

Automated detection of ulcers and erosions in capsule endoscopy images using a convolutional neural network

Dr.Sivakumar .P , Professor & Head, Dept. of ECE, GRT institute of Engineering and Technology, Tiruttani. Email: sivathellore@yahoo.com

Abstract

The wireless capsule endoscope, often known as WCE, has been successfully tested and validated in clinical settings for so many years. This new study approach, which is based on artificial intelligence technology and has excellent levels of accuracy and resilience, has the potential to minimise detection costs and help the general population. Furthermore, automatic detection techniques that look for signs of gastrointestinal sickness in WCE pictures have been accepted by these professionals as an ideal enhancing tool. Within the scope of this study, a novel computer-aided diagnostic approach for ulcer identification in WCE pictures is put up as a potential solution. The method was evaluated by an independent party with 256 591 new pictures. When implemented to the validation data, the learning algorithm had a diagnosis accuracy of 98.067 percent. According to the results of our correlations of F1, F2, and ROC-AUC, the proposed approach performs good than a number of off-the-shelf CNN models, such as VGG, Inception-ResNet-v2 and DenseNet, as well as traditional machine learning methods for WCE image classification that involve handcrafted features. Overall, the results of this research show that it is possible to recognise ulcers in WCE photos by using the deep CNN approach. In addition, the HANet architecture that we have developed specifically for this issue provides an excellent option for the construction of the network structure. The CNN networks that were employed in this study were pre-trained using millions of annotated natural photos in order to compensate for the small number of images of ulcers that were included in the accessible data sets (ImageNet). Following the extraction of the deep features, a randomized forests classifier is used to determine whether or not WCE pictures include ulcers. The outcomes are quite encouraging (96.73 percent and 95.34 percent in terms of precisions and recall accordingly).

Keywords

Detection network, gastric ulcer, wireless capsule endoscope, convolutional neural network, complications.

1 Introduction

One of the most frequent abnormalities that may be seen in the (GI) gastrointestinal tract is called an ulcer. About ten percent of the world's population is affected with ulcers at any one time (Kaplan, 2017). Erosion of the mucosal lining is what they are called, and they may happen anywhere throughout the GI system. The duodenum, the abdomen, the oesophagus, as well as the jejunum are the organs in which they are found the most often.

Ulcers are a risk factor for developing more severe conditions such as Chron's disease and peptic ulcers (Yuan et al., 2015). Endoscopy is the medical procedure that is used to diagnose ulcers in the gastrointestinal system. However, due to their low imaging resolutions., this is because it is easy for these methods to impact or even misdirect the outcomes of a diagnosis made by a medical professional. WCE only tries to make up for such deficiencies, which has several benefits like being risk-free, painless, and non-invasive, amongst others. -Y. Yeh (2014). In the meantime, it is able to delve deep into sections of the small bowel that are difficult to get to by the standard methods, and it can successfully accomplish the identification of the complete digestive system. The WCE is a major step forward in the treatment of clinical digestive tract illnesses and the attainment of favourable outcomes. Injury to the intestinal lining as well as mucosal surfaces caused by the digestive system may take several forms, including bleeding, ulcers, tumours, and polyps. The primary purpose of capsule endoscopy is to determine the reason for internal bleeding inside the digestive tract, as well as to diagnose and treat inflammation or malignancies located in the small intestine. Pannu H S, 2020. However, this overview or assessment of these videos by practitioners can take many hours, which would be exhausting, causing them to miss components in which irregularities of the intestinal system are present. Because these components are very often only available in a few images of the image sequences, it is possible that they will miss such components.

Non-steroidal anti-inflammatory medicines, often known as NSAIDS, are among the pharmaceuticals that are used one of most regularly across the globe and can typically be acquired without a prescription. These drugs are associated with a high incidence of gastrointestinal side effects, which might include either the lower or upper intestinal tract. These 3 anchoring will allocate items of varying sizes to distinct predictions layers, which is appropriate given that the identification of objects in the scene should be uniformly dispersed across the scale. If the anchor method for calculating strategy that is focused on the standard Kmeans algorithm is used, and the majority of the lesion areas consist of large and medium-sized items, then the rest sections for identifying small parts will not be well skilled, or there will be no learning at all, that will waste the network's resources. P. Hurtik, "Poly-YOLO (2020). In addition, it is an excellent answer for situations in which the amount of data used for training is limited, such as the one we have. In the absence of this foundation, it could be challenging to learn these kinds of filtration with the restricted number of ulcer photos at one's disposal. In addition, the pre-trained network is capable of more accurately retain the complicated colon scenario than just its shallow version because to its deeper learning and larger dimensionality. We fixed the values of the early layers of the deep learning model to avoid over-fitting and speeding up the training process. Specifically, this was done to avoid over-fitting. In order to accomplish the fine-tuning, we put a number of different optimizers that included regularisation terms through their paces for the loss function. Because it is superior to the other methods, we used the (Adam) adaptive moment estimation technique in this study. This technique was previously used in El Jaafari I, 2020 optimization.

2 Literature Survey

Artificial intelligence (AI), also known as machine learning, has been playing an increasingly important part in the development of therapeutic procedures. In particular, the

utilization of AI in capsules endoscopy has garnered interest due to its ability to autonomously diagnose illnesses based on the videos and reduce the length of the procedure. Tsuboi, A. (2020). Reading time allotted by WCE Hricak, H. (2018). Reading WCE pictures is one of the primary applications of the (CNN) convolutional neural network developed by Lim D-H, (2020), which is the primary deep learning method used for image assessment. This application was documented in previous research. It has been shown that some CNN-based techniques are capable of accurately diagnosing a range of conditions affecting the small intestine, such as ulcers, polyps, Crohn's disease, angioectasia, and haemorrhage. Ding, Z. (2019); Hwang, Y., (2018); Tsuboi, A. (2020); Aoki, T. (2019); Leenhardt, R. (2018).

