ISSN No:-2456-2165

Digestive Tract Abnormalities Classification Using Wireless Capsule Endoscopy Data

Harshkumar Modi
Computer Science Department
SRM Institute of Science & Technology
Chennai, India

Shashwat Misra
Computer Science Department
SRM Institute of Science & Technology
Chennai, India

Shashwat Gaur
Computer Science Department
SRM Institute of Science & Technology
Chennai, India

Abstract:- Wireless capsule endoscopy (WCE) is proving to an extremely beneficial method for perceiving ailments inside of human gastric pathway. Through WCE Doctors identify the diverse irregularities like chronic diarrhoea, ulcer, bleeding, polyps, cancer/tumour in small intestine and Crohn's disease, any of it in the gastrointestinal tract that is invasive. It is a method that used a pill-size wireless camera to get visual data of the digestive tract of a human body. It is a method that helps doctor to see inside your intestine, an area which is normally difficult to reach to through standard endoscopy procedure or imaging methods. Scholars are finding new ways to progress and enhance the metrics of performance of WCE using models that detect these ailments at a quicker rate of improvement, on their own. With datasets produced from this dataset are imporving with better technology and imaging services, machine learning grants us an opportunity to diagonis and get better insights from this data. Utilizing these methods, we propose a system to train and predict 13 classes of anomalies inside the digestive tract. The proposed model has achieved an accuracy of 97.82%.

Keywords:- *Wireless Capsule Endoscopy, Machine Learning, Digestive Tract.*

I. INTRODUCTION

Many individuals all over the world are afflicted by disorders of the gastrointestinal system, particularly the abdominal, small and large intestine. In Bangladesh, over 27.8% of individuals pass away from gastrointestinal infections. Around 1.6 million people in the United States are affected by the Inflammatory bowel disease (IBD), while a rise of over 200,000 accounted since the year 2011. Morethan 70,000 new cases of IBD are identified each year. Children with Crohn's disease (CD) or ulcerative colitis (UC) account for 80,000 of these. Globally, IBD is on the rise. It is necessary to discover GI tract disorders as soon as feasible in order to cure them.

Artificial intelligence (AI) has proved to be the key in recent advanements of the Wireless capsule endoscopy (WCE) technology. The probable lies in fine detection of the irregularities while reducing or dropping manual labour. As the process requires finely labelled data, it is hard to be found as the process requires trained professionals/medics to make sure that the data is labelled correctly, who in most cases do not have the time to do it. And even it being done, access to such data for researchers is hard to get to. Howover is some cases we do find access to some data, to which, we in our research utilize Kvasir-Capsule dataset, a huge WCE dataset collected from inspections at Hospitals in Norway.It consists of 118 videos, used to extract a total of 44,228 frames with a bounding box around detected irregularities from 13 diverse classes of discoveries. Initial work reveals the possible benefits of AI-based computer-assisted analysis systems for VCE. Yet, they also display that there is boundless potential for developments, and well labelled datasets can prove to be the key in the advancements of the WCE technology.

II. RELATED WORK

Old endoscopy methods were difficult on the patients as they were hard to perform and caused pain to them. It is to eliminate these drawbacks, wireless capsule endoscopy examinations were presented to provide trouble-free, and accurate diagnosis of the stomach tract. The process is uniquely pain-free as it uses a uses a small wireless camera to capture images of your gastric pathway. In the capsule endoscopy, a camera is fitted inside a pill sized capsule that you swallow in your body. As the capsule passes along your digestive tract, the camera starts recording a video which is then directly communicated to a recorder you wear on a belt across your abdomen.

Capsule endoscopy allows medics to have a detailed view inside your small intestine — an area that is hard to reach with more-traditional endoscopy measures. With image data being generated from this procedure, processing this data using computer aided and machine learning methods allow us to predict or diagonis the abnormality earlier without much efforts taken by the physician. The methods previously used started with basic image manipulation to extract data and give conclusions from the same.

