International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 06 Issue 08 August 2023
DOI: 10.47191/ijcsrr/V6-i8-74, Impact Factor: 6.789
IJCSRR @ 2023

Detection and Classification of Gastrointestinal Diseases by using Machine Learning: A Review

Salah Mashal Raheem[1], Ziad Mohammed Abood[2]
[1,2] Department of Physics, College of Education, University of Al-Mustansiriyah, Baghdad, Iraq

ABSTRACT: Currently, gastrointestinal diseases claim the lives of up to two million people worldwide. GI disease treatment can be challenging, time-consuming, and expensive. One of the most recent advancements in medical imaging is the use of video endoscopy to diagnose gastrointestinal illnesses such stomach ulcers, bleeding, and polyps. Doctors require a lot of time to review all the images produced by medical video endoscopy since there are so many of them. This makes manual diagnosis difficult and has encouraged research into computer-aided approaches to diagnose all of the generated images quickly and accurately. The innovative aspect of the suggested methodology is the creation of a system for the diagnosis of digestive disorders. Machine learning techniques have the potential to significantly lower the cost of examination procedures while increasing the accuracy and speed of diagnosis. This paper describes a method for classifying GI illnesses using machine learning techniques.

KEY WORDS: Artificial Intelligent, Classification, Gastrointestinal diseases, Machine Learning.

1. INTRODUCTION

One of the most prevalent diseases in humans, gastrointestinal (GI) disease causes one of the most significant healthcare issues. It can be classified as benign GI disorders, precancerous lesions, early GI cancer, and advanced GI cancer depending on the degree of the lesion. In the short term, benign GI conditions including ulcers, gastritis, and bleedings won't turn into cancer. If not identified and treated in a timely manner, precancerous GI lesions may progress into early GI cancer or even advanced GI cancer [1].Detection and recognition of gastrointestinal (GI) tract infections in medical image processing are currently active research areas [2].

In the majority of cases, gastrointestinal polyps are thought to be the early stages of cancer development. Thus, early polyp discovery and removal can lessen the likelihood of malignancy. In actuality, stomach and colonic mucosal polyps are aberrant growths of tissue. The majority of the time, before they reach a considerable size, this growth is a leisurely process without any symptoms. However, if polyps could be detected early on, cancer may be prevented and treated [3].

As the third most common cause of cancer-related fatalities globally, gastric cancer is typically detected at an advanced stage and is not amenable to curative resection [4]. The 5-year survival rate of this cancer, however, surpasses 95% if (GC) can be detected and subsequently treated effectively at an early stage [5].

A routine method for checking a patient's gastrointestinal (GI) tract for potential problems, such as cancer, ulcers, and bleeding, is a video endoscopy [6]. The examination of the digestive system (through an endoscope) is a time-consuming and laborious task for medical professionals. In actuality, video endoscopy produces a substantial amount of frames, which are then carefully reviewed by a gastroenterologist [7].

When screening a large number of patients, an automated vision-based system can be useful in finding cancerous frames throughout all video frames, saving time for the medical professionals [8]. In a computer-aided diagnosis system, the extraction of significant characteristics from medical images is a crucial step for the detection of anomalies.[9]

Currently, video endoscopy is a crucial tool for the diagnosis and treatment of various diseases; however, when performing endoscopic examinations, some lesions are not visible to the naked eye; therefore, when observed in specific shades of color using Color filters, the image provides valuable information to medical specialists in endoscopy for the more accurate detection of lesions and diseases of the internal cavities. [10].

There are video endoscopy systems that identify and display the images that are in focus using approaches based on edges and clusters, enhancing the conditions for the specialist to make better decisions. By utilizing chromatic filters, those are able to display images in various color shades, providing endoscopy specialists with invaluable information for the detection of tumors and disorders in internal cavities [11]. Currently, upper digestive endoscopy (UDE) is the most crucial operation to assess and diagnose

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 06 Issue 08 August 2023
DOI: 10.47191/ijcsrr/V6-i8-74, Impact Factor: 6.789
IJCSRR @ 2023

stomach cancer and take biopsy samples. It is a vital tool for the identification and treatment of various disorders. Novice endoscopists may not be able to accurately evaluate the whole mucosa during UDE. Blind spots are the hidden regions that could conceal future neoplastic lesions. However, the capable AI system could be able to get beyond this obstacle and perhaps improve the quality of UDE [12].

Despite the fact that there are few or no studies in Iraq focused on using machine learning (ML) to diagnose gastrointestinal disorders, these conditions constitute a serious hazard to human health. For instance, according to [13], the intestinal metaplasia and mucosal atrophy brought on by the helicobacter pylori (HP) infection of the gastric mucosa both enhance the risk of gastric cancer. Therefore, the purpose of this review was to evaluate the whether the performance of a ML-based AI for detecting early gastric diseases is better than that of endoscopists based on endoscopic video images.

2. COMPUTER – AIDED DETECTION / DIAGNOSE

The computer-based approach known as computer-aided detection (CADe) or computer-aided diagnosis (CADx) in the field of medical imaging aids doctors in making judgments quickly [14, 15]. Medical imaging is concerned with information in images that doctors and medical professionals must quickly assess and analyze for irregularity. In the medical field, imaging analysis is a very important activity because imaging is a fundamental tool for early disease diagnosis, but image acquisition does not harm the human body. Imaging methods like MRI, X-ray, endoscope, ultrasound, etc., when obtained with high energy will produce good-quality images but they will hurt the human body; as a result, images are taken with low energy, which results in images with poor quality and little contrast. CAD systems are used to enhance the image quality, which aids in appropriately interpreting medical imaging and processing the images to highlight the portions that stand out [16]. A variety of principles from artificial intelligence (AI), computer vision, and medical image processing are all included in the technology known as CAD. Finding abnormalities in the human body is the primary application of CAD technology. The most common use of these is tumor detection because, if missed during routine screening, it can result in cancer [17].

