An Empirical Study on Hookworm Detection

Anuja Kharatmol¹, A. M. Bagader²

¹PG Student, Department of Information Technology, Pune Institute of Computer Technology, Pune, India
²Professor, Department of Information Technology, Pune Institute of Computer Technology, Pune, India

Abstract: Hookworms are the curse to the human health, the process of detection of hook worm is always a challenging task to the medical science. So to overcome this wireless capsule endoscopy is being applied in the advanced clinical procedures. Even though that it remains a challenging task to identify the hook worms even by the pictorial sights. So to overcome this many methodologies are existed out of that most advanced and recent one is using of the Convolution neural network (CNN). The one major limitation of the CNN is it usually needs a large amount of labeled data for training and fine tuning the parameters. And most of the time it is not possible to provide such a large data for the detection process. So proposed methodology put forwards an idea of detecting the hookworm using moderate amount data which includes the steps of image data weighing feature selection, Region of interest, Convolution neural network and finally fuzzy classification through which hookworms are detected efficiently.

Keywords: CNN, Hookworms, Region of interest (ROI), Endoscopy.

1. Introduction

Hookworms are intestinal blood feeding parasitic nematode worms that live off other living things and have strong buccal hooks that they use for attaching to the host intestinal lining. The hook contains dangerous blood sucking pests. There are many serious effects of hookworm in our lungs, small intestine, and lethargic anemic state in humans. The main infection caused by hookworm in human is named as helmintiases. The two main types of round worm which causes hookworm infection in human is belong to genera Necator and Ancylostoma. The main source of hookworm infection in human is hookworm larvae found in contaminated dirt of feces. The usual route of hookworm to enter into host body is through the skin. There are 576 to 740 million people in the world infected by hookworms people in developing countries are more prone to hookworm infection due to bad sanitation system in these countries.

The main symptoms occur after hookworm infections are itchiness in skin, nausea, abdominal pain, blood in stool, fever, diarrhea etc. It is important to detect hookworm infection on time, otherwise if it last long the person could become anemic. Anemia causes a drop in red blood cell count in our body, which can contribute to heart failure in extreme case. Another complication arises due to hookworm infections are ascites and nutritional deficiencies. In ascites the protein loss occurs in the body due to which fluid buildup in our abdomen. Children who are more prone to hookworm infection show the less physical and mental growth of their body due to loss of protein and iron. The treatment of hookworm infected host start by improving nutrition, oral medications prescribed by doctor like mebendazole and albendazole and treat anemia complications. In case of ascites the additional protein is also included in our diet. The better sanitization system also prevents hookworm infection in people. Many developing countries like India start educating people for using toilet to prevent them for hookworm infection. Centers for Disease Control and Prevention mentioned standard method for hookworm detection is by microscopic examination of hookworm eggs in the infected person stool sample. In case of light infection, it is difficult to find eggs, so a concentrated procedure is recommended. The blood sample is also taken to found iron deficiency and anemia. The stool sample should be properly examined for the consistency, color, pus presence, blood, presence of worms. There are many microscopic methods are present in examined sample stool. Some of these are:

- **Direct Microscopic:** The direct wet smear method such as Lugol stain it is fast and take less time to prepare and inexpensive in comparison with other technique. But the main demerit of this technique is it didn’t find hook worm eggs if the concentration is too low, or if too much fat is present in sample stool.
- **FEC Method:** FEC (Formol Ether Concentration) technique is utilized to detect hookworm infection in stool which has high fat or debris presence. It is a sedimentation method used to find protozoan cysts and hookworm eggs presence in sample stool. The main demerit of this technique is it missed hookworm eggs presence in stool samples because the hookworm egg shell is very thin and it’s broken at the time to do concentration.
- **Kato_Katz Technique:** In this method a slide having stool samples are cleaned by glycerol for an hour, otherwise hookworm eggs are disappearing and disintegrate. The main drawback of this method is if more time is taken in examination of samples than it a much lower chance to detect hookworm eggs because hookworm eggs tend to disintegrate in long time delay.
- **McMaster Method:** In this method a counting chamber is used to examine a volume of fecal suspension microscopically. It is easier and performance comparable method then Kato-Katz technique, but its
main demerit is it required a special counting chamber for the examination.

