognition, image classification, and detection.

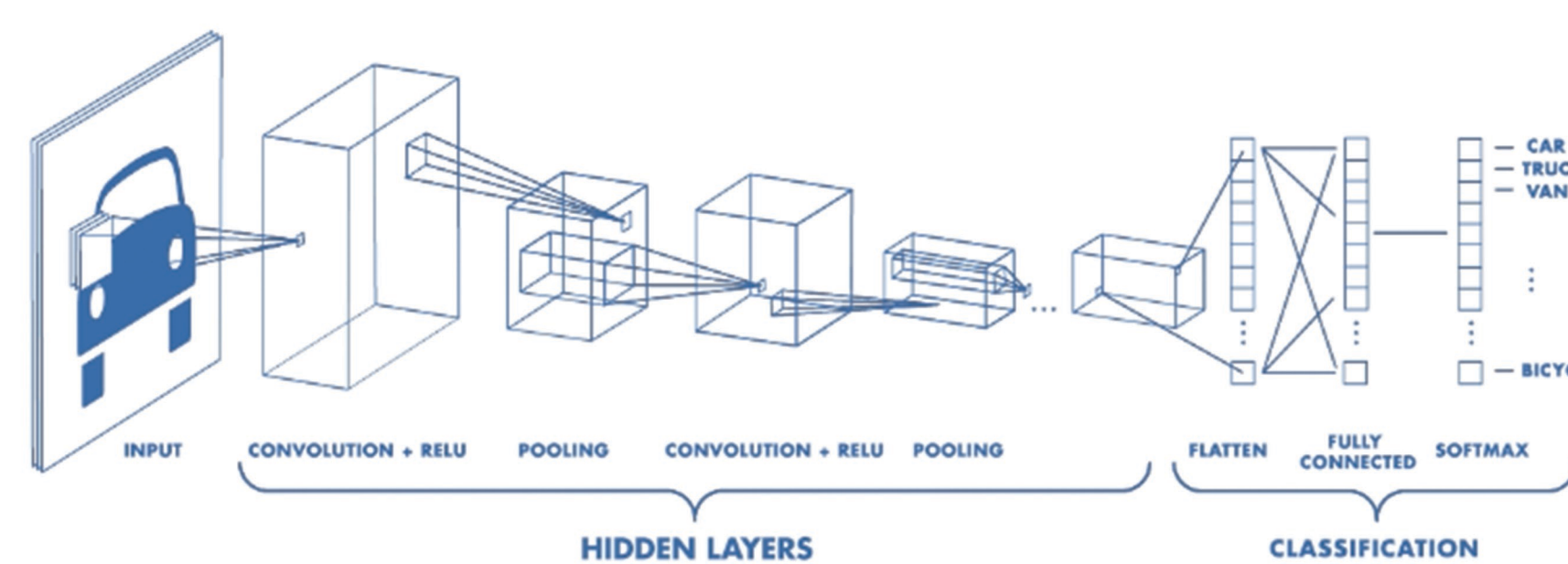


Figure 2 : Convolutional Neural Network (CNN)

InceptionResnetV2 [3]

Inception-ResNet-v2 is a convolutional neural network that is trained on more than a million images from the ImageNet database. The network has 164 layers deep and can classify images into 1000 object categories, such as the keyboard, mouse, pencil, and many animals. As a result, the network has learned rich feature representations for a wide range of images. The network has an image input size of 256-by-256, and the output is a list of estimated class probabilities.