The early research was mainly focused on finding areas of bleeding in the digestive tract, with a considerably low sensitivity and specificity of ~22% and ~42%. Later a feature extractor called MPEG-7 visual descriptor was used to find the abnormalities like ulcer and bleeding. The ulcer detected from the capsule endoscopy pictures having indication of bleeding were processed using chromaticity moment by means of HSI (Hue-Saturation-Intensity) color space. In statistical feature-based bleeding detection from capsule endoscopy video, recognition is obtained by histogram from indexed images in capsule endoscopy videos. The inference from the histogram was used to differentiate normal areas from bleeding/ulcer areas. In texture and color-based image division and pathology recognition in capsule endoscopy videos have anticipated a Color Coherence Vector (CCV) based feature for classifying ulcer pictures.

Initial methods used this HSV to process data and then run it through a logistic regression classifier. Video data was converted to image frames, after which a color threshold was applied to identify regions where blood was there. The image, after being processed, was converted to binary and trained through the classifier that separated ulcer regions from the non-ulcer ones. Having trained and tested on publicly available data, the obtained accuracies were 87.70%.

Similar work was observed in one of the other researches where the method proposed to detect frames of bleeding by extracting regions of interest. These ROIs were extracted using by means of Exception Maximization (EM) created image separation method and Gaussian mixture distribution. Once done, the model was trained using Support Vector Machine, giving a resultant model that classifies bleeding frames from the non-bleeding normal frames.

This dataset was however to small to train, having only ~4000 images in total with both normal and bleeding lesions frames.

These approaches may have not achieved higher accuracies but they did find a unique approach of using regions of interest to get better features extraction from the dataset.

This finding gave another unique approach that did not rely on computer aided analysis or machine learning to extract features. The approach presented was to use specific wavelength LED in the pill, in order to define and color the areas of lesion or bleeding correctly in such order that they are reflected clearly in the outcoming data. Once done, various machine-learning algorithms were used to train and asses the data. Through these features were extracted more efficiently, helping to attain better accuracies.

Furthermore, with increase in deep learning approaches, the image dataset available were trained and tested on various such models, proving how deep learning helped in attaining better accuracies in compared to machine learning approaches. Applied models used to classify any ulcer or worm disease in the digestive tract. The approaches had lesser data of the ulcer catagories which was a hinderance to achive higher accuracies. But with more publicly available data upcoming in the world, we hope to achieve better results with the same algorithms.

From what is observed from these works is that support vector machine (SVMs), KNN, AlexNet, all have proved to be the most efficient methods till now.

However, despite advancements in the classifying approaches, one another obstacle to tackle is to locate the exact location of the pill where it finds the lesions inside the body. The location would be the only way through which doctors can look for solutions to cure the problem. Using methods like belt around the waist that continuously takes the data from the pill, more approaches need to be researched to make the whole process accurate to predict the disease.

III. PROPOSED WORK

Analyzing the methods used by researchers to overcome the fundamental problems of processing Wireless Capsule Endoscopy, this paper takes the Kvasir-capsule data, the largest available data in the market to train upon a CNN network that was hand-crafted after multiple approaches to achieve higher accuracies.

The model is developed using the ‘TensorFlow’ and ‘Keras’ libraries available on python .

TensorFlow is an open-source machine learning platform that works from beginning to end. It features a big, flexible network of libraries, tools, and public resources that allow academics to advance in machine learning and developers to quickly construct and present ML applications. TensorFlow has many stages of abstraction, allowing one to pick the best fitting way that aligns with the requirements. Creating and assessing models with the Keras API makes getting started with TensorFlow and machine learning simple.

Keras is a human-level understandable API, not a machine-centric one. Keras follows to best practises for minimising cognitive load, such as presenting uniform and simple APIs, decreasing the amount of user steps necessary for typical use cases, and providing clear and actionable fault messages. It comes with a fair share of documentation and developer instructions.

A Convolutional Neural Network (ConvNet/CNN) is a deep learning system that can take an input picture, assign significance (learnable weights and biases) to several features/objects in the image, and differentiate amongst them. When associated to other classification approaches, the total pre-processing required by a ConvNet is pointedly less. While basic techniques need hand-engineering of filters, ConvNets can learn these filters/characteristics with enough training.

The proposed work uses libraries from Keras that allows to process the data and load it for training. Using the ImageDataGenerator library we prepare the data by first rescaling it and splitting it for validation. As the data was split in a folder and the labels corresponding to the images were stored in a csv file, we combined the data using “flow_from_dataframe”. The function processed the data and training and validation.

