



EEEE4008

Final Year Individual Project Proposal

**Using Generative Adversarial Networks to Map
Gastrointestinal Diseases during Endoscopy**

AUTHOR: Ms Catarina Gomes da Costa

ID NUMBER: 20313190

SUPERVISOR: Dr George Gordon

MODERATOR: Dr. Kevin Webb

DATE: 3rd November 2022

1 Introduction

The project proposed in this document aims to detect diseases in the gastrointestinal system by using artificial intelligence (AI), specifically generative adversarial networks (GANs), on a labelled dataset of images collected during endoscopies. In the following sections, a review of the project's background and industrial relevance will be done to contextualise the project. The remaining sections will focus on the project's intrinsic details, such as aims, milestones, risks and time plan.

2 Background Review

2.1 Project Context

Capsule endoscopy is a recent method developed to perform endoscopies without causing patients discomfort. In capsule endoscopy, the patient takes a capsule with a camera and LEDs inside to photograph the gastrointestinal tract. Whilst the capsule travels through the tract, it sends the images collected to a belt the patient wears, which contains a data recorder. Thus, it stores all the images the capsule collects and receives the data wirelessly [1]. Figure 1 illustrates this process, with a patient wearing the data recorder while taking the capsule, the capsule's width and the capsule travelling through the gastrointestinal tract.

The capsule can take 6 to 12 hours to finish its journey [1]. Should a patient not be sure if the capsule is no longer in their body after a bowel movement, the doctor can request a CT scan to clear any doubts and tell the patient what to do next. Once the doctor gets the data recorder back, the processing of the data collected can commence. This process is highly time-consuming for doctors, as thousands of images are collected in each endoscopy when sometimes, only a few are concerning.

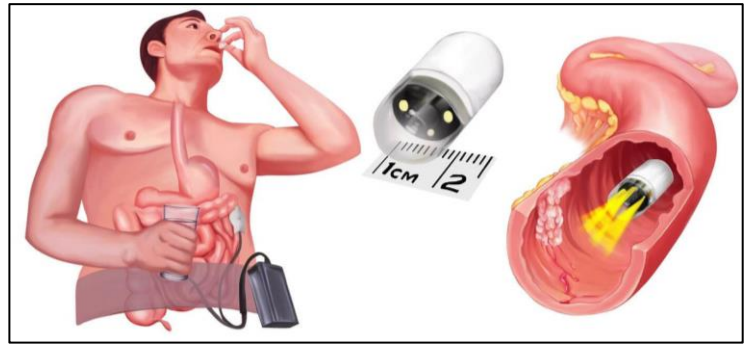


Figure 1 Capsule Endoscopy Overview (includes data recorder, capsule size reference and capsule traveling through the gastrointestinal tract, respectively) [1]

Thus, this project aims to create a model to aid doctors in processing the data collected to diagnose patients in much less time and accurately. Moreover, due to the nature of the procedure, it is possible to develop tools to help doctors process so much data. In the case of traditional endoscopy, such would not be possible/ relevant because doctors process the data in real time. Therefore, in this project, the model to be created will map gastrointestinal diseases from raw endoscopic images (input) to then output a replica of each image but with different colours to highlight the concerning tissue in the images. Such model uses generative adversarial networks (GANs), a form of deep learning within AI, to map these diseases. The model will be trained on a specific list of pathologies present in the KID dataset [2].

2.2 GANs Overview [3]

GANs are a relatively recent framework with a particular architecture comprised of two neural networks. They are capable of generating artificial data inspired by their input data. However, their ability to produce hyper-realistic data comes from the game-like interaction between the two neural networks as they compete with each other. These neural networks are called generator and discriminator. The generator generates artificial data (new instances) from a noise input, while the discriminator is responsible for classifying each instance as real or fake. To do so, the discriminator has two inputs, the new instances created by the generator and real images collected beforehand.

As the discriminator labels each instance, it penalises the generator every time it detects a fake instance. This way, the generator learns how to produce realistic data.

Figure 2 shows the complete GAN architecture. As can be seen, two loss functions are coming from the discriminator. These are vital to training the GAN, and due to its unique architecture, so is their training process. As there are two neural networks, the GAN is trained by training the discriminator and generator independently, for one or more epochs each, and repeating the process as desired.