One of the most significant issues in the field of image recognition is the classification of medical images into several groups in order to assist physicians in the diagnosis of disease or in further research. In general, there are two processes in the classification of medical images. Extraction of useful features from the image is the initial stage. The next stage is to create classification models for the dataset of images using the features. [18].

The identification of GI disease and the fast, effective, and secure excision of lesions can be aided by the automatic classification of GI diseases utilizing CAD [19]. Four stages may be distinguished in the combined CAD system's operation during a clinical image examination: image pre-processing, segmentation, feature extraction / selection, and classification of lesions.

Computer-aided diagnostic (CAD) systems that automatically select, identify, and classify lesion images and give doctors an unaltered point of reference have been created to increase efficiency and diagnosis accuracy. These CAD technologies could not only reduce the load on physicians but also increase the effectiveness of diagnostic procedures [20].

3. RELATED WORKS

Bchir et al. (2016) [21] assessed the effectiveness of nine visual features for ulcer detection in WCE video frames, including local binary patterns, CIE lab color histograms, curvelet transforms, chromaticity moments color, scalable color descriptors, color coherence vectors, homogeneous texture descriptors, YCbCr color histograms, and HSV (hue, saturation, and intensity value) color histograms. Utilizing SVM, they reported 96% accuracy.

Automated feature extraction was performed by Jia and Meng (2016) [22] using a CNN architecture, and the retrieved features were subsequently utilized to train an SVM to recognize inflammatory gastrointestinal illness in WCE videos. Accuracy levels in this study reached up to 90%.

Wimmer et al. (2016) [23] trained a CNN using various filter dimensions, layers, and layer counts to identify celiac disease, They united SVM and CNN. And the performance that was most successful was 97%.

A color feature-based automated method of identifying bleeding from WCE frames was proposed by Suman et al. in 2017 [24].The system being discussed involves use of statistical color feature analysis and an SVM classifier for classification. The test results show that the proposed methodology is beneficial since it provides higher accuracy when compared to current methods. Combining the selection of statistical color features with the 97.67% accurate SVM of the RBF kernel function.

The effectiveness of various visual descriptors for the detection of ulcers utilizing WCE (Wireless Capsule Endoscopy) frames was compared by Bchir et al. in 2018 [25]. The goal of this research was to identify which visual description more accurately detects gastrointestinal ulcers and portrays WCE frames. SVM classifiers were combined with a variety of visual descriptors during the studies. The authors used an LBP descriptor together with an SVM classifier to reach a maximum accuracy of 98.85%.

In training, Kanesaka et al. (2018) [26] identified 66 abnormal conditions out of 126 images, and in testing, they identified 61 abnormalities out of 81 images. Additionally, the study used a Support Vector Machine (SVM) with 96.3% accuracy, 96.7% sensitivity, and 95% specificity to identify early gastric cancer. Comparing this outcome to works that used CNN, which typically show between 85% and 93.8% accuracy, is pretty impressive.

Takiyama et al. (2018) [27] looked at stomach cancer, In order to identify the anatomic multi-location of the disease, Location-specific accuracy for stomach cancer detection was determined to be 90% and Area Under Curve (AUC) of 0.99, a sensitivity of 96.9%, and a specificity of 98.5%.

Xing et al. (2018) [28] proposed automatic system for bleeding frame identification and area segmentation when using of the super pixel color histogram feature and a subspace K-nearest neighbor classifier. 99% accuracy using a subspace KNN-based technique and the super pixel histogram.

A novel method based on the fusion of deep CNN and geometrical functions was proposed by Sharif et al. in 2019[29]. Using a novel technique called improved contrast color features, the illness regions are initially retrieved from the supplied WCE images. The segmented illness region served as the source for the geometric features. The deep CNN functions VGG16 and VGG19 were then uniquely combined based on the Euclidean Fisher Vector. Utilizing the conditional entropy technique, the best features are selected by combining the geometric and unique features. The chosen features were categorized using K-Nearest Neighbor (kNN). A privately produced collection of 5500 WCE photos was utilized to evaluate the recommended approach, and the results showed classification accuracy of 99.42% and precision rate of 99.51%. However, the authors were only able to classify the data into three categories: ulcers, bleeding, and health.

Souaidi et al. (2019)[30] developed a method for identifying ulcers using many scales. The full pyramids of LBP and Laplacian, for example, were retrieved by the authors and then categorized using SVM. On two WCE datasets, they tested the system, and they discovered that it had accuracy levels of 95.11% and 93.88% respectively.

4. ARTIFICIAL INTELLIGENT

Artificial intelligence (AI) algorithms began to develop as a result of the constant advancements in information technology and their effects on every aspect of our lives. This was due to the demand for better machine performance. Human brain performance can be impacted by weariness, stress, and lack of experience, unlike that of computers.