The study's findings demonstrated that the CNN had a sensitivity of 88.9 percent, a specificity of 87.4 percent, an accuracy of 87.7 percent, and a specific diagnostic time of 194 seconds, while the same indicators for 23 endoscopists were 79.0 percent, 83.2 percent, 82.4 percent, and 230 65 minutes, respectively. This shown that manually identification by practicing physicians is more time-consuming and less accurate than utilising a CNN to identify H. pylori infection, which also saved time. There is a lot more research that are the same as this one on AI and predicting HP infections. All of these findings, which were published in Zhang L. (2021) as well as Nakashima H. (2020), illustrate the higher sensitivity and accuracy of AI (Supplementary Table S1). Furthermore, AI's capacity to identify HP infections has still certain limits that will need to be addressed in the future. These constraints need to be overcome. For instance, the preparation and testing set for the aforementioned research were all collected from a single hospital or nation, and it is important to get more consistent and stringent external validation from a variety of sources in order to guarantee the legitimacy of the findings.

Wang S (2019) created a CNN framework to examine the different diagnostic capability of the algorithm for four distinct kinds of abdomen infectious diseases (lesion, polyp, esophagitis, and haemorrhaging) and seven various kinds of gastrointestinal ulcers (destruction, polyp, abscess, submucosal tumour, xanthoma, regular mucous lining, and haemorrhaging) Lim D-H, (202), accordingly. All of the study findings demonstrated that CNN is capable of providing superior diagnostic evaluation. The former achieved an accuracy of 96.5 percent, whilst the latter had some sensitivities of 96.2 percent, a specificity of 76.2 percent, as well as a correctness of 77.1 percent, respectively. There are numerous more research that are comparable. Guo L., (2021); Zhang L., (2021); Wang S., (2019), and the Supplementary provides additional details in greater depth.

Abstinence from NSAIDs is indicated before doing a VCE, often for a period of at least 4 weeks. NSAID-induced enteropathy represents the most significant complicating factor for small intestinal ulcers in individuals tested for suspected (CD) Crohn's disease. Sturm A, Maaser C (2019) Lim D-H (2020). Because of the high risk of intestinal bleeding that is connected with NSAID-induced mucosal damage (10), a number of therapeutic medications, including as misoprostol and probiotics, have been investigated for their potential to protect the mucosal lining of people who use NSAIDs. Lim D-H (2020) Mortensen B, (2019); Taha AS, (2018).

3 Methodology

This section will go into more depth on the suggested technique. The method may be broken down into three primary stages, all of which are represented in Figure 1. The very first step is the pre-processing of the picture that was supplied. After that, a CNN models that has been fine-tuned is used to extract feature representations. In the last step, the vector that was created is input into a random forest classification algorithm in order to identify ulcers.

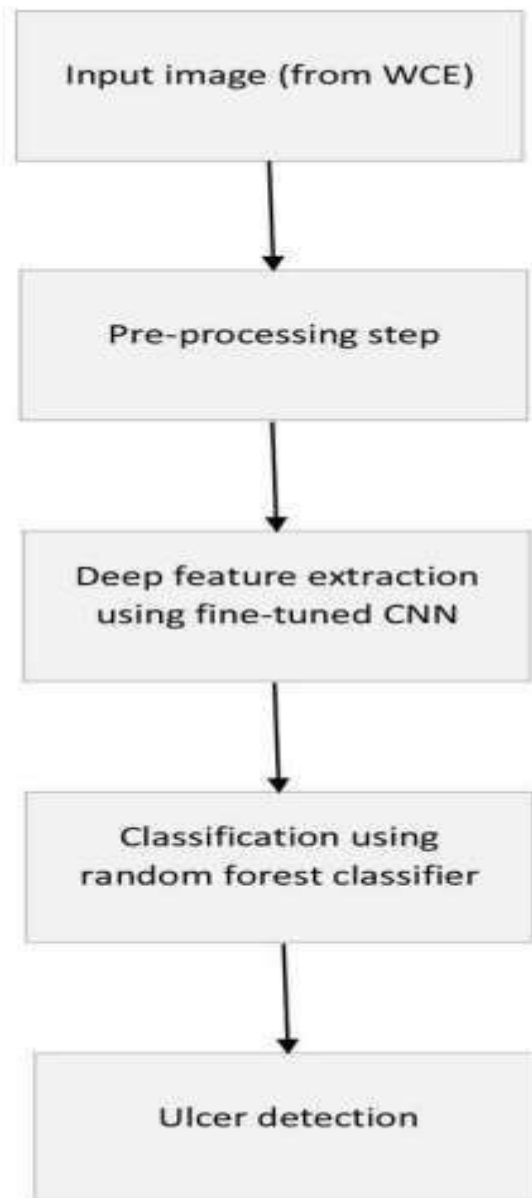


Figure 1: An algorithmic approach to the diagnosis of ulcers

As a result, one of our primary goals is to create a fresh model based on CNN that is not only useful but also suitable for clinical use. As a part of this effort, we built a CNN-based binary classification scheme that was able to sort pictures into those that included clinically relevant lesions and those that did not have such lesions. Because of this, the amount of time spent reading may be cut down significantly, as gastroenterologists will only need to look at

the photographs that were chosen by AI. As a result, we devised a CNN-based categorical classification algorithm for capsules endoscopic and tested it using situations where the capsule was not visible from the outside. These photos were extracted using MiroView 4.0 and saved in the PNG format at such a resolution of 320 by 320 pixels. Following the anonymization of the information, all of the images of the small bowel were examined by three seasoned gastroenterologists, who then manually labelled the images into one of seven distinct categories: regular mucosa, biliary dominant, air pockets, debris, inflamed mucous lining, unusual vasculature, as well as haemorrhaging.

3.1 Development of the CNN-based AI for auxiliary reading model

For the purpose of this study, the binary classification of CE images was performed using the InceptionResnet-V2 models that is part of the TensorFlow-Slim (TF-Slim) libraries. This prototype is well-known for its ability to quickly and efficiently collect multi-level features from an inception module and inner layer of a Resnet module to prepare an artificial intelligence. We used learning algorithm to employ a pre-trained models obtained from the imageNet database in order to improve the recognition rate.