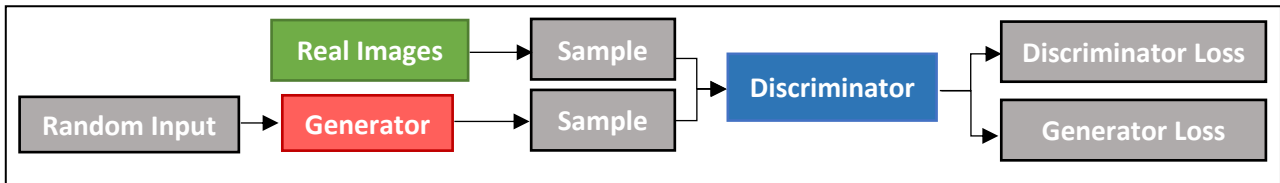


Figure 2 GAN Complete Architecture [3]

Despite their potential, GANs still have limitations, such as:

- Failure to converge- mitigated by changing the discriminator's penalising weights and adding noise
- Generator Failure- caused by the discriminator being too good for the generator, causing vanishing gradients. Mitigated by using Wasserstein loss which prevents vanishing gradients
- Mode Collapse- occurs when the generator does not produce varied instances. The Wasserstein loss can also be used to alleviate this.

Nevertheless, overall, GANs are a tool with great potential, and they have been proving so by improving the quality of the artificial data generated in just a few years. Besides their use for this project, GANs can also be used for data augmentation, denoising, text-to-image translation and computer vision, to name a few [3].

For this specific task, the GAN architecture will be slightly different. Instead of a generator, it will have a variational autoencoder (VAE), with the KID dataset as input. The VAE will be responsible for creating the final mapped image. This image will highlight the concerning tissues in a binary system by classifying each pixel as disease or not disease. Furthermore, if a more complex system should later be created, just using a VAE in a GAN will not suffice as it is only capable of performing binary mapping. Thus, an additional object detection algorithm might be required to create a map with multiple classes, i.e., Alex. However, using an additional neural network may not be necessary GANs can be used for object detection [4].

On the other hand, different GANs have varying (sometimes low); thus, before choosing the object detection algorithm, possible architectures and current literature must be considered. Finally, the discriminator will be responsible for ensuring a mapped image and a raw image from the dataset correspond to a correct pair to perform the mapping correctly.

2.3 GANs' Current Research for Medical Imaging

Due to their potential and surprising results, GANs have been a popular area of research in recent years. In the medical field, specifically medical imaging, GANs can be used in different ways. GANs are used to generate high-resolution medical data (a form of data augmentation), such as MRI scans. Even though it is a common way to collect data, slow acquisition is known for affecting data quality; hence, GANs are also used in data reconstruction and denoising [5].

Their most successful application is reliable data synthesis [5] which can help AI models overcome underfitting due to insufficient data. This was a recurring limitation when applying AI in the medical field that GANs are helping to fix. In addition, there is an increase in funding to collect greater datasets from thousands of patients (i.e. UK BioBank), which can also help GANs to synthesise a broader range of possible pathologies.

Furthermore, there is significant research done in domain transformation [5], which refers to the transformation of an image type to another, i.e. MRI to CT scan (the most common medical images in transformation research) [5]. This capability is beneficial for both patients and clinicians due to the characteristic of each data collection type, i.e. MRI requiring patients to not move for long periods and not have any metallic object. In 2020, researchers proposed a GAN function expansion, where adversarial learning helps build a generalisable classifier to help predict retinal diseases [5].

Moreover, a critical technique in medical imaging is segmentation, and a wide range of methods exist to achieve it. Segmentation refers to the setting of boundaries for different objects in an image. When applied in medical images, the non-linearity of tissue makes the segmentation precision a challenge. However, GANs' synthesising ability can help in improving segmentation accuracy [6]. On the other hand, researchers also propose using conditional GANs' generators to segment images by taking advantage of their U-net shape. Such method has been applied in tumour segmentation and spinal shape knowledge [5].