Up next, we defined a sequential model having 3 convolve layers, with pooling and dropout layers after each convolve layers. With a kernel of (2,2) and pooling size of

ISSN No:-2456-2165

(3,3), we extracted the information from the images. Later using the dense function of 128 units we trained it, before finally doing the multip classification using another dense layer. The activation functions used were "relu" and "softmax".

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 335, 335, 16)      208
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 111, 111, 16)      0
_________________________________________________________________
dropout (Dropout)            (None, 111, 111, 16)      0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 110, 110, 32)      2080
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 36, 36, 32)        0
_________________________________________________________________
dropout_1 (Dropout)          (None, 36, 36, 32)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 35, 35, 64)        8256
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 11, 11, 64)        0
_________________________________________________________________
dropout_2 (Dropout)          (None, 11, 11, 64)        0
_________________________________________________________________
flatten (Flatten)            (None, 7744)              0
_________________________________________________________________
dense (Dense)                (None, 128)               991360
_________________________________________________________________
dense_1 (Dense)              (None, 14)                1806
=================================================================
Total params: 1,003,710
Trainable params: 1,003,710
Non-trainable params: 0
```

Figure 1:Model Architecture

Using the "categorical_crossentropy" loss function and RMSprop optimizer, the learning rate was set as 0.001. The metrics used to measure the performance of he model were accuracy. Model was trained with a batch size of 64 and for 30 epochs.

IV. DATASET

The Open Science Framework (OSF) provides an open access to the Kvasir-Capsule dataset.The dataset contains of more than 4,700,000 key data records, i.e., almost 47,200+ pictures labels around, the 43 matching categorized videos and 74 non-categorized videos. 4,694,266 non-categorized pictures can further be take out from all the videos collective. Compiling it all, the dataset has a entire size of ~89 GB.

Due to limited training resources we have trained and tested our model on the labelled dataset of images and videos only.

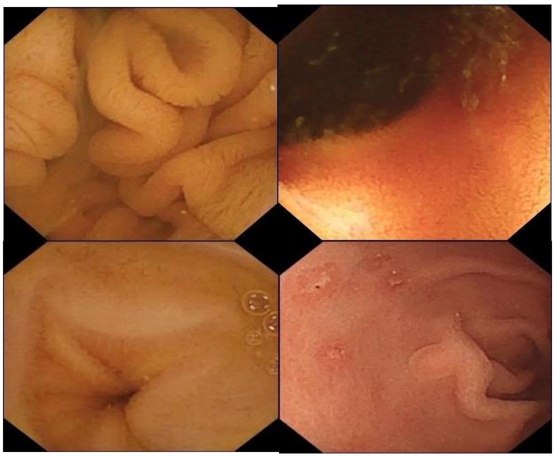

Figure 2: images in the dataset

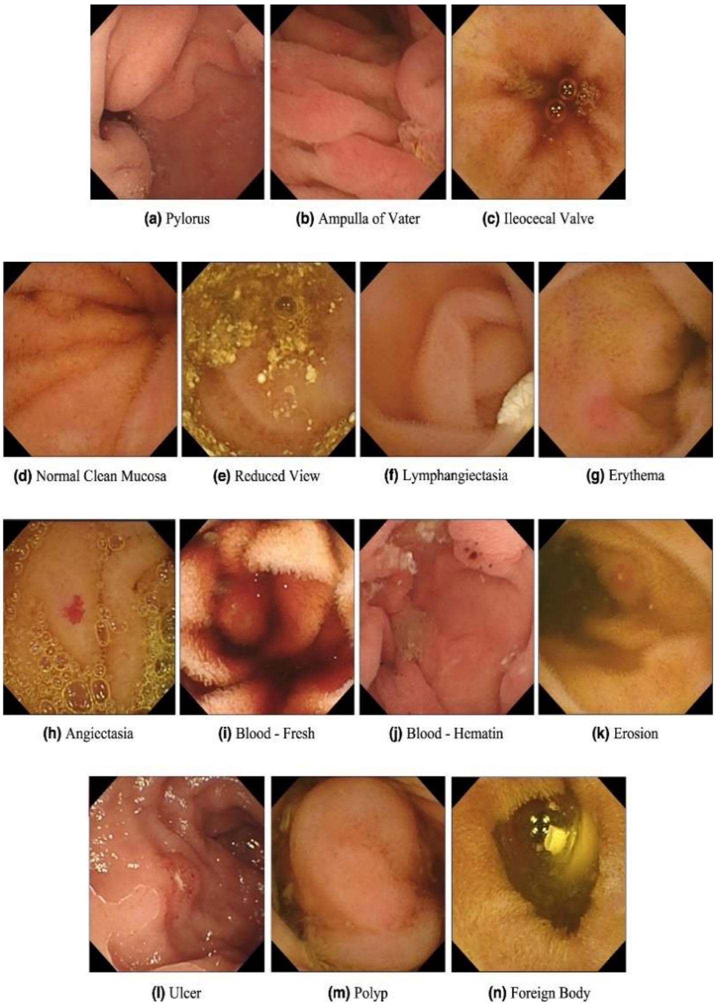

Figure 3:Some examples of Lesions

V. IMPLEMENTATION

The workflow pipeline has been divided into four modules:

a) DATA PREPROCESSING