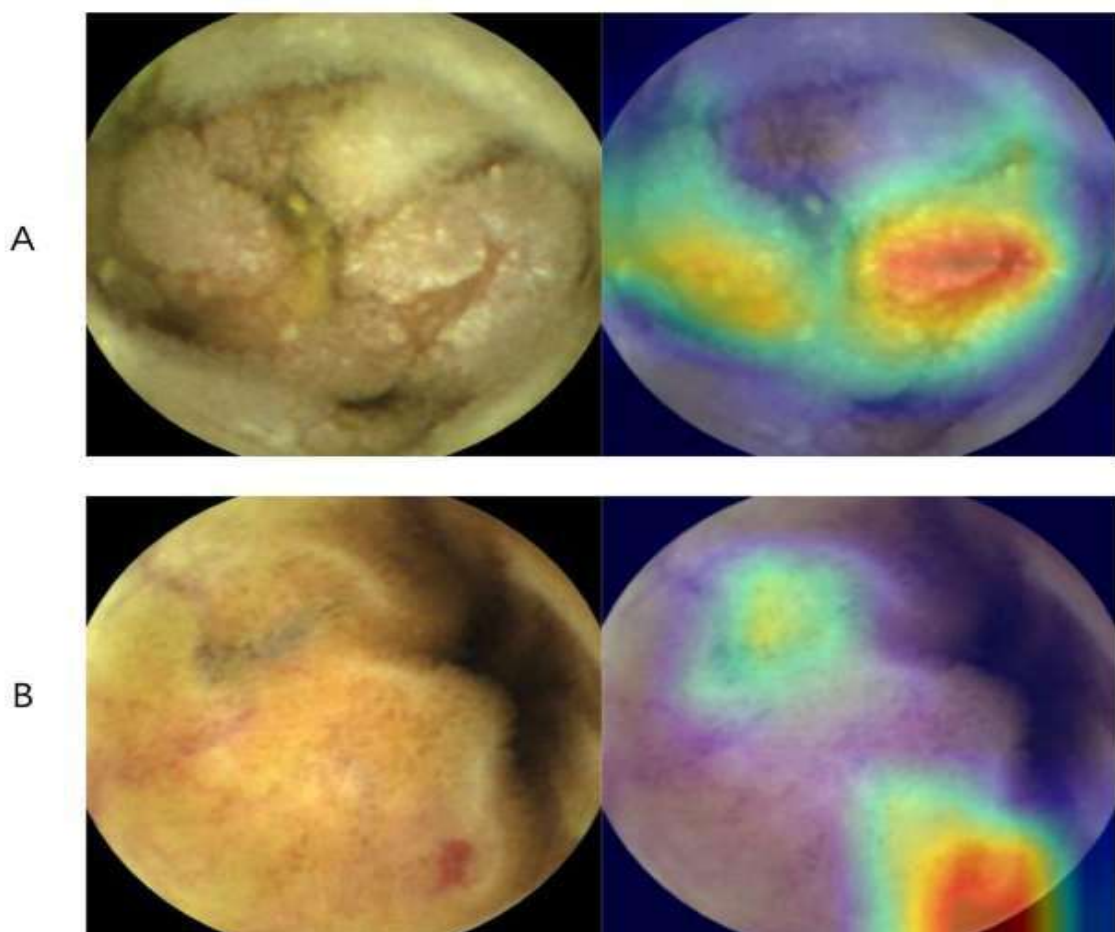


Figure 2. Activation profile of significant lesions for the class. (Image of a capsule endoscopy obtained with a Mirocam MC1600 and processed using MiroView 4.0—<http://www.intromedic.com/eng/main>)

To begin, we trained the final 'logits' layer for 10,000 iterations at a learning rate of 0.01, which was employed throughout the whole process. After that, we trained all of the layers using a training set of 0.0001 for a total of 200,000 steps. In the same way, while we were developing the first version of the InceptionResnetV2 model, we didn't make use of dropout and instead used the rectified linear unit (ReLU) as a control system. The total amount of time that transpired throughout the process of transforming the picture data in to the TFrecord (binary data easily accessible in Tensorflow) was 2621.208 seconds for the 80,000 test photos. During this time period, the amount of processing time necessary to evaluate the converted data was 431.258 seconds. Therefore, the suggested artificial intelligence demonstrated a processing capability of 26.208 frames per second. The following is an explanation of how (CAP) class activation maps were created using a channel-wise aggregation technique (Fig. 2).

3.2 Pre-processing

When it comes to identifying ulcers in WCE photos, visual information constitutes one of the most essential characteristics to look for. Furthermore, the colour features that are retrieved from several photographs are not necessarily same as, although if they depict the same anomaly. This is the case even when the images reflect the same abnormalities. This is because to differences in the lighting as well as the WCE devices that were used. As a result, we processed the photographs that were given to us utilising a few different stages. To begin, we improved the pictures obtained by the WCE by using several enhancement methods. In this instance, we made use of two distinct ways for enhancing the image: increasing the contrasts in the HSV colour space and balancing out the hue component. Lastly, in order to accommodate the pre-trained CNNs' expectation that the input data would be of the same size as those used during the training procedure, we scaled all of the testing and training pictures to a resolution of 224 by 224 pixels. During this step of the resizing process, the aspect ratio of any non-square photos was modified. The pre-processed picture provides a decrease in the amount of time needed for execution as well as the load of computing work. In addition, since it provides information on the illnesses that we desire to identify, it is helpful in determining which qualities may be classified as discriminants.

