REPUBLIC OF TURKEY

ALTINBAŞ UNIVERSITY

Institute of Graduate Studies

Electrical and Computer Engineering

DIAGNOSE COLON DISEASE BY FEATURE SELECTION BASED ON ARTIFICIAL NEURAL NETWORK AND GROUP TEACHING OPTIMIZATION ALGORITHM

Alaa Badr EYSA

Doctor of Philosophy Thesis

Supervisor

Asst. Prof. Dr. Sefer KURNAZ

Istanbul, 2022

DIAGNOSE COLON DISEASE BY FEATURE SELECTION BASED ON ARTIFICIAL NEURAL NETWORK AND GROUP TEACHING OPTIMIZATION ALGORITHM

Alaa Badr EYSA

Electrical and Computer Engineering

Doctor of Philosophy

ALTINBAŞ UNIVERSITY

2022

The dissertation titled DIAGNOSE COLON DISEASE BY FEATURE SELECTION BASED ON ARTIFICIAL NEURAL NETWORK AND GROUP TEACHING OPTIMIZATION ALGORITHM prepared BY ALAA BADR