- **FLOTAC Method**: In this method, apparatus is utilized for sample flotation in a centrifuge, after that apical portion of floating suspension is cut and eggs quantization is done. This method is successful in detecting helminthes and intestinal protozoa together. The main disadvantages of using this method need for a centrifuge, high cost, and require more time to prepare the sample. The comparison studies of all the above mentioned technique prove that the FLOTAC technique has high sensitivity and direct microscopy has less sensitivity in comparison to others. In this paper, section 2 is dedicated for literature review of past work, section 3 for Comparative Study and Finally Section 4 concludes this paper.

2. Literature review

[1] Presents a technique for the self-executing sectionalisation of Magnetic Resonance brain images into their respective tissue classes utilizing convolution neural network. To verify that the approach achieves accurate segmentation specifics as well as spatial consistency, the network operates multiple convolution kernel sizes and multiple patch sizes to achieve multi-scale data about each voxel. The method is not reliant on explicit features, but learns to perceive the information that is significant for the distribution based on training data.

[2] Explores trends concerning brain tumors. The primary focus of this research are gliomas they are the most aggressive and frequent, which leads to a very short life-span in their highest grade. Thus, treatment planning is an important hurdle to enhance the standard of life of oncological patients. MRI is a most prevalent imaging technique to detect these tumors', but the huge amount of data composed by MRI restrains manual segmentation in a plausible time, limiting the use of precise quantitative measurements in the clinical practice. So, automatic and dependable sectionalisation techniques are required; but, the large amount of variability, both spatial and structural in brain tumors makes automatic segmentation a formidable obstacle. In this paper, the authors present an automatic segmentation procedure established on CNN, exploring tiny 3x3 kernels. The utilization of compact kernels allows formulating a deeper architecture, in addition to having a positive impact against over fitting, because of the lower number of weights in the network.

[3] Presents a new technique for the self-regulating segmentation of MS lesions relied on deep convolutional encoder networks in conjunction with shortcut connections. The combined training of the prediction pathways and feature extraction allows for the learning of characteristic features at different extents automatically, that are tuned for a given combination of segmentation task and image types. Shortcuts between the two pathways allow low and high-level characteristics to be leveraged simultaneously for more consistent performance across scales. In conjunction, the authors have introduced a new objective function supported on the combination of sensitivity and specificity, which makes the objective function inherently robust to unbalanced classes like, for example, MS lesions, which frequently represent not more than 1%of all image voxel. The researchers have examined the method on a large data set from an MS clinical trial and, two publicly available data sets with the results demonstrating that the method performs comparatively to the best conventional methods, even for relatively small training set sizes.

[4] Presents systematic and robust method to automatically detect CMBs by influencing 3D CNNs from MR volumes. With the continuous accumulation of medical data, 3D CNN is an optimal solution for tasks from 3D volumetric data involving any detection and segmentation capabilities, as it is capable of representing high level features with rich spatial information of targets in a data directed way. However, the expensive computational cost of 3DCNN prohibits its use in clinical practice. We propose a two-stage framework to reduce its computational cost and improve the detection performance. By influencing an innovative 3D FCN strategy, the first stage retrieves a number of patients with higher prospects of being CMBs. Contrasted with conventional sliding window approach, the 3D FCN strategy disposes a large number of redundant convolutional computations, and hence dramatically speeds up the detection operation. In the second phase, a well-developed model is implemented on the candidates to distinguish CMBs from hard mimics. Experimental results established that the introduced design outperforms previous implementations by a large margin with higher detection sensitivity and fewer false positives. The proposed method can be easily adopted to other detection and segmentation tasks and boost the utilization of 3D CNNs on volumetric medical data.

[5] Introduces deep learning concepts sensitive to the local neighborhood for nucleus detection and classification in routine stained histology images of color ectaladenocarcinomas. Histopathological data are often insufficient and contaminated with inviduality, which is also the case for the dataset used in this particular study. This is inescapable due to the sheer number of cell nuclei and the enormous discrepancy in inviduality, which is also the case for the dataset used in this particular study. This is inescapable due to the sheer number of cell nuclei and the enormous discrepancy in morphology, making it difficult to identify all cells with certainty. Most pathologists only use high power cellular arrangement specifics to confirm or reject the initial response, but rely heavily on low power architecture to frame the main picture of what is being developed. Identifying individual cells on high power features alone without architectural clues will increase their misclassification. The interpretation was conducted on a large dataset with more than 20; 000 samples of different histological grades. The comparison is in favor of the presented spatially-constrained CNN for nucleus detection and the softmax CNN with the proposed neighboring ensemble predictor for nucleus classification. The combination of the two could potentially offer a systematic quantitative analysis of
tissue morphology, and tissue constituents, lending itself to be a useful tool for better perception of the microenvironment.