3.3 Deep feature extraction

The MobileNet models is the CNN structure that is being used this time for the purpose of deep learning extracting. It is a network that was suggested in Howard A G, (2017) by researchers working for Google Brain in 2017. In 2018, a modified version of such a system was made publically available. As its name implies, the purpose of this structure is to deliver machine learning to smart phones, which come with their own set of limitations in terms of memory, computational power, and energy supplies. In order to accomplish this goal, the authors focused their efforts on the most difficult aspect of the computing of convolutional networks, which is the convolution. One strategy for lowering the cost of the computation would be to minimize the amount of layers as well as the thicknesses of those layers. On the other hand, the effectiveness of machine learning is proportional to the number of levels as well as the representations filters that are used. The remedy that the team working on the Google Brain project decided to go with was therefore to develop a new new type of

convolution that they call (DSC) Depthwise Separable Convolution (Figure. 3). This new type of convolutions does not minimize the size, but it does decrease the number of computation and variables, as can be seen in the following example:

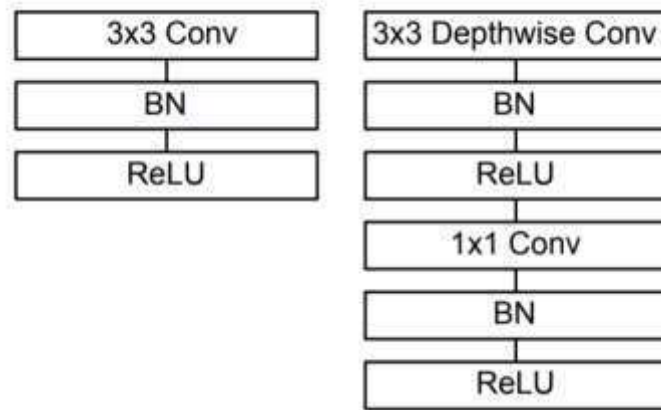


Figure 3: Convolutional layer with batch norm and ReLU activation, as shown on the left. To the right: horizontally Convolutions that are split up into their Depthwise and Pointwise components, which are then followed by batchnorm and ReLU.

3.4 Abnormality Recognition in WCE Videos

Conventional machine learning approaches using hand - crafted features as well as deep-learning techniques may broadly be separated into two categories to describe the derived from previous techniques for irregularity identification in WCE films. Conventional approaches to machine learning often begin with the human selection and creation of features, which is then followed by the deployment of a classifier. Color and texture are examples of elements that are often used in traditional approaches. In most cases, the colour of a lesion region will be different from the colour of the normal areas that surround it. For instance, hemorrhage regions may appear as red, while inflamed areas may show as yellow or white. The researchers Fu et al. (2019) devised a technique for the quick detection of bleeding that extracts colour features using the RGB colour space. In addition to the RGB colour space, additional colour spaces, such as the HSI/HSV [9] and the YCbCr colour space, are also often utilised in the process of feature extraction. Another form of feature that is often utilised for pattern recognition is called a texture. (LBP) Local binary patterns as well as filter-based characteristics are both examples of texture features, according to Aoki, T. et al (2019). A basic binary coding method that examines every pixel including its neighbours serves as the foundation for an LBP description. Figure 4 provides an illustration of sample ulcer frames taken from WCE films.

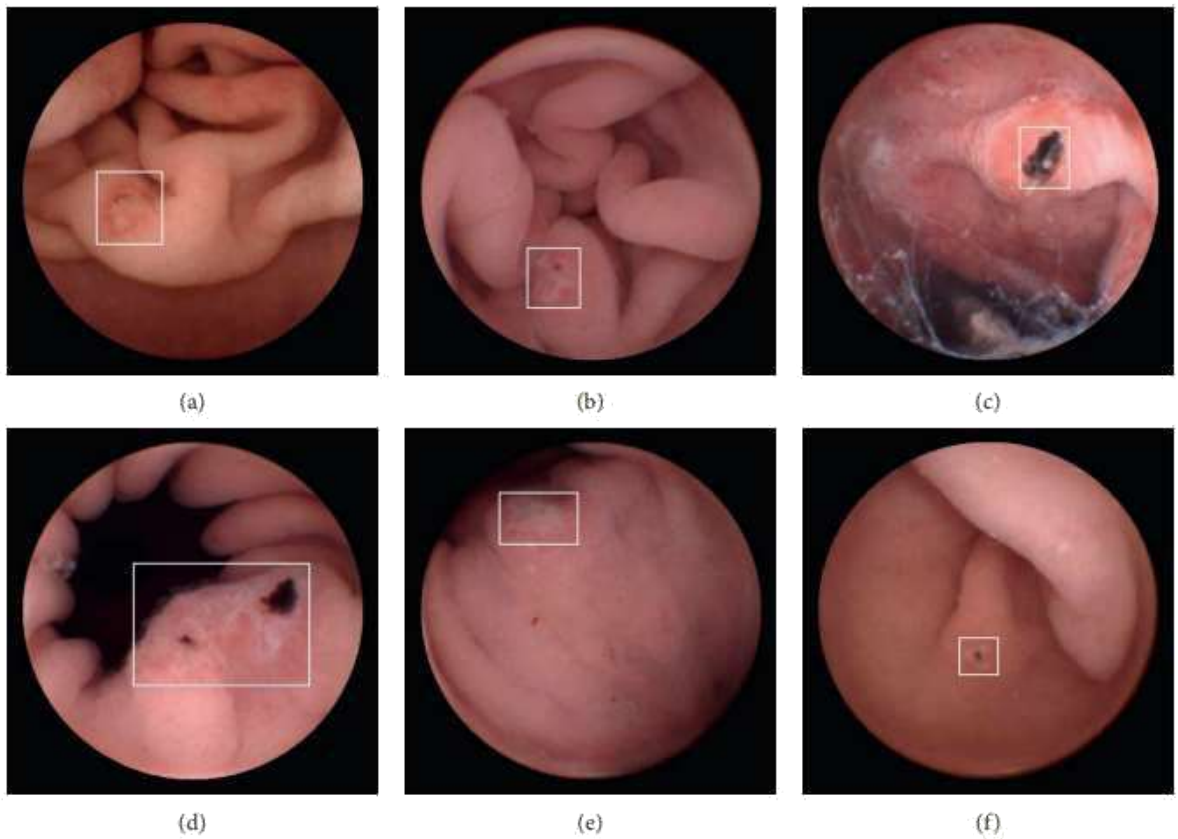


Fig 4: Images of an ulcer similar with WCE. A white box is superimposed over each picture to denote the ulcerated regions.