3 Industrial Relevance

A national census conducted in 2019 showed that gastrointestinal (GI) endoscopies have increased by 12-15% compared to 2017 (\approx 2.1 million endoscopies in 2019, many for bowel cancer) and that fewer services are meeting target wait times [7]. Such census was conducted before the COVID-19 pandemic, and no further census has updated these numbers. However, it is safe to assume that the problem has worsened based on the stress the NHS was under for those two years. The total cost of an endoscopy depends on the device used to conduct the procedure. There are two main ways of conducting a GI endoscopy: using a capsule with a camera that the patient ingests and the traditional flexible cable.

According to the National Schedule of NHS costs from 2020/21, the price difference between conventional and capsule does not vary significantly, with cost per procedure being £1,192 and £949, respectively [8]. Although single-procedure cost savings isn't significant when opting for capsule endoscopy, the comfort it provides to patients makes it more appealing, increasing the number of procedures performed. Such has led to diagnosing and treating more pathologies in their early stages. Thus reducing the number of surgeries costs by approximately £16,000 [7], [8]. The manufacturer, Medtronic of PillCam, is known for the screening program with NHS England [9]. Moreover, a few other manufacturers stand out in the capsule endoscopy industry, where the price per capsule does not vary significantly from \$500 (USD) [10]. These are Olympus America and Jinshan [10]; thus a significant part of the cost of capsule endoscopy is the capsule itself.

Using capsule endoscopy, doctors have to go over thousands of pictures per patient, making the diagnosis incredibly time-consuming. Thus, with this project offering a more time and cost-effective procedure where more patients can be tested without any discomfort, this problem can be mitigated. Such means the AI can be used to detect other worrying and/or life-threatening pathologies, such as bowel cancer in its early stages, saving lives and additional treatment costs, already done in 2021. There is an average of \approx 43,000 new cases each year, and 54% are preventable [11]. Overall, bowel cancer cost the UK £1.74 billion in 2018 alone [12], and with the increase in screening, more doctors suspect more patients have the disease. Thus, deploying a product to help doctors make a more rapid and accurate diagnosis on a mass scale can help the UK save lives and a substantial amount of capital through disease prevention.

4 Project Aims and objectives

The project **aims** to create an AI system that takes raw endoscopic images collected during capsule endoscopy as an input and outputs the maps of their gastrointestinal diseases. As GANs have a complex architecture, the AI implementation should start with simpler networks and progress to more advanced ones until a GAN has been implemented, all using TensorFlow. Therefore, to be precise, the project's objectives are:

1. Implement and test an Image Classifier using convolutional neural networks on an example dataset, i.e. Fashion MNIST dataset
2. Implement and test an Image Classifier using convolutional neural networks on the labelled KID dataset to classify the different diseases present in it
3. Implement and test a Variational Auto Encoder (VAE) using an example dataset, i.e. MNIST dataset. (Note: VAE is one of the neural networks that constitute the GAN to be implemented)
4. Implement and test the Pix2Pix GAN using an example dataset, i.e. CelebA dataset
5. Implement and test the Pix2Pix GAN on the labelled KID dataset to map gastrointestinal diseases
6. **Additional:** Combine classification with disease mapping to create colour-coded maps that show different diseases
7. **Additional:** Run the final trained model on a different, unlabelled data set on the Pix2Pix GAN to compare performance with the KID data set and discuss results

5 Deliverables, Milestones and Risks

From a technical point of view, the project has 4 milestones (one of them additional) which correspond to the successful implementation of each neural network:

1. Complete classifier for KID dataset (with the best performance achieved)
2. Complete an example of a Variational Auto Encoder (with the best performance achieved)
3. Complete Pix2Pix GAN for the KID dataset (with the best performance achieved)
4. **Additional:** Merging of classifier and Pix2Pix GAN into one program (with the best performance achieved)

The previous section presented the project's objectives with SMART deliverables as a reference. It is essential to highlight that the order in which the selected neural networks are implemented helps ensure the time given to complete the project is sufficient, contributing to its success.

Thus, the project's deliverables include the following:

1. Literature Review on GANs and Gastrointestinal endoscopy
2. Complete code capable of classifying diseases in the KID dataset
3. Complete code capable of mapping gastrointestinal diseases
4. Final year project report: including analysis and results of neural networks used

5.1 Risks and Risk Mitigation:

Table 1 shows the project's risks and how they are expected to be mitigated in greater detail.