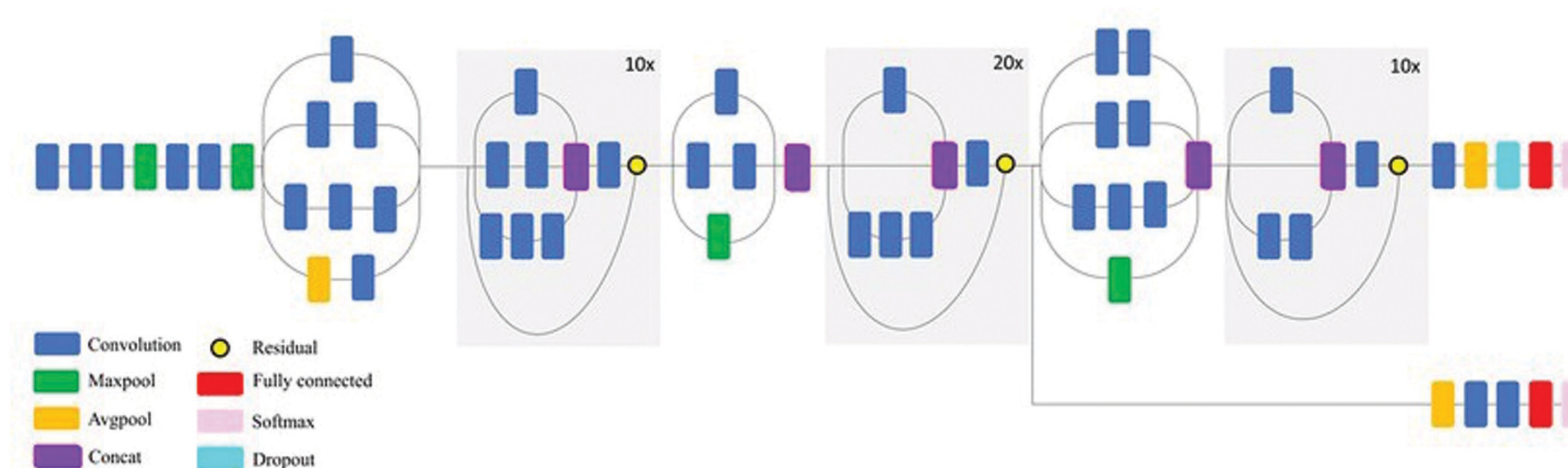


Figure 3 : InceptionResnetV2 Architecture

Results and Discussion

Example of predicting a Esophagus, Stomach, Small Bowel and Colon image along with class probabilities are shown in Fig. 4 with highest probabilities. The confusion matrix of InceptionResnetV2 visualizing the performance of our algorithm is shown below. The model classified capsule endoscopy images into 4 classes with accuracy 93.22%, precision 97.78%, recall (sensitivity) 93.22%, specificity = 97.98% and F1 95.25%. The average prediction time was 20 milliseconds per image. The deep learning model demonstrated an excellent classification performance, which can be used as a building block for the ultimate goal of creating a fully automated model for interpreting capsule endoscopy images.