The dataset was well preprocessed with all the images of the same size, so no preprocessing was required to clean the data or make changes to the given images. Using the ImageDataGenerator library, we matched the image file with their corresponding label in the csv file.

b) MODEL DEVELOPMENT

The model was trained using CNN Sequential model, which had all the layers and parameters custom tuned according to the responses on different training parameters before. It consisted of convolve layer, pooling layer and dropout layer that extracted features which finally were trained using a dense layer network for multiclassification of the different types of abnormalities in the digestive tracts.

c) TRAINING AND VALIDATION

The dataset is fairly imbalanced as majority of the images belonged to the normal mucas class. This was a limitation the dataset, despite which having a fair amount of images in the other datasets, the model for able to fairly detect the correct class of abnormality. Having received a lesser accuracy on the Adadelta optimizer, we later used RMSprop optimizer to find better accuracy. The conclusive training accuracy was of 87.82% with validation accuracy of 94.44%.

VI. RESULTS DISCUSSION

The presented system was trained using jupyter notebook on anaconda envioursnment on an Apple M1 chip GPU. The initial experiments gave an accuracy of 72%. However, repeated experiments to find the optimal learning rate and optimizer eventually led the model to achieve an accuracy of 97.82%.

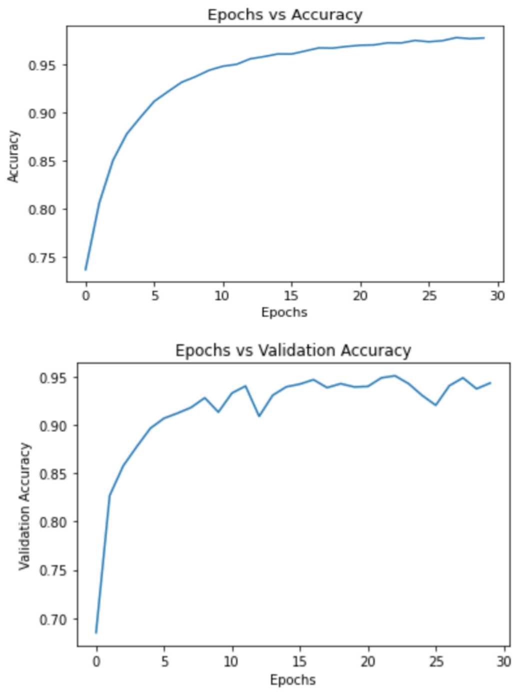

Figure 4: Final training metrics achieved

The dataset was divided into 75% and 25%, with 75% for training and 25% for validation.

Each batch sized trained 64 images at once, which were at the limitation of the GPU used to train the data. The training could be better if larger batch size training could be done.

VII. CONCLUSION

A technique for a challenging dataset has been developed here by means of CNN. The proposed can classify 13 classes of abnormalities in the digestive tract with an accuracy of 97.82%. The performance of the proposed method amplified when the right optimizer and learning rates were applied. The change in the filters of the convolve layer also added to the accuracy of the model. The model's accuracy could be increased using a more balanced dataset for training and better images of improved quality.

REFERENCES

[1]. Z. Amiri, H. Hassanpour and A. Beghdadi, "A Computer- Aided Method to Detect Bleeding Frames in Capsule Endoscopy Images," 2019 8th European Workshop on Visual Information Processing (EUVIP), 2019, pp. 217-221, doi: 10.1109/EUVIP47703.2019.8946168.

[2]. Y. Yuan, B. Li and M. Q. -. Meng, "Bleeding Frame and Region Detection in the Wireless Capsule Endoscopy Video," in IEEE Journal of Biomedical and Health Informatics, vol. 20, no. 2, pp. 624-630, March 2016, doi: 10.1109/JBHI.2015.2399502.