RISK	MITIGATION
TRAINING TIMES (ESPECIALLY DURING TESTING)	The training of any neural network is highly time-consuming. Thus, the first way to mitigate this issue is by following the time plan to know when to stop spending more time on a task and not compromise the project's success. Moreover, using the project's budget to purchase a Colab subscription to access cloud computing (currently waiting for approval) is the ideal way of reducing training times.
DATASET TOO SMALL	The KID dataset has a couple of thousands of labelled images, which for an AI model, might be too little. This is because the dataset must be split into test and train groups (20:80, respectively), and it may not be enough data to train the model appropriately. In other words, insufficient training data will lead to the model having a poor approximation (overfit). For this reason, it is essential to be aware of datasets with similar images, such as Kvasir (n=6,000) and Cad-cap (n=25,000), to overcome this issue.
OVERFITTING	Overfitting occurs when the model "memorises" the training data instead of "learning" it. Such means that once the model is tested, its accuracy during testing is significantly lower than in training. Thus, to overcome this, it is imperative to use cross-validation to prevent overfitting. Overfitting is due to the dataset the model was trained on being too small (refer to the previous risk to overcome this issue).
UNKNOWN OPTIMAL ACCURACY	Regardless of the format, any raw data has a certain amount of noise which in this case is unknown. The presence of noise affects the model's accuracy making it difficult to <i>a priori</i> make any realistic predictions. Moreover, similar datasets have been used as input to AI models. Thus, it could be to use their final accuracy as a guideline.

Table 1 Project's Risks and Their Mitigation

6 Time plan

Figure 3 shows the project's proposed Gantt Chart with two timelines. The first timeline (in blue) includes all core and additional objectives, and the second timeline (in grey with red borders) only includes the core objectives as a contingency plan, meaning, should there be any delays and the ideal plan not possible to follow. Therefore, the Gantt chart includes two types of milestones, one for each timeline, green and red, respectively. In the ideal plan (blue), approximately three weeks were given to complete each task, while four or more weeks were given in the contingency plan.

The contingency plan divides the workload between the two semesters by achieving the first milestone in the autumn semester whilst conducting a literature review (while neural networks are retrained as training times are long). The second milestone should be achieved shortly after the exam period in January, as work is expected to continue during Christmas break. The final milestone will be achieved close to the start of Easter break, still leaving enough time to complete the thesis write-up. The ideal plan divides the workload between the two semesters by achieving two milestones in each semester.

Do note that both plans include a buffer at the end of March to ensure that any fundamental tasks in each plan are completed. Three weeks have been reserved for writing the thesis exclusively, with extra days to proofread and submit. In both cases, the literature review is written during Christmas break. Lastly, note that objectives 3 and 4 can be done simultaneously as variational autoencoders are one of the neural networks that constitute the Pix2Pix GAN. However, such is more likely to happen should the contingency plan have to be used, as it is an additional strategy that can be used to ensure the core objectives are completed in due time.