. The LBP descriptor, along with its expanded forms such like standard LBP Aoki, T. (2019) as well as monogenic LBP, has already been employed in a variety of different WCE identification tasks. Because of their capacity to characterise pictures in multistage space, filter-based characteristics, such like Gabor filters or wavelet transformations, are often used in WCE image identification tasks. Additionally, the many textural elements that make up an object may be merged in order to improve identification performance. As was established in the combining of discrete wavelet transform as well as standard LBP, automated polyp identification may be accomplished with a high degree of precision. Lesions are among the most frequent types of lesions that may occur in the gastrointestinal system, and it is estimated that ulcers affect around one out of every ten people. An region of tissue that has been damaged by gastric acid and has a discontinuities or rupture in a biological membrane is what is known as an ulcer. The colour and consistency of the ulcerated region are not the same as those found in a regular GI tract in the same location. The identification of certain types of ulcers needs the categorization of each picture included within a WCE video as either ulcerated or not, in a manner that is analogous to the categorization work required in tasks involving computer vision.

3.5 Classification

The primary objective of this stage is to get at a conclusion by generating a judgement that is based on the in-depth characteristics that were derived first from CNN. In this study, we used two different classifiers that competed with one another. The SVM is the initial classifier that plays a role in our work. Vapnik is credited as being the first person to present it. It is a method of classification that involves distinguishing between two or more groups of data sets by using one or even more hyper - plane as the dividing line (s). A hyper- plane which divides a group of data elements also has two hyperplanes that are parallel to one another. The objective here is to reduce as much as possible the inverted of the margins that exists between these hyperplans. In a variety of classification and detection tasks, the SVM classifier demonstrated strong performance. In this section, both the (RBF) radial basis function as well as the sigmoid kernel functions are evaluated by experimenting with different values for the regularisation parameter. The random forest classifier is the one that comes in second. It is an established classifier that is used for a variety of detecting responsibilities.

4 Result & Discussion

The primary purpose of this initial round of assessments is to choose amongst the several CNN architectures that are available. The outcomes of such an assessment are shown in Fig. 5 (a) and (b), respectively. The MobileNet, ResNet50, as well as VGG16 networks were put through their paces so that we could determine which CNN would serve our purposes the best. These designs shown excellent achievement in earlier categorization challenges, for example the ILSVRC challenge that took place in the current year.

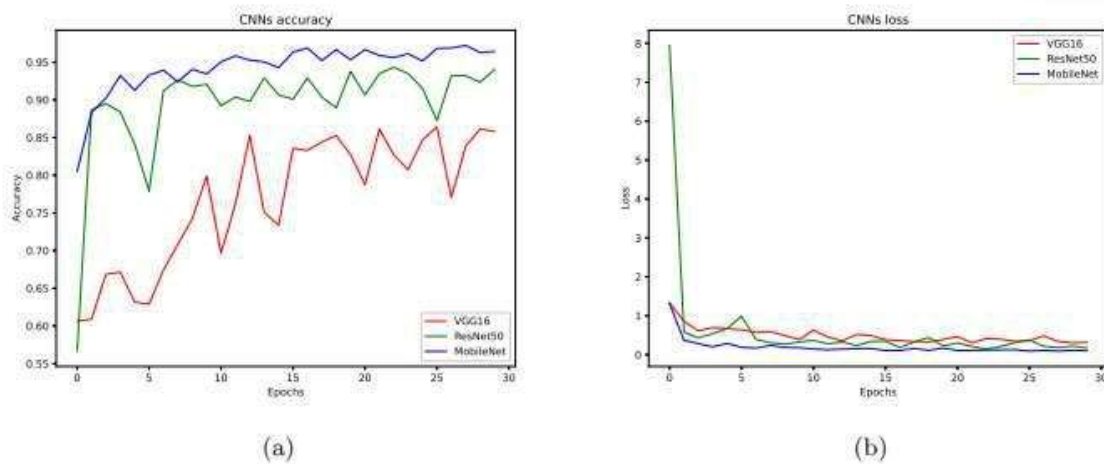


Fig 5: The findings of CNNs are broken down into two categories: (a) training accuracy and (b) training loss.

Each of the systems was first pre-trained by using ImageNet set of data, and then additional photos were used to fine-tune the models. The whole networks, such as its pooling, completely linked, and soft - max layer, was employed in this research. We note that the Mobile Net network had the greatest results, outperforming the ResNet50 networks by 0.341 and 0.276 accordingly. This is because the Mobile Net networks had an accuracy rate of 0.962 as well as

a loss of 0.056, whereas ResNet50 only achieved 0.056. The VGG16 network had the lowest accuracy rate and losses scores among all of the Convolutional networks that were considered for this study. Its scores were 0.837 and 0.497, accordingly. The next round of assessments consisted of comparing things between such a large number of classifiers in order to decide which one would be appropriate for use in our suggested strategy. The metrics of accuracy and recall, in addition to the F-measure, are used in the performance of this set of assessments. The following is a definition of those measures:

$$Precision = \frac{TP}{TP + FP}$$

$$F - measure = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

The numbers FP, FN, TP, and TN correspond, accordingly, to the total numbers of true positives, actual number of true negative, total number quantity of false false positive, and total number of false negatives. The results of a comparing between such a randomized forests containing 600 trees as well as a support vector machine that uses an RBF kernel, with C equal to 7 and G equal to 0.09, are shown in Table 2. The comparison looks at accuracy, recall, and F-measures. In order to extract characteristics that would later be supplied to all of those classifiers, fine-tuned MobileNet algorithm was used. These classifiers were used to take the position of the CNN's top two layers, which they had previously occupied. A comparison of this kind is carried out in order to provide justification for the classifier that we selected. We make the observation that perhaps the random forests performs superiorly than the SVM in all measures, outperforming it by a mean of 2.14 percent.