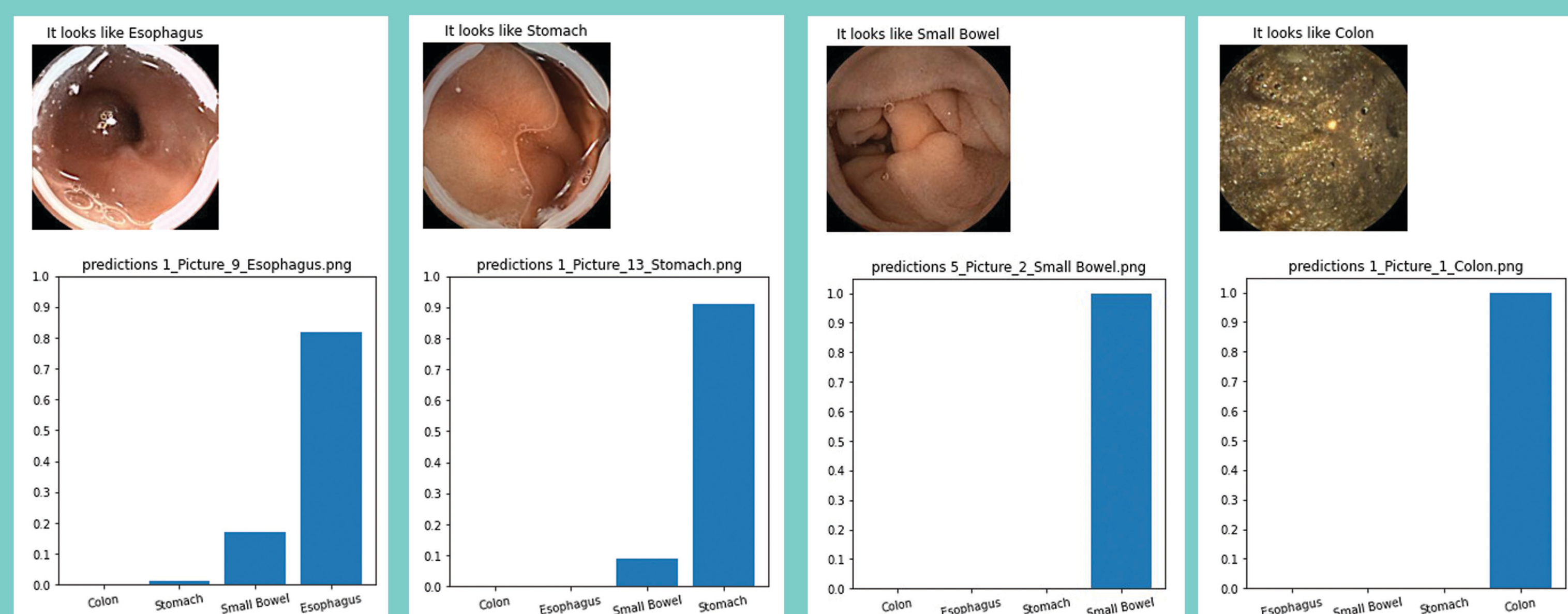


Figure 4 : Examples of Predicted Results

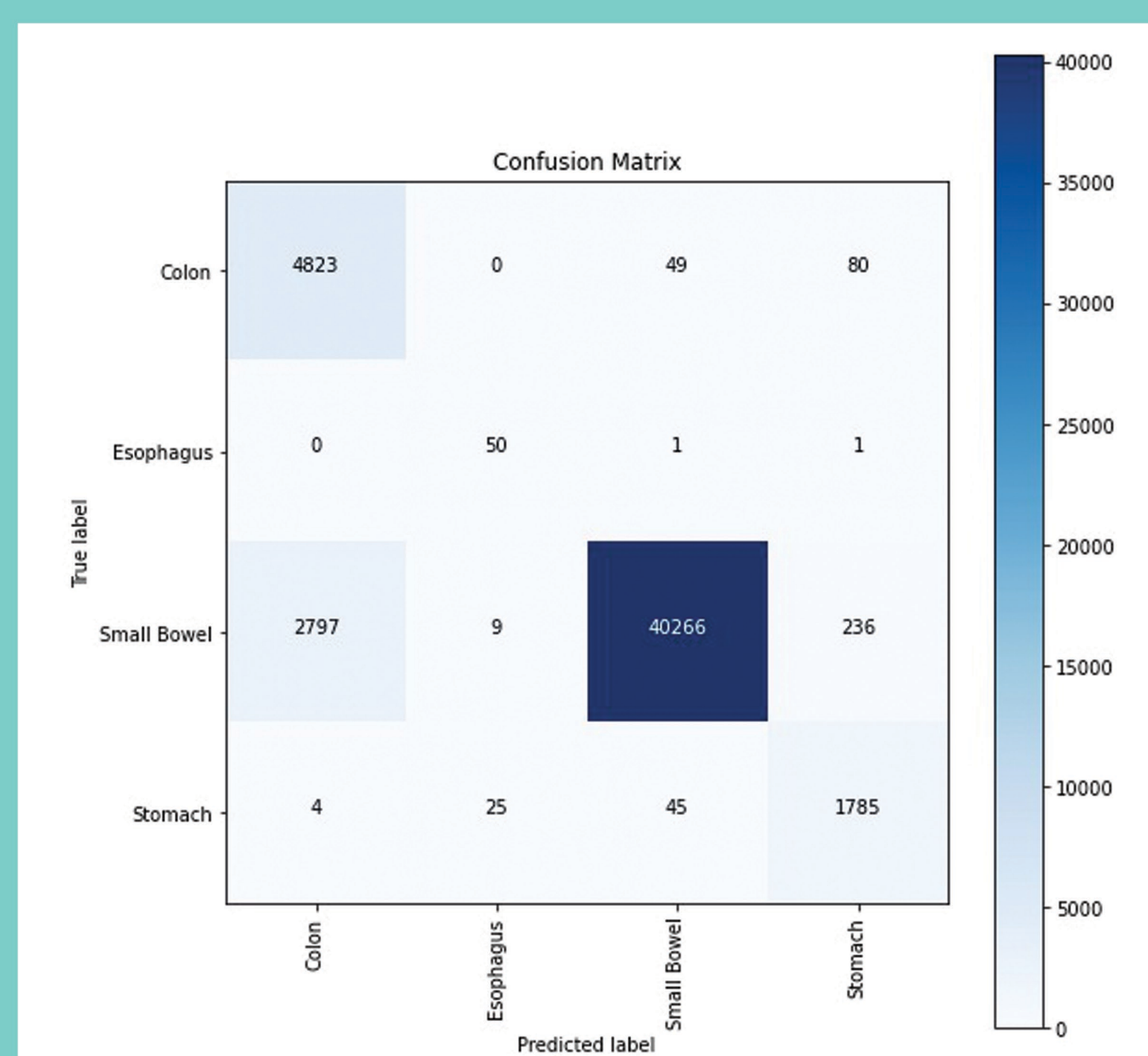


Figure 5 : Confusion Matrix

Bibliography

[1] Capsule endoscopy. Retrieved June 7, 2020, from <https://www.mayoclinic.org/tests-procedures/capsule-endoscopy/about/pac-20393366>

[2] Convolutional NeuralNetworkArchitecture: Forging Pathways to the Future. Retrieved June 17, 2020, from <https://missinglink.ai/guides/convolutional-neural-networks/convolutional-neural-network-architecture-forging-pathways-future/?fbclid=IwAR2fTx7l7RDmXgVwyxqF4QaCAlucvIN3Mw7OopnRt4jBS5Rv6UN6RokjYOY>

[3] Elhamraoui,Z.(2020,May16).I n c e p t i o n ResNetV2 Simple Introduction.Retrieved March 1,2021.from <https://medium.com/@zahraelhamraoui1997/inceptionresnetv2-simple-introdcion-9a2000edcdeb>