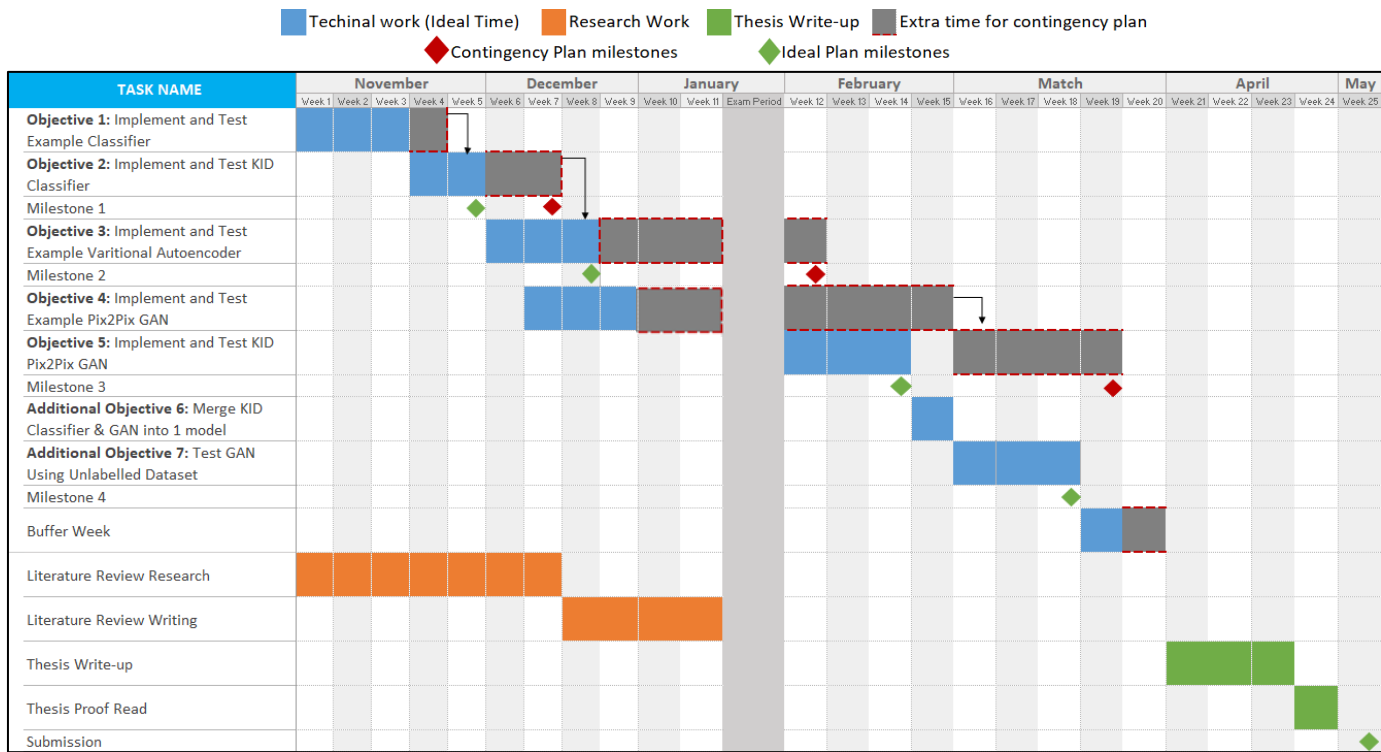


Figure 3 Project's Proposed Gantt Chart Including Ideal and Contingency Plans

7 References

- [1] D. V. Jayasekaran, "Video capsule endoscopy," [Online]. Available: <https://www.gutworks.com.au/video-capsule-endoscopy-murdoch-perth/>. [Accessed 23 10 2022].
- [2] MDSS Research Group, "A Capsule Endoscopy Database for Medical Decision Support," MDSS Research Group, [Online]. Available: <https://mdss.uth.gr/datasets/endoscopy/kid/>. [Accessed 23 10 2022].
- [3] Google, "Introduction," Google, 18 7 2022. [Online]. Available: <https://developers.google.com/machine-learning/gan>. [Accessed 23 10 2022].
- [4] L. J. K. Charan D. Prakash, "It GAN DO Better: GAN-based Detection of Objects on Images with Varying Quality," *IEEE Transactions on Image Processing*, vol. 30, pp. 9220-9230, 2021.
- [5] X. L. et. al., "When medical images meet generative adversarial network: recent development and research opportunities," *Discover Artificial Intelligence*, vol. 1, no. 5, 2021.
- [6] S. X. e. al., "Generative adversarial networks in medical image segmentation: A review," *Computers in Biology and Medicine*, vol. 140, 2022.
- [7] S. R. et. al., "National census of UK endoscopy services in 2019," *BMJ*, vol. 12, no. 6, 2019.
- [8] NHS England, "National Cost Collection for the NHS," 2021. [Online]. Available: <https://www.england.nhs.uk/costing-in-the-nhs/national-cost-collection/>. [Accessed 25 10 2022].
- [9] Medtronic, "Frequently asked questions PillCam remote reading platform," [Online]. Available: <https://www.medtronic.com/content/dam/covidien/library/emea/en/product/capsule-endoscopy/weu-pillcam-remote-reading-platform-faq-for-hcps.pdf>. [Accessed 22 10 2022].
- [10] R. Martin, "Can Capsule Endoscopy Technology Keep You Out of Surgery?," 30 11 2018. [Online]. Available: <https://igniteoutsourcing.com/healthcare/capsule-endoscopy-technology/>. [Accessed 21 10 2022].
- [11] Cancer Research UK, "Bowel Cancer Statistics," [Online]. Available: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer#heading-Five>. [Accessed 22 10 2022].
- [12] Bowel Cancer UK, "Bowel cancer costs the UK £1.74 billion a year," Bowel Cancer UK, 15 10 2020. [Online]. Available: <https://www.bowelcanceruk.org.uk/news-and-blogs/news/bowel-cancer-costs-the-uk-%C2%A31.74-billion-a-year/>. [Accessed 22 10 2022].