It demonstrates findings that are able to distinguish lesions of varying sizes, from big to tiny, and the recognition frameworks do not simply choose the region of the lesions that is being examined; rather, they also examine the tissue that is around the lesion. The combination of both large and tiny labelling allows for the identification of minor recurrent lesions as well. During the procedure of identification, numerous lesion sites are combined together to generate a larger is shown in detail in the accompanying fig.6.

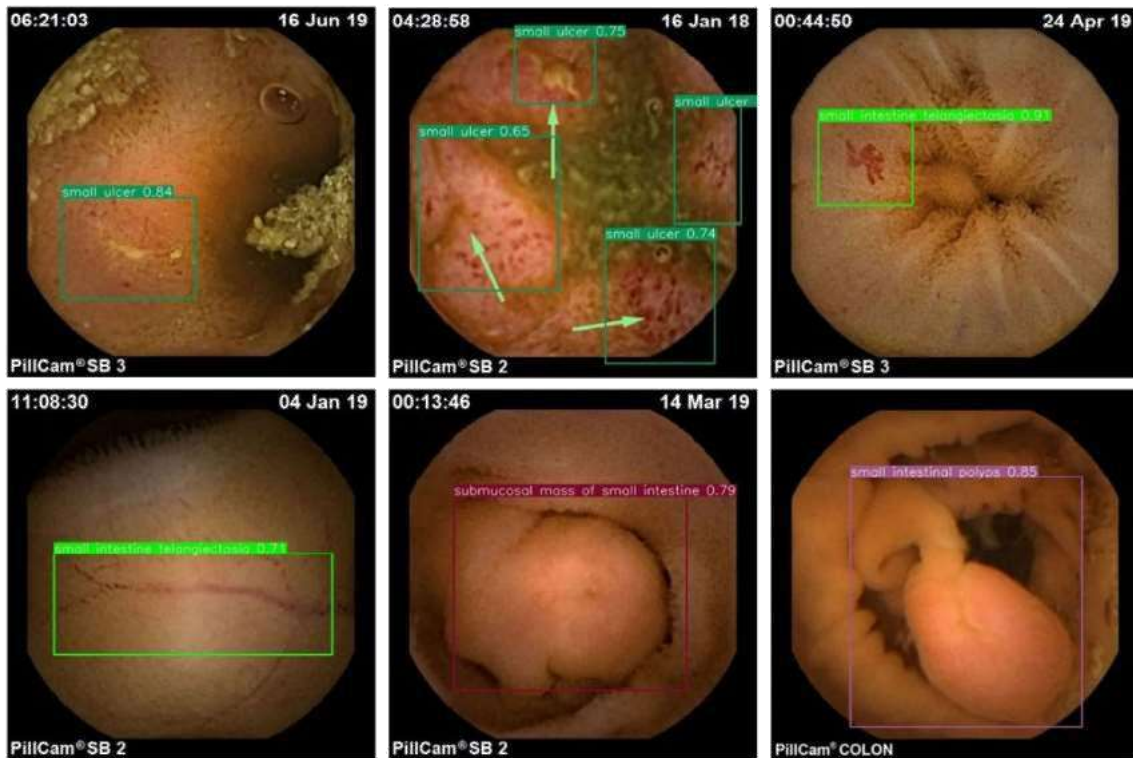


Fig 6. Results of the detection of lesions in the small intestine based on the enhanced YOLOv3.

When it comes to deep learning, the test data that are used during the training phase have a significant effect on the quality of training. This is something that is readily apparent for the particular target characteristics that are set against a simple backgrounds, as it makes it simple to learn during the training phase and produces a high detection accuracy. However, because of the complexities of the backdrop as well as the presence of a great deal of organic residues as well as other interfering variables, the process of learning the features is very challenging, and the detection performance is poor, which necessitates a range of approaches.

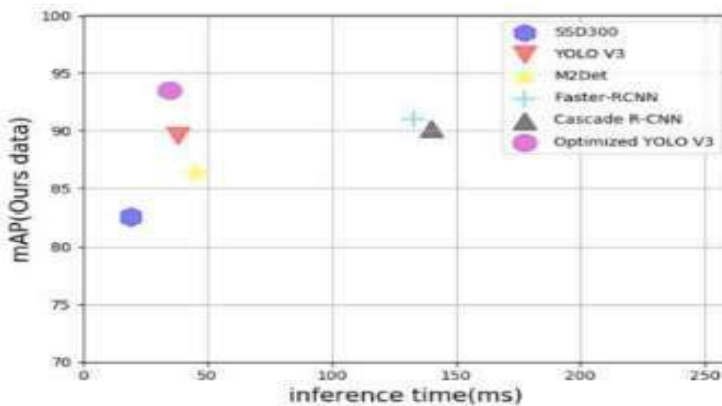


Fig. 7. The end results of efforts to enhance network detection

Following on from the previous figure, the one below is the outcome of enhanced network identification. As a result, the only photos of the small intestine that are impractical are those, whereas the article broadens the scope of the training and testing specimens. It incorporates all of the pictures of such digestive system organ as well as the typical lesions associated with them into the first training set. In total, there seem to be 27 groups, and the outcomes of the learning and examination are shown in Figure 7.

4.1 Outcomes and statistical analysis

The most important findings, which include the areas under the (ROC) receiver operating characteristic curves, as well as the sensitivities, specificities, as well as correctness of the CNN-based AI system, show whether or not each picture had lesions of clinical importance. Both the area of clinical relevance and the likelihood rating of the lesion were moulded by the trained model (range: 0–1). During the validation process, we determined that a certain score threshold would provide the best outcome. The Statistical Packages for Social Sciences, Version 20.0 was used for the analysis of the data (SPSS Inc., Chicago, IL, USA).