Artificial intelligence (AI) technology would make up for human limitations, prevent human errors, allow machines some trustworthy autonomy, and boost productivity and efficiency at work. Therefore, AI may be the greatest choice when looking for a quick and trustworthy assistant to handle the steadily increasing number of patients. Gastrointestinal (GI) endoscopy could benefit greatly from the use of AI technology. It can lower inter-operator variability, improve diagnosis accuracy, and support making prompt, accurate therapy decisions. AI would also speed up, lower the cost, and simplify endoscopic treatments.[31] The classification of gastrointestinal disorders has been the subject of numerous studies, all of which aim to significantly advance the accurate diagnosis and treatment of gastrointestinal disorders. The bulk of research studies, however, sought to categorize a very small number of gastrointestinal illnesses in a particular area of the human gastrointestinal system. Several studies, thus, utilized limited data [32].

Artificial intelligence (AI) is fundamentally a technique that uses mathematics and statistics to learn human thought patterns and transmit human experience. The three main pillars of AI are algorithm iteration, increasing data, and advancing computing power. AI is a subset of machine learning (ML)[33], and deep learning is a subset of ML to actualize ML[34], where several algorithms are organized together in intricate layers. Computer-aided diagnostic (CAD) systems use deep learning and machine learning to evaluate medical images[35].

5. MACHINE LEARNING

A branch of artificial intelligence known as machine learning (ML) automates decision-making with little to no human involvement by learning from data by spotting patterns [36]. The ability of a machine learning (ML) model to independently adapt, learn from prior calculations, and generate accurate results when repeatedly exposed to fresh datasets is its most crucial quality. The two key elements are (i) utilizing Computer Aided Design (CAD) and Machine Learning (ML) approaches, doctors can quickly evaluate medical images. and (ii) algorithms used for difficult jobs like CT scan with segmentation [37].

Traditional ML models operated on structured datasets with predetermined methodologies for each step; if a step was skipped, the model fails. It is crucial for ML and DL algorithms to evaluate the quality of the data they employ [38]. New algorithms, however, adjust data omission based on the algorithm's need for robustness [39].
As effective classifier and grouping algorithms, machine learning techniques are frequently applied in the field of medical imaging research [40]. Support vector machines and clustering techniques like k-nearest neighbor (k-NN) are the best classifiers currently being employed [41].

Artificial Intelligence (AI) and machine learning (ML) have advanced significantly in recent years. In the fields of medical image processing, computer-aided diagnosis, image interpretation, image fusion, image registration, image segmentation, image retrieval, and image analysis, machine learning and artificial intelligence have played significant roles.

ML collects information from the images and efficiently and effectively represents the information. Together, ML and AI can detect diseases, estimate their risks accurately and quickly, and take timely preventative measures. These methods help medical professionals and researchers better understand how to identify the common changes that cause disease. These methods consist of traditional, non-learning algorithms like Support Vector Machines (SVM), Neural Networks (NN), and KNN, among others.

5.1 Machine Learning Applications

Machine learning systems are not like thermometers, which accurately measure temperature using physical laws that apply to all objects, nor are they like skilled physicians, who can elegantly adjust to changing conditions. Instead, these systems should be seen as a set of guidelines that can function properly in one Center but completely fall apart in another. These guidelines were developed to operate in specific situations and rely on specific assumptions[42]. As a result, ML infiltrated numerous life-related joints to fulfill goals quickly and precisely, as shown in figure (1).

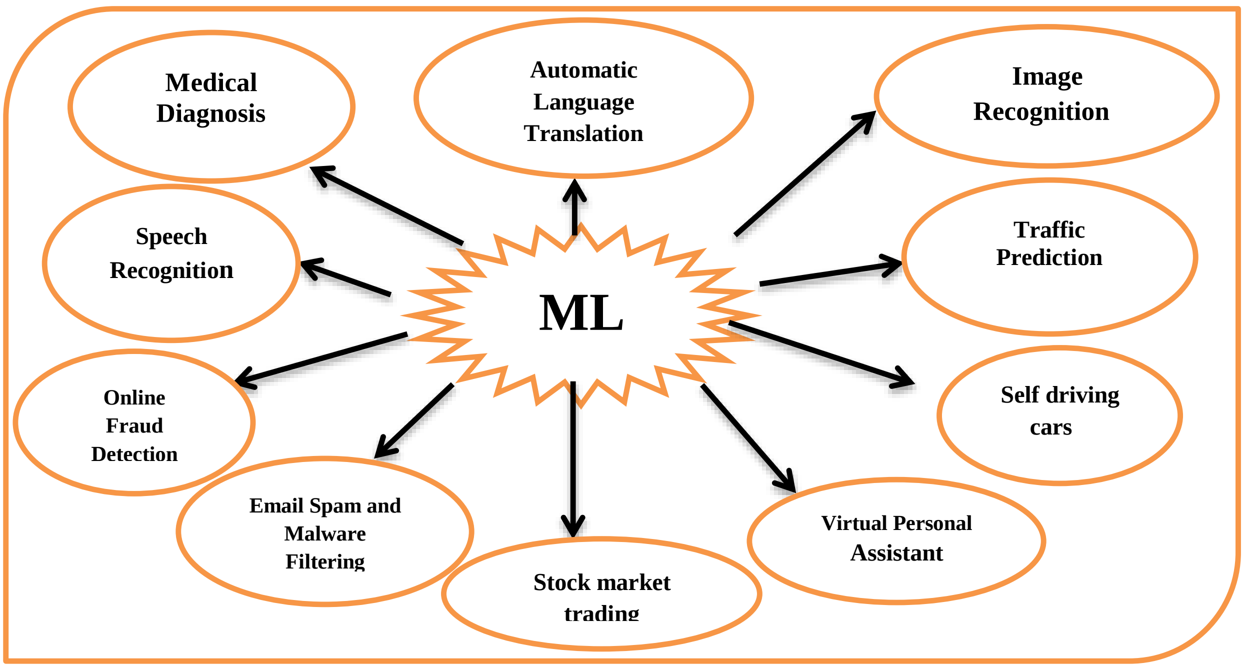

Figure (1): Application of Machine Learning

International Journal of Current Science Research and Review
ISSN: 2581-8341
Volume 06 Issue 08 August 2023
DOI: 10.47191/ijcsrr/V6-i8-74, Impact Factor: 6.789
IJCSRR @ 2023