Using Generative Adversarial Networks to Map Gastrointestinal Diseases during Endoscopy

George Gordon

Area: signal processing, machine learning/AI

Nature of project: mathematics, coding, data analysis, image processing

Relevant modules in Year3/4: EEEE4119, EEEE4127

Skills to be gained during project: imaging, artificial intelligence, machine learning, neural networks, python/TensorFlow

This project is well-suited for those confident with coding and mathematics.

Project description: Capsule endoscopy is a relatively new, low-cost procedure that can be used to diagnose diseases of the gastrointestinal tract, particularly parts that cannot be easily reached by conventional endoscopes (e.g. distal small bowel). Each year over 12,000 capsule endoscopies are conducted in the UK [5] and the procedure consists of swallowing a video-recording capsule. Currently, a doctor has to manually sift through over 20,000 recorded frames to find often only 1 that contains a concerning pathology. This leads to significant delays in diagnosis and is prone to human error.

Recently, significant advances have been made in the field of AI image processing due to increased computational power and availability of large datasets. This has had a particularly large impact in healthcare, where AI has been applied for patient treatment, hospital management, disease detection and surgery. This project aims to detect diseases in a labelled dataset of images taken during capsule endoscopy using AI, the KID dataset [6]. Specifically, the student will work with neural networks capable of processing large image datasets, with increasingly complex architectures to reach the final goal. The relevant networks are:

- Simple classifier of the diseases present in the dataset provided
- Variational Autoencoder
- Pix2Pix GAN- Generative Adversarial Network

Objectives:

1. Understand the concept of convolutional neural networks and how they can be used to create classifiers for images
2. Understand the concept of variational auto encoder by implementing an example system in Python / TensorFlow
3. Implement the Pix2pix GAN on a simple example to understand the workings and implementation details of GANs
4. Implement the Pix2Pix GAN on the labelled KID dataset to produce maps of disease structures
5. **Additional:** Combine classification with disease mapping to create colour-coded maps show different diseases
6. **Additional:** Run the trained model on a different, unlabelled data set [7] on the Pix2Pix neural net and compare performance with the original (labelled) data set

Key references

1. Intuitive explanation of convolutional neural networks: <https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/> and <https://www.youtube.com/watch?v=aircAruvnKk>
2. Tutorials for building neural networks in python: <https://machinelearningmastery.com/tutorial-first-neural-network-python-keras/> and <https://www.pyimagesearch.com/2016/09/26/a-simple-neural-network-with-python-and-keras/>
3. Classic textbook on neural networks (if you like maths): <https://www.amazon.co.uk/Networks-Recognition-Advanced-Econometrics-Paperback/dp/0198538642>
4. <https://www.mayoclinic.org/tests-procedures/capsule-endoscopy/about/pac-20393366>
5. Ravindran S, Bassett P, Shaw T, et al National census of UK endoscopy services in 2019 Frontline Gastroenterology 2021;12:451-460.
6. <https://mdss.uth.gr/datasets/endoscopy/kid/>
7. Murra-Saca, D., 2022. The Gastrointestinal Atlas - gastrointestinalatlas.com. [online] Gastrointestinalatlas.com. Available at: <https://www.gastrointestinalatlas.com/english/english.html>

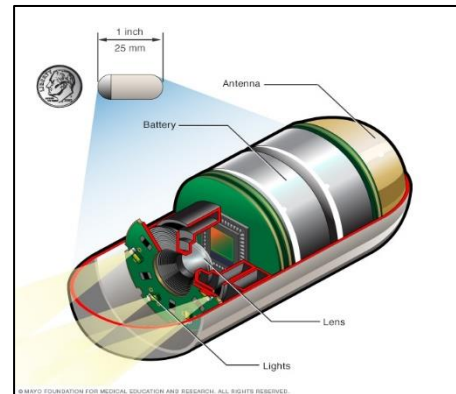


Figure 1. Example of Capsule Endoscopy [4]