5 Comparison with state-of-the-art works

In order to evaluate the efficacy of the findings, we evaluated the method that was suggested with the most recent and cutting-edge methods. In order to do a useful comparison of this kind, you will need to utilise a same set of data. We contrasted the suggested approach with those that were discovered via the utilisation of the set of data. The following are some of the strategies that were employed in such works:

- a) a) Mostly in mean shift detection technique that is used, a multi-levels super-pixel representations that uses a max-pooling approach that is merged with the (LLC) Locality-constrained Linear Coding algorithm is used.
- b) The entire LBP, the laplacian pyramid, and the support vector machine are used in a multi-scale technique.
- c) A convolutional neural networks is employed in AlexNet for the purpose of detecting erosion and ulceration of the small intestine. The comparison study may be found in table 1, which can be seen below.

Table 1: A comparison of the suggested approach with more recent ones that have been published.

Method	Precision	Recall	F-measure
[26]	95%	95%	95%
[1]	79%	95%	87%
[27]	91%	96%	94%
The proposed method	96 %	95 %	96 %

In the result, it was validated on a fresh data set in order to evaluate the degree to which CNN as well as endoscopists vary in their ability to detect H. Pylori infections. The study's findings demonstrated that the CNN had a sensitivity of 88.9 percent, a specificity of 87.4 percent, an accurateness of 87.7 percent, as well as a specific diagnostic time of 194 seconds, while the same markers for 23 gastroenterologists were 79.0 percent, 83.2 percent, 82.4 percent, and 230 65 minutes, respectively. This shown that manually diagnostic by practitioners is more time-consuming and less accurate than utilising a CNN for diagnosis H. pylori infections, which also saved time.

6 Future scope

Even while efficiency and area under the curve (AUC) are typical measurements, they are not foolproof. One of the various problems that lie ahead is the creation of a framework that, in addition to taking into account the reliability of the internal algorithmic functioning, is also capable of quantifying the uncertainty level that is associated with machine learning. In conclusion, we established the practical application of a CNN-based complete binary classification framework inside the small intestinal CE. This model is an exciting new tool that has the potential to be implemented into regular clinical practise in the not too distant future. In our ongoing research, we want to investigate more methods in network architecture to further enhance ulcer identification. The most important goals for the near future are to achieve the robustness, adaptability, and variety of multipurpose AI technologies in PU and to perform multicentre prospective clinical studies as fast as feasible. In addition to this, the AUC of the system for identifying CD ulcer is comparable to such AUC of a system for identifying NSAID ulcer. These conclusions should be kept in mind in the upcoming planning and evaluation of VCE future research directions since machine learning is now undergoing a transformation that will alter gastro intestinal endoscopies.

7 Conclusions

As a result of this study, we suggested a brand new computer-aided method that is capable of automatically identifying ulcer illnesses based on WCE pictures. In order to collect the deep features, we employed a MobileNet model that had been fine-tuned. Because the accessible ulcer data sets only comprise a small number of pictures, the network that was used for this purpose went through a preliminary training phase that used the ImageNet set of data.

An analysis of randomised forest as well as SVMs classifier was carried out in order to facilitate the selection of the one that would be employed in the detection stage. This research suffers from a number of shortcomings. To begin, the artificial intelligence was trained using just photographs from a single organisation. Second, the labelling of the pictures was completed by 3 professionals, this means that there is the potential for some inter-observer's variances. In addition, sometimes it utilised data that had been acquired in the past, and throughout the training and reporting procedures, it did not use clinical data at all. It is quite crucial to ensure the dependability of the decisions that are made by AI in order to determine whether or not it can be employed in the actual world. It is a very challenging task to undertake a full uncertainty quantification and trustworthiness analysis, and there are debates over whether AI Technology should always be considered less or more. Even while accuracy and (AUC) area under the curve are typical measurements, they are not fool proof. One of the various problems that lie ahead is the creation of a framework that, in addition to taking into account the reliability of the internal algorithmic working, is also capable of quantifying the uncertainty level that is associated with machine intelligence.

In conclusion, it has been established that a CNN-based complete binary classification model may be practically applied in the small intestinal CE. This system is a promising technique that has the potential to be implemented in daily clinical practise in the not too distant future. This represents a retrospective study of information that was acquired in a prospective manner. It is necessary to conduct prospective interventional investigations in order to analyse the previously collected data and demonstrate the findings' applicability in the actual world. The number of photographs included in the research is rather low, especially for NSAID-induced ulcers. In addition, DL is a constantly developing discipline, which means that new algorithms are being created all the time. Since such Xception framework was employed in a prior effort that was connected to this one, we decided to apply it here as well (17). Nevertheless, this approach suffers from the following drawbacks: 1. The Xception Cnn architecture networks is now only suitable with said TensorFlow backend. 2. The Xception algorithm requires input with such a resolution of 299 by 299 pixels. 3. The images weights for the Xception algorithm range from 90 to 100 megabytes. It's possible that other systems will provide better outcomes. Last but not least, prospective research for upcoming projects need to explore network in real-world settings.

References

- [1] Kaplan, G.G., 2017. Does breathing polluted air increase the risk of upper gastrointestinal bleeding from peptic ulcer disease? *The Lancet Planetary Health* 1 (2), e54–e55.
- [2] Yuan, Y., Wang, J., Li, B., Meng, M.Q.-H., 2015. Saliency based ulcer detection for wireless capsule endoscopy diagnosis. *IEEE Transactions on Medical Imaging* 34 (10), 2046–2057.
- [3] J.-Y. Yeh, T.-H. Wu, and W.-J. Tsai, “Bleeding and ulcer detection using wireless capsule endoscopy images,” *J. Softw. Eng. Appl.*, vol. 7, no. 5, pp. 422–432, 2014.