There are numerous applications in pharmaceuticals and medicine for the use of machine learning in the medical industry, including:
1. Recognizing Illnesses and Making Diagnoses
2. Drug Development and Production
3. Diagnostic Medical Imaging
4. Individualized Healthcare
5. Behavioral modification based on machine learning
6. Smart Health Records
7. Clinical trials and research
8. Data collection from the general public
9. Improved Radiotherapy
10. Forecast for an outbreak

5.2 Machine Learning types

With the incorporation of artificial intelligence, more notably machine learning, information technology improvements in healthcare have significantly impacted these developments. Machine learning algorithms will be used more frequently in the medical area due to the ongoing growth and complexity of data [43].

The study of algorithms and statistical models that are utilized to carry out a certain task without explicit instructions is known as machine learning [44]. A set of data with inputs and the desired output is used by machine learning algorithms to create a mathematical model [45]. Machine learning models can be categorized into one of three primary groups fall under Figure (2).

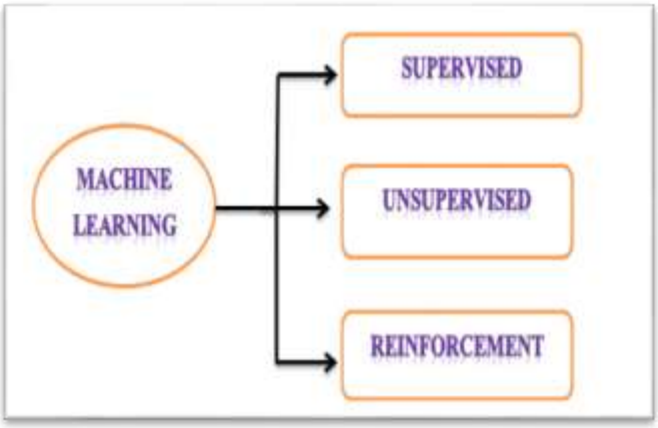

Figure (2): Machine Learning Types

1. **Supervised learning**

The training data used by conventional supervised learning algorithms contains labels that indicate the desired output for each corresponding input. The field's most widespread and well-established technique of education is this one. Classification and regression are two categories of supervised learning techniques[46].

Choosing a random sample of the data, selecting an algorithm, training the model, and then assessing the model produced are some common techniques [47]. There are a variety of algorithms from which to pick. One approach is to use a variety of pertinent algorithms to train the model, and then use the confusion matrix and receiver operating curve (ROC) to assess how well it performed [48]. A final model is created by an iterative procedure. To provide the model the most amount of predictive power, the optimum method with the combination of parameters is chosen.16 To reduce the amount of time and computational effort required for hyper parameter tweaking, some methodology articles have provided default values for certain parameters [49].

2. **Unsupervised Learning**

Unsupervised learning refers to a machine learning procedure that can forecast knowledge. There are no predetermined target values because there are no real data provided. The phrase "learning without a teacher" is another name for this method. Clustering is a well-known use of unsupervised learning. In order to classify the input values according to frequent patterns, clustering looks for similarities in the input values [50].

3. Reinforcement Learning

A kind of machine learning and a real-world application is reinforcement learning. To get at some decisions, a series of models is created. The system acts and then receives input in response to that activity. Based on the input, the decisions and subsequent actions are updated. This artificially intelligent system begins off by making mistakes, and it then learns from the input it receives. The feedback is taken into account as subsequent judgments and actions are changed until the system responds favorably. When many machine learning algorithms are creatively integrated and arranged to make various judgments, receive regular feedback, adjust the decisions, and experiment on various decision spaces until the final decisions are optimal based on the results [51]. The decision-making process in healthcare is not linear, thus a variety of considerations must be taken into account. A decision support system can be created using reinforcement learning to offer treatment suggestions to the doctors. Adopting reinforcement learning in the healthcare industry presents a number of difficulties, including assessing the decision made by the system and selecting the reward to change the decision and action. Understanding disease dynamics and contextually establishing the causal links between the pertinent factor and the outcome are also necessary for this approach [52].

6. CLASSIFICATION

The difficulty of identifying items in a group into various categories is known as classification. The classification objective is to correctly predict the target class for each sample of the data. In machine learning, classification has been regarded as one of the examples of supervised learning, which is learning where a training set of observations that have been correctly identified is available. Two steps are used to illustrate it: the Learning step and the Classification step. A variety of methods have been used to develop the classifier through learning the training set that is accessible in the learning step (training phase), which is where the construction of the classification model is done. In order for the model to correctly anticipate outcomes, it must be trained. To categorize the testing data and determine the classification rule precision, which is based on labels for specified classes, a model is utilized in the classification stage [53].