[6] Proposes experiments with convolution classification RBMs, which we trained with discriminative and generative learning objectives. Feature learning is generally done with a purely generative learning objective, which favors a representation that gives the most accurate description of the data but is inconsistent with the representation that is best for the goal of the system. This paper showed how the standard generative learning objective of an RBM can be combined with a discriminative learning objective. In our experiments, evaluating the categorization efficiency of random forests utilization RBM-learned features, the researchers found that a mixture of discriminative and generative learning objectives often gave a better categorization efficiency than generative or discriminative learning alone. The features learned with the mixed learning objective gave better results than several standard filter banks. The results suggest that adding discriminative learning is most useful when learning smaller representations, with fewer filters or hidden nodes.

[7] Explores a deep CNN to classify lung CT image patches into 7 classes, including 6 different ILD patterns and healthy tissue. A novel network architecture was designed that captures the low-level textural features of the lung tissue. The network is made up of 5 convolution layers with Leaky ReLU and activations 2x2 kernels, with size equivalent in comparison to the size of eventual, concluding feature maps and 3 dense layers, accompanied by just one average pooling. The training was performed by minimizing the categorical cross entropy with the Adam optimizer. The proposed approach gave promising results, outperforming the conventional methods on a very challenging dataset of 120 CT scans from different hospitals and scanners. The method can be easily trained on additional textural lung patterns while performance could be further improved by a more extensive investigation of the involved parameters. The huge number of parameters and the relatively slow training (typically a few hours) could be considered as a drawback of this kind of DL approaches, together with the slight fluctuation of the results, for the same input, due to the arbitrary initialization of the weights. In future studies, the authors plan to enhance the method to consider three-dimensional data from MDCT volume scans and finally to integrate it into a CAD system.

[8] Exploits three essentials, but hitherto understudied aspects of deploying deep convolution neural networks to computer-aided detection issues. The authors first examine and assess different CNN composition. The studied models consist of more than 150 million parameters, and fluctuate in numbers of layers. We then gauge the significance of dataset scale and spatial image context on performance. The authors studied two specific CADE problems, specifically ILD categorization and LN identification. The researchers achieved the conventional performance on the LN identification, report the first five times cross-validation categorization results on anticipating axial CT slices with ILD categories, with a maximum of 3 false alarms per patient at 85% sensitivity. Through their extensive empirical assessment, CNN model examination and valuable revelation can be broadened to the design of high-performance CAD systems for other medical imaging tasks.

[9] Introduces the helminths, the hookworms that live inside the intestines of humans as this negatively affects the health and is one of the leading causes of child and maternal mortality. There have been ways designed to diagnose such an ailment, but the Wireless Capsule Endoscopy still remains a demanding task. The researchers propose an innovative framework for detection of hookworms based on Deep Convolution Neural Networks. The technique identifies the hookworms in WCE images using two CNN networks, one for the visual appearance and the other designed to detect the tubular structure of the hookworms. Both the networks are seamlessly integrated to achieve the best possible performance. The proposed technique provides impressive performance which surpasses many traditional techniques in its efficiency and effectiveness.

[10] Explores the detection of hookworms in an individual's digestive tract with use of a Wireless Capsule Endoscopy. The researchers state that the WCE procedure is pain free and non-invasive but adds quite a lot of work for the technicians, as they have to examine each and every frame for the incidence of a hookworm. Therefore, the authors present an innovative technique for automatic recognition of the hookworm by detection of the bleeding regions encountered. The technique is a two-step process that utilizes colour images and application of K-means clustering technique on the pixels to extract the bleeding centres. The second stage generates a saliency map that is utilized to highlight the bleeding clusters. The method was tested to outperform the traditional techniques by a big margin.

[11] Presents a novel method for automatic detection of hookworm presence in an individual's digestive tract with the utilization of Wireless Capsule Endoscopy. The process of detection in WCE is difficult because of the complex design of the gastrointestinal tract, presence of foreign matters, poor quality if images and different appearances in terms of texture and color. The researchers utilize a multiscale dual matched filter to localize the tubular structure. Then, to those regions extracted from the previous step, a parallel detection method is applied piecewise. To reduce the affects of the variability of the gastrointestinal tract a histogram is produced with the classification of images done by Reboots to ameliorate the effects of imbalanced data; the proposed method achieves successful detection with high sensitivity.