- [4] Pannu H S, Ahuja S, Dang N, Soni S and Malhi A K 2020 *Multimedia Tools and Applications* 79 8145–8160.
- [5] P. Hurtik, V. Molek, J. Hula, M. Vajgl, P. Vlasanek, and T. Nejezchleba, “Poly-YOLO: Higher speed, more precise detection and instance segmentation for YOLOv3,” 2020, arXiv:2005.13243. [Online]. Available: <http://arxiv.org/abs/2005.13243>.
- [6] El Jaafari I, Ellahyani A and Charfi S 2020 *Signal, Image and Video Processing* 1–6.
- [7] Xiao, J. & Meng, M. Q. A deep convolutional neural network for bleeding detection in wireless capsule endoscopy images. In *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference* 639–642. <https://doi.org/10.1109/embc.2016.7590783> (2016).
- [8] Fan, S., Xu, L., Fan, Y., Wei, K. & Li, L. Computer-aided detection of small intestinal ulcer and erosion in wireless capsule endoscopy images. *Phys. Med. Biol.* 63, 165001. <https://doi.org/10.1088/1361-6560/aad51c> (2018).
- [9] Alaskar, H., Hussain, A., Al-Aseem, N., Liatsis, P. & Al-Jumeily, D. Application of convolutional neural networks for automated ulcer detection in wireless capsule endoscopy images. *Sens. (Basel, Switz.)* <https://doi.org/10.3390/s19061265> (2019).
- [10] Aoki, T. et al. Automatic detection of erosions and ulcerations in wireless capsule endoscopy images based on a deep convolutional neural network. *Gastrointest. Endosc.* 89, 357-363.e352. <https://doi.org/10.1016/j.gie.2018.10.027> (2019).
- [11] Klang, E. et al. Deep learning algorithms for automated detection of Crohn’s disease ulcers by video capsule endoscopy. *Gastrointest. Endosc.* 91, 606-613.e602. <https://doi.org/10.1016/j.gie.2019.11.012> (2020).
- [12] Leenhardt, R. et al. A neural network algorithm for detection of GI angiectasia during small-bowel capsule endoscopy. *Gastrointest. Endosc.* 89, 189–194. <https://doi.org/10.1016/j.gie.2018.06.036> (2019).
- [13] Aoki, T. et al. Automatic detection of blood content in capsule endoscopy images based on a deep convolutional neural network. *J. Gastroenterol. Hepatol.* <https://doi.org/10.1111/jgh.14941> (2019).
- [14] Tsuboi, A. et al. Artificial intelligence using a convolutional neural network for automatic detection of small-bowel angioectasia in capsule endoscopy images. *Dig. Endosc.: Of. J. Jpn. Gastroenterol. Endosc. Soc.* 32, 382–390. <https://doi.org/10.1111/den.13507> (2020).
- [15] Ding, Z. et al. Gastroenterologist-level identification of small-bowel diseases and normal variants by capsule endoscopy using a deep-learning model. *Gastroenterology* 157, 1044-1054.e1045. <https://doi.org/10.1053/j.gastro.2019.06.025> (2019).

- [16] Hwang, Y., Park, J., Lim, Y. J. & Chun, H. J. Application of artificial intelligence in capsule endoscopy: Where are we now?. *Clin. Endosc.* 51, 547–551. <https://doi.org/10.5946/ce.2018.173> (2018).
- [17] Hricak, H. 2016 New horizons lecture: Beyond imaging-radiology of tomorrow. *Radiology* 286, 764–775. <https://doi.org/10.1148/radiol.2017171503> (2018)
- [18] Nakashima H, Kawahira H, Kawachi H, Sakaki N. Endoscopic threecategorical diagnosis of *Helicobacter pylori* infection using linked color imaging and deep learning: a single-center prospective study (with video). *Gastric Cancer.* (2020) 23(6):1033–40. doi: 10.1007/s10120-020-01077-1.
- [19] Wang S, Xing Y, Zhang L, Gao H, Zhang H. Deep convolutional neural network for ulcer recognition in wireless capsule endoscopy: experimental feasibility and optimization. *Comput Math Methods Med.* (2019) 2019:7546215. doi: 10.1155/2019/7546215
- [20]. Guo L, Gong H, Wang Q, Zhang Q, Tong H, Li J, et al. Detection of multiple lesions of gastrointestinal tract for endoscopy using artificial intelligence model: a pilot study. *Surg Endosc.* (2021) 35(12):6532–8. doi: 10.1007/s00464-020-08150-x
- [21]. Zhang L, Zhang Y, Wang L, Wang J, Liu Y. Diagnosis of gastric lesions through a deep convolutional neural network. *Dig Endosc.* (2021) 33:788–96. doi: 10.1111/den.13844.
- [22] Lim D-H, Jung K, Lee SB, Park IK, Cha HJ, Park JH, et al. Non-steroidal anti-inflammatory drug-induced enteropathy as a major risk factor for small bowel bleeding: a retrospective study. *BMC Gastroenterol.* (2020) 20:178. doi: 10.1186/s12876-020-01329-5
- [23] Mortensen B, Murphy C, O’Grady J, Lucey M, Elsafi G, Barry L, et al. *Bifidobacterium breve* Bif195 protects against small-intestinal damage caused by acetylsalicylic acid in healthy volunteers. *Gastroenterology.* (2019) 157:637– 646.e4. doi: 10.1053/j.gastro.2019.05.008.
- [24] Taha AS, McCloskey C, McSkimming P, McConnachie A. Misoprostol for small bowel ulcers in patients with obscure bleeding taking aspirin and non-steroidal anti-inflammatory drugs (MASTERS): a randomised, doubleblind, placebo-controlled, phase 3 trial. *Lancet Gastroenterol Hepatol.* (2018) 3:469–76. doi: 10.1016/S2468-1253(18)30119-5.
- [25] Sturm A, Maaser C, Calabrese E, Annese V, Fiorino G, Kucharzik T, et al. ECCO-ESGAR guideline for diagnostic assessment in IBD part 2: IBD scores and general principles and technical aspects. *J Crohns Colitis.* (2019) 13:273– 84. doi: 10.1093/ecco-jcc/jjy114
- [26] Maaser C, Sturm A, Vavricka SR, Kucharzik T, Fiorino G, Annese V, et al. ECCO-ESGAR guideline for diagnostic assessment in IBD part 1: initial diagnosis, monitoring of known IBD, detection of complications. *J Crohns Colitis.* (2019) 13:144–64. doi: 10.1093/ecco-jcc/jjy113.

[27] Howard A G, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M and Adam H 2017 Mobilenets: Efficient convolutional neural networks for mobile vision applications (Preprint 1704.04861)