Each component of the dataset is classified using a machine learning model into one of the specified classifications [54]. These classifiers operate concurrently and create several models using the training dataset [55]. Classification methods are tested on the dataset for the classification of gastrointestinal illnesses.Three different categorization algorithm types are highlighted in this research.

6.1 Support Vector Machine

Every programmer and machine learning specialist should be knowledgeable about the Support Vector Machine (SVM), a significant, uncomplicated supervised learning technique. It can be applied to tasks involving classification and regression. The Support Vector Machine technique seeks out a hyper plane in an N-dimensional space that clearly divides data points, but it is primarily employed for classification reasons. [56].

Based on statistical learning theory, the Support Vector Machines classifier (SVM) is a learning model used to separate two classes. SVM addresses classification, learning, and prediction issues [57]. Finding the hyper-plane that provides the greatest degree of separation between the two classes is the main classification approach employed by SVM. In SVM, the hyper-plane is often constructed using a set of data called the training dataset, and its generalizability is verified using a separate subset known as the testing dataset [58].

Support vector machines are used to find a hyperplane line that most evenly splits the positive and negative data [59]. The support vector machine (SVM) seeks a hyperplane in the Multidimensional Features space that efficiently classifies data points [60].

There are two types of separation methods: linear separation and non-linear separation.

a- Linear separation or linear SVMs

The simplest SVM instances are those where the classes can be separated linearly. In these cases, all that has to be done is to maximize the classifier margin in order to choose a suitable separating hyperplane .

b- Nonlinear separation or nonlinear SVMs:

Since classes in real SVM applications cannot be separated linearly, we are using nonlinear SVM to get around this issue. Specifically, we are applying a nonlinear transformation to the data to change dimensions and quickly find a hyperplane classification in this new space. We are also giving the classifier more latitude to correctly classify the points even if they are initially points on the incorrect side of the initial hyperplane (nonseparable categories).

An SVM's generalizability and convergence are key factors in its success [61]. Numerous additional effective classification methods exist in addition to SVMs, including the KNN algorithm [62, 63], Bayesian networks [64, 65], Random Forest, artificial neural networks [66, 67], and decision trees [68].

6.2 Random Forest

The Random Forest Classifier (RF) may classify and analyze data in a variety of ways. It also provides a method for missing value estimation. It is also resistant to noise and training data reduction. A multi-decision tree is created by a community classifier known as a (RF) employing a number of different variables [69]. According to Reis et al. (2018)[70], Random Forest is a classifier that consists of a number of classifiers with a tree-like structure, identically distributed independent random vectors, and each tree that casts a unit vote for the most popular class at input x. An upper bound is extracted for Random Forests to get the generalization error in terms of two parameters, Exactitude and interdependence of individual classifiers. This generalization error is expressed in terms of two parameters, Exactitude and interdependence of individual classifiers (71).

6.3 K-Nearest Neighbor’s Algorithm (KNN)

KNN is one of many machine-learning methods that are classified as non-parametric, which means that the dataset determines the model's structure and the sample data distribution has no presumptions. One of the most straightforward and basic data mining approaches is KNN, often known as memory-based classification because the training samples must be kept in memory while the algorithm is running [72].

KNN's goals are to analyze a dataset that categorizes vectors into distinct classes and to identify the categorization of fresh data points by applying prediction techniques. KNN employs a database of data points divided into several classes with the goal of predicting the classification of a fresh sample point [73], defying conventional theoretical presumptions (as with linear regression models).

The KNN algorithm, which is utilized for classification and is part of the supervised machine learning approach [74], is one algorithm that can be used for prediction. The KNN algorithm operates by classifying the attributes that are closest to each other in the training set [75].

The ease of implementation and interpretation of this straightforward prediction method is well known [76]. Setting numerous parameters is not necessary for this model. The more predictor variables there are, the slower this method becomes. The outcome is categorized in accordance with the majority decision of the close-by data points [77].

CONCLUSION

Modern machine learning techniques are quite resilient to real-world situations, and the learning process actually benefits from the forced removal of some data. Machine learning systems will do jobs that were once thought to be exclusive to humans due to the rapid rate of technological breakthroughs. Endoscopy is currently utilizing machine learning, and in the near future, these applications are likely to expand quickly. In order to ensure that patients receive the best care possible, it is crucial that we engage in this field of study because the application of machine learning in endoscopy has significant implications for the practice of medicine. Understanding the characteristics of machine learning technologies is essential to guaranteeing their safest and most efficient application.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018; 68: 394–424.
2. Khan, M. A., Sharif, M., Akram, T., Yasmin, M., & Nayak, R. S. (2019). Stomach deformities recognition using rank-based deep features selection. Journal of Medical Systems.