[12] Proposes an innovative technique for detection and classification of images in the ImageNet Large-Scale Visual Recognition Challenge. The researchers designed the algorithm to perform the operation on the network, by using the computational resources of the network itself. To achieve this level of computation from the limited resources the network has to offer, the authors deepened and widened the network to keep...
the computational budget under control. For the optimization of
this heavily budgeted framework, the researchers based their architectural designs on the Hebbian principle. The proposed framework was tested on the 22 layered networks, Google Net and achieved promising results.

[13] Explores the realm of image recognition, more specifically edge detections and aims to address the predominant issues, such as, multi-scale/ multi-level feature learning and holistic image training procedure. The researchers propose an innovative technique to ameliorate the issues encountered by HED holistically- nested edge detection. HED implements deep learning on an image-to-image prediction that is completely reliant on deeply supervised nets and Convolution neural networks. The learning networks ingrained in the method allow it to learn and resolve the ambiguity it encounters in boundary detection and edge recognition. This technique has been demonstrated to significantly outperform traditional techniques.

[14] Proposes an improved bag of feature (BoF) method to assist classification of polyps in WCE images. Instead of utilizing a single scale-invariant feature transform (SIFT) feature in the traditional BoF method, they extract different textural features from the neighborhoods of the key points and integrate them together as synthetic descriptors to carry out classification tasks. Specifically, researchers study influence of the number of visual words, the patch size and different classification methods in terms of classification performance. The researchers test the proposed improved BoF model with two classifiers: SVM and FLDA. SVM is a relatively new machine learning method based on the foundation of statistical Learning theory. The best performance of polyp classification could be obtained with the visual words set of 120, vocabulary set size set of 8*8, integrated feature of SIFT+CLBP and the SVM method. The proposed scheme shows an outstanding performance compared with the traditional textural methods.

[15] Presents a novel coding method, namely, saliency and adaptive locality-constrained linear coding (SALLC) for WCE image classification, which considers the saliency and local information about patch feature simultaneously. Different from the existing approaches that treat each local feature equally; this encoding method incorporates the saliency constraints to emphasize the important features of the images. Instead of calculating the response of each descriptor to a fixed number of codewords, the proposed SALLC strategy projects each descriptor into an adaptive determined number of coding bases, which is calculated based on the distances between the feature and the visual words. In addition, this estimation of the adaptive coding bases includes only one tuning parameter, but it can provide better characterization of the images. The proposed method can be applied to other image classification problems as well to improve the corresponding performance.

[16] Suggests neural network classifier technique to identify the potential region having hookworm bodies. NNs can be used to model complex relationships between inputs and outputs or to find patterns in data, thereby making it suitable for hookworm segmentation. Neural networks are employed in this method to classify the hookworm tissues and non-hookworm regions of the liver. The performance of the suggested system is analyzed in terms of sensitivity, specificity and accuracy. The suggested hookworm detection technique achieves 83% of sensitivity, 98% of specificity and 99% of accuracy.

[17] Presents an approach for automated segmentation of blood regions in WCE images via a deep learning strategy. The proposed method first classify the bleeding samples into active and inactive subgroups based on the statistical features derived from the histogram probability of the color space. Then for each subgroup, they highlight the blood regions via fully convolutional networks (FCNs). Experimental results on the clinical WCE dataset demonstrate the efficacy of approach, which achieves comparable or better performance than the state-of-the-art methods.

[18] proposes a robust intestinal hemorrhage detection model via Convolutional Neural Networks(CNNs). They Explored different CNN architectures, including LeNet, AlexNet, GoogLeNet, and VGGNet. These architectures are different in the input format, depth, and modules. In addition, to improve the robustness of the learned model, they augmented the data with various transformations, including rotation, blurring, luminance change, and Poison noise.

3. Conclusion

Due to ease increasing unhygienic in the surrounding that leads to the growth of the dirt which are often contaminated by the feces. Because of these hookworms are entered in human body and creates the vulnerable diseases, So most of the times endoscopy is required to identify the hookworms. This research paper analyzes most of the past work to know still there is a lot of work need to do in the hook detection process. So to make this process easier proposed model analyzed all the past works and put forward an idea of detecting hookworms using convolution neural network with the help of the fuzzy classification theory.

References