[3]. A. H. Khan, M. H. A. Sohag, S. S. Vedaei, M. R. Mohebbian and K. A. Wahid, "Automatic Detection of Intestinal Bleeding using an Optical Sensor for Wireless Capsule Endoscopy," 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2020, pp. 4345-4348, doi: 10.1109/EMBC44109.2020.9176340.

[4]. A. A. Mamun and M. S. Hossain, "Ulcer Detection in Image Converted from Video Footage of Wireless Capsule Endoscopy," 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), 2019, pp. 1-4, doi: 10.1109/ICASERT.2019.8934597.

[5]. Anastasios Koulaouzidis, Dimitris K Iakovidis, Alexandros Karargyris & John N Plevris (2015) Optimizing lesion detection in small-bowel capsule endoscopy: from present problems to future solutions, Expert Review of Gastroenterology & Hepatology, 9:2, 217-235, DOI: 10.1586/17474124.2014.952281

[6]. 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. *Sensors* **2019**, *19*, 1265. https://doi.org/10.3390/s19061265

[7]. A. Karargyris and N. Bourbakis, "Detection of Small Bowel Polyps and Ulcers in Wireless Capsule Endoscopy Videos," in IEEE Transactions on Biomedical Engineering, vol. 58, no. 10, pp. 2777-2786, Oct. 2011, doi: 10.1109/TBME.2011.2155064.

[8]. Bchir, O., Ben Ismail, M.M. & AlZahrani, N. Multiple bleeding detection in wireless capsule endoscopy. *SIViP* **13,** 121–126 (2019). https://doi.org/10.1007/s11760-018-1336-3.

[9]. Dimitris K. Iakovidis, Anastasios Koulaouzidis, Automatic lesion detection in capsule endoscopy based on color saliency: closer to an essential adjunct for reviewing software, Gastrointestinal Endoscopy, Volume 80, Issue 5, 2014, Pages 877-883,ISSN0016-5107,https://doi.org/10.1016/j.gie.2014.06.026.

[10]. Moglia, A., Menciassi, A., Dario, P. *et al.* Capsule endoscopy: progress update and challenges ahead. *Nat Rev Gastroenterol Hepatol* **6,** 353–361 (2009). https://doi.org/10.1038/nrgastro.2009.69

[11]. Iakovidis, D., Koulaouzidis, A. Software for enhanced video capsule endoscopy: challenges for essential progress. *Nat Rev Gastroenterol Hepatol* **12,** 172–186 (2015). https://doi.org/10.1038/nrgastro.2015.13

[12]. Hale MF, Sidhu R, McAlindon ME. Capsule endoscopy: current practice and future directions. *World J Gastroenterol*. 2014;20(24):7752-7759. doi:10.3748/wjg.v20.i24.7752

[13]. Smedsrud, P.H., Thambawita, V., Hicks, S.A. *et al.* *Kvasir-Capsule*, a video capsule endoscopy dataset. *Sci Data* **8,** 142 (2021). https://doi.org/10.1038/s41597-021-00920-z

[14]. Deep learning for wireless capsule endoscopy: A systematic review and meta-analysis, Gastrointestinal Endoscopy, Shelly Soffer, Eyal Klang, Orit Shimon, Noy Nachmias, Rami Eliakim, Shomron Ben-Horin, Uri Kopylov, Yiftach Barash, Volume 92, Issue 4, 2020, Pages 831-839.e8, ISSN 0016-5107, https://doi.org/10.1016/j.gie.2020.04.039