3. Billah,M., Waheed,S and Rahman,M.M. (2017). An Automatic Gastrointestinal Polyp Detection System in Video Endoscopy Using Fusion of Color Wavelet and Convolutional Neural Network Features. International Journal of Biomedical Imaging Volume 2017, P. 9.
4. Itoh, T., Kawahira, H., Nakashima, H. and Yata, N. ``Deep learning analyzes Helicobacter pylori infection by upper gastrointestinal endoscopy images," Endosc. Int. Open, vol. 6, no. 2, pp. E139_E144, 2018.
5. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. CA Cancer J Clin (2020) 70(1):7–30.
6. Van Cutsem E, Sagaert X, Topal B, Haustermans K, Prenen H. Gastric cancer. Lancet (2016) 388(10060):2654–64.
7. Sergio Coda, An Investigation of the Diagnostic Potential of AutofluorescenceLifetime Spectroscopy and Imaging for Label-Free Contrast of Disease. PhDthesis, Thesis submitted for the award of Doctor of Philosophy (PhD) ImperialCollege of Science, Technology and Medicine, 2014.
8. Ghosh, S.A. Fattah, C. Shahnaz, A.K. Kundu, M.N. Rizve, Block basedhistogram feature extraction method for bleeding detection in wirelesscapsule endoscopy, IEEE Region 10 Conference TENCON. IEEE (2015) 1–4.
9. Shuai Wang, Yang Cong, Jun Cao, Yunsheng Yang, Yandong Tang, Huaici Zhao,Haibin Yu, Scalable gastroscopic video summarization via similar-inhibitiondictionary selection, Artif. Intel. Med. 66 (2016) 1–13.
10. Horie Y, Yoshio T, Aoyama K et al. The diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. Gastrointest. Endosc. 2018 .
11. Lai, K., Bo, L., Ren, X., & Fox, D. (2011). A largescale hierarchical multi-view RGB-D object dataset. In Proceedings - IEEE International Conference on Robotics and Automation (pp. 1817–1824).
12. Fuller, D., Buote, R. and Stanley, K. (2021). A glossary for big data in population and public health: discussion and commentary on terminology and research methods | Journal of Epidemiology & Community Health. Accessed December 18, 2021.
13. X. Liu, C. Wang, Y. Hu, Z. Zeng, J. Bai, and G. Liao, ``Transfer learning with convolutional neural network for early gastric cancer classification on magnifying narrow-band imaging images," in Proc. Int. Conf. Image Process. (ICIP), Athens, Greece, Oct. 2018, pp. 1388_1392.
14. Doi K. Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. Computerized Medical Imaging and Graphics. 2017;31(4):198-211
15. [15] Li Q, Nishikawa RM, editors. Computer-Aided Detection and Diagnosis in Medical Imaging. Taylor & Francis, CRC Press, New York; 2015
16. [16] Chen C-M, Chen, Yi-Hong Chou, Norio Tagawa, and Younghae Do. Computer-aided detection and diagnosis in medical imaging. Computational and Mathematical Methods in Medicine. vol. 2013, Article ID 790608, 2 pages, 2013.
17. Giger ML, et al. Computer-aided diagnosis in mammography. Handbook of Medical Imaging. 2nd ed. SPIE Digital Library, Europe, 2000. 915-1004
18. E. Ahn, A. Kumar, J. Kim, C. Li, D. Feng, and M. Fulham, "Xray image classification using domain transferred convolutional neural networks and local sparse spatial pyramid," in roceedings of 2016 IEEE International Symposium on Biomedical Imaging (ISBI), pp. 855–858, IEEE, Prague, Czech Republic, April 2016.
19. Hwang, J.H.; Jamidar, P.; Baig, K.R.K.K.; Leung, F.W.; Lightdale, J.R.; Maranki, J.L.; Okolo III, P.I.; Swanstrom, L.L.; Chak, A. GIE Editorial Board Top 10 Topics: Advances in GI Endoscopy in 2019. Gastrointest. Endosc. 2020, 92, 241–251.
20. Du, W., Rao, N., Liu, D., Jiang H., Luo, C., Li, Z., Gan, T.,and Zeng, B. (2019). Review on the Applications of Deep Learning in the Analysis of Gastrointestinal Endoscopy Images. Digital Object Identifier, IEEE, 7: 142053.
21. Georgakopoulos, S.V.; Iakovidis, D.; Vasilakakis, M.; Plagianakos, V.P.; Koulaouzidis, A. Weakly-Supervised Convolutional Learning for Detection of Inflammatory Gastrointestinal Lesions. In Proceedings of the IEEE International Conference on Imaging Systems and Techniques (IST), Chania, Crete Island, Greece, 4–6 October 2016; pp. 510–514.
22. Jia, X.; Meng, M. A Deep Convolutional Neural Network for Bleeding Detection in Wireless Capsule Endoscopy Images. In Proceedings of the IEEE 38th Annual International Conference on the Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 16–20 August 2016; pp. 639–642.

23. Wimmer, G.; Hegenbart, S.; Vécsei, A.; Uhl, A. Convolutional Neural Network Architectures for the Automated Diagnosis of Celiac Disease. In Proceedings of the InternationalWorkshop on Computer-Assisted and Robotic Endoscopy, Munich, Germany, 5 October 2016; pp. 104–113.
24. S. Suman, F. A. B. Hussin, A. S. Malik, K. Pogorelc, M. Riegler, S. H. Ho, I. Hilmi, and K. L. Goh, "Detection and classification of bleeding region in wce images using color feature," in Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing –CBMI' 17 , 2017, pp. 16.
25. Bchir, O.; Ismail, M.; AL_Aseem, N. Empirical Comparison of Visual Descriptors for Ulcer Recognition in Wireless Capsule Endoscopy Video. In Proceedings of the 4th International Conference on Image Processing and Pattern Recognition (IPPR 2018), Copenhagen, Denmark, 28–29 April 2018.
26. Kanesaka T., Lee T.-C., Uedo N., Lin K.P., Chen H.Z., Lee J.Y., Wang H.-P., Chang H.T. Computer-aided diagnosis for identifying and delineating early gastric cancers in magnifying narrow-band imaging. *Gastrointest Endosc.* 2018;87:1339–1344.
27. Takiyama, H., Ozawa, T., Ishihara, S., Fujishiro, M., Shichijo, S., Nomura, S., Miura, M., & Tada, T. (2018). Automatic anatomical classification of esophagogastrodenoscopy images using deep convolutional neural networks. *Scientific Reports*, *8*(1), 1–8.
28. Xing, X., Jia, X., and Meng,M. (2018). Bleeding Detection in Wireless Capsule Endoscopy Image Video Using Superpixel-Color Histogram and a Subspace KNN Classifier. 978-1-5386-3646.
29. Sharif, M., Attique Khan, M., Rashid, M., Yasmin, M., Afza, F., & Tanik, U. J. (2019). Deep CNN and geometric features-based gastrointestinal tract diseases detection and classification from wireless capsule endoscopy images. *Journal of Experimental and Theoretical Artificial Intelligence*, 1–23.
30. Souaidi, M., Ait, A., & El Ansari, M. (2019). Multi-scale completed local binary patterns for ulcer detection in wireless capsule endoscopy images. *Multimedia Tools and Applications*, *78*(10), 13091–13108.
31. El Hajjar, A. and Rey, J. (2020). Artificial intelligence in gastrointestinal endoscopy: general overview. Chinese Medical Journal;133(3):326.
32. Borgli, H., Vajira Thambawita, V., Smedsrud, P. H., Hicks, S., Jha, D., Sigrun L. Eskeland,S. L., Randel, K. R., Pogorelov, K., Lux, M., Nguyen,D. T. D., Johansen, D., Griwodz, C., Stensland, H. K., Garcia-Ceja, E., Schmidt, P. T., Hammer, H. L., Riegler, M. A., Halvorsen, P., Thomas de Lange. (2020). HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy. *Scientific Data*, *7*(1), 1–14.
33. Bi Q, Goodman KE, Kaminsky J, Lessler J. What is Machine Learning? *Am J Epidemiol* 2019; 188: 2222-2239 .
34. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015; 521: 436-444.
35. Ahmad OF, Soares AS, Mazomenos E, Brandao P, Vega R, Seward E, Stoyanov D, Chand M, Lovat LB. Artificial intelligence and computer-aided diagnosis in colonoscopy: current evidence and future directions. *Lancet Gastroenterol Hepatol* 2019; **4**: 71-80
36. Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G … Zhao S (2019) Applications of machine learning in drug discovery and development. Nat Rev Drug Discov 18(6):463–477.
37. Schoepf UJ, Zwerner PL, Savino G, Herzog C, Kerl JM, Costello P (2007) Coronary CT angiography. Radiology 244(1):48–63.
38. Bhushan M, Duarte JÁG, Samant P, Kumar A, Negi A (2021) Classifying and resolving software product line redundancies using an ontological first-order logic rule based method. Expert Syst Appl 168:114167.
39. Negi A, Kaur K (2017) Method to resolve software product line errors. In: International conference on information, communication and computing technology. Springer, Singapore, pp 258–268 .
40. Wang S, Summers RM. Machine learning and radiology. Med Image Anal 2012;16:933–51.
41. J. Ker, L. Wang, J. Rao and T. Lim, "Deep Learning Application in Medical Image Analysis," in IEEE Access, vol. 6, pp. 9375-9389, 2018.
42. Futoma, J., Simons, M., Panch, T., Doshi-Velez, F., Celi, L.A. (2020) The myth of generalisability in clinical research and machine learning in health care . Lancet Digital Health 2020; 2: e489–92

43. T. Davenport and R. Kalakota, "The potential for artificial intelligence in healthcare," Future Healthcare Journal, vol. 6, no. 2, pp. 94–98, 2019.
44. D. Cielen, A. D. Meysman, and M. Ali, Introducing Data Science. Introducing Data Science, Manning Publications, 2016.
45. C. M. Bishop, Pattern Recognition and Machine Learning. Information Science and Statistics, Springer, 2006.
46. Jiang T, Gradus JL, Rosellini AJ. Supervised machine learning: a brief primer. Behav Ther 2020;51(5):675–87.
47. Alanazi , A . (2022). Using machine learning for healthcare challenges and opportunities. Informatics in Medicine Unlocked 30:100924.
48. Kendale S, Kulkarni P, Rosenberg AD, Wang J. Supervised machine-learning predictive analytics for prediction of postinduction hypotension. Anesthesiology 2018;129(4):675–88.
49. Probst P, Boulesteix AL, Bischl B. Tunability: importance of hyperparameters of machine learning algorithms. :32.
50. Unsupervised machine learning. DataRobot AI cloud. Accessed December 17, 2021.
51. Osi´nski B, Budek K. What is reinforcement learning? The complete guide. deepsense.ai. Published July 5, 2018. Accessed December 17, 2021.
52. Riachi E, Mamdani M, Fralick M, Rudzicz F. Challenges for Reinforcement Learning in Healthcare. arXiv:210305612 [cs]. Published online March 9, 2021. Accessed December 17, 2021.
53. Bui, N., Cesana, M., Hosseini, S. A., Liao, Q., Malanchini, I., and Widmer, J. (2017). "A survey of anticipatory mobile networking: Context-based classification, prediction methodologies, and optimization techniques". *IEEE Communications Surveys & Tutorials*, 19(3), 1790-1821.
54. D. Gnanambal, D. Thangaraj, Meenatchi V T, and D. Gayathri, "Classification Algorithms with Attribute Selection: an evaluation study using WEKA," *Int. J. Adv. Netw. Appl.*, vol. 09, no. 06, pp. 3640–3644, 2018.
55. A. K. Jain, H. Dhawan, and B. Sowmiya, "DDoS Detection Using Machine Learning Ensemble," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 12, pp. 1647–1655, 2021.
56. Yassin B., Mohamed C. and Yassine A. (2021). A Nonlinear Support Vector Machine Analysis Using Kernel Functions for Nature and Medicine.
57. N. Kausar, B. Belhaouari Samir, A. Abdullah, I. Ahmad, and M. Hussain, "A review of classification approaches using support vector machine in intrusion detection," Commun. Comput. Inf. Sci., vol. 253 CCIS, no. PART 3, pp. 24–34, 2011
58. Oommen,T., Misra, D., Twarakavi, N. K. C., Prakash, A., Sahoo, B. and Bandopadhyay, S. "An objective analysis of support vector machine based classification for remote sensing," Math. Geosci., vol. 40, no. 4, pp. 409–424, 2008 .
59. Jader, R.F., Abd, M.H.M. and Jumaa, I.H., 2022. Signal Modulation Recognition System Based on Different Signal Noise Rate Using Artificial Intelligent Approach. Journal of Studies in Science and Engineering, 2(4), pp.37-49.
60. Tuysuzoglu, G., Birant, D. and Pala, A. (2019) 'Majority voting based multi-task clustering of air quality monitoring network in Turkey', Applied Sciences (Switzerland), 9(8), pp. 1–21 .
61. Cristianini, N. and Shawe-Taylor, J. *An Introduction to Support Vector Machines and Other Kernel based Learning Methods*. Cambridge University Press, 2000.
62. Altman, N. S. "An introduction to kernel and nearest-neighbor nonparametric regression," *Am. Stat.*, vol. 46, no. 3, pp. 175–185, 1992 .
63. Zhang, Y., Cao, G., Wang, B. and Li, X. "A novel ensemble method for k-nearest neighbor," Pattern Recognit., vol. 85, pp. 13–25, Jan. 2019.
64. Marcot, B. G. and Penman, T. D. "Advances in Bayesian network modelling: Integration of modeling technologies," *Environmental Modelling and Software*, vol. 111. Elsevier Ltd, pp. 386–393, Jan. 01, 2019 .
65. Drury,B., Valverde-Rebaza, J., Moura, M. F. and de Andrade Lopes, A. "A survey of the applications of Bayesian networks in agriculture," *Eng. Appl. Artif. Intell.*, vol. 65, pp. 29–42, Oct. 2017 .
66. Du,J., Zhai, C. and Wan, Y. "Radial basis probabilistic neural networks committee for palmprint recognition," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2007, vol. 4492 LNCS, no. PART 2, pp. 819–824 .
67. "Neural Networks for Pattern Recognition Guide books." (accessed Dec. 29, 2020).

68. M. Paliwoda, “Decision Trees Learning System,” in *Intelligent Information Systems*, Physica- Verlag HD, 2002, pp. 77–90.
69. T. Aytaç, M. A. Aydın, and A. H. Zaim, “Detection DDOS attacks using machine learning methods,” *Electrica*, vol. 20, no. 2, pp. 159–167, 2020 .
70. Reis, I., Baron, D., & Shahaf, S. (2018). Probabilistic Random Forest: A Machine Learning Algorithm for Noisy Data Sets. *The Astronomical Journal*, *157*(1), 16.
71. Ozgode Yigin, B., Algin, O., & Saygili, G. (2020). Comparison of morphometric parameters in prediction of hydrocephalus using random forests. *Computers in Biology and Medicine*, *116*, 103547.
72. Shouman, M., Turner, T. and Stocker, R., 2012. Applying k-Nearest Neighbour in Diagnosing Heart Disease Patients. International Journal of Information and Education Technology, pp.220-223.
73. 2012. Introduction To K Nearest Neighbour Classification And Condensed Nearest Neighbour Data Reduction. [ebook] University of Leicester, p.1. Available at: https://staff.fmi.uvt.ro/~daniela.zaharie/dm2018/ro/TemeProiecte/Biblio/kNN/CondensedNearestNeigh-bor.pdf>].
74. Budianto, A., Ariyuana, R., & Maryono, D. 2018. Perbandingan K-Nearest Neighbor (Knn) Dan Support Vector Machine (Svm) Dalam Pengenalan Karakter Plat Kendaraan Bermotor [Comparison Of K-Nearest Neighbor (Knn) And Support Vector Machine (Svm) In Introduction To Plat Vehicle Characters]. *Jurnal Ilmiah Pendidikan Teknik dan Kejuruan*, 11(1): 27-35.
75. Anggoro, D. A., Rahmatullah, P. I. 2020. The Implementation of Subspace Outlier Detection in K-Nearest Neighbors to Improve Accuracy in Bank Marketing Data. *International Journal of Emerging Trends in Engineering Research*, 8(2).
76. Choudhary R, Gianey HK. Comprehensive review on supervised machine learning algorithms. In: 2017 international conference on machine learning and data science. MLDS); 2017. p. 37–43.
77. Harrison O. Machine learning basics with the K-nearest neighbors algorithm. Medium., 16, 2021.

Cite this Article: Salah Mashal Raheem, Ziad Mohammed Abood (2023). Detection and Classification of Gastrointestinal Diseases by using Machine Learning: A Review. International Journal of Current Science Research and Review, 6(8), 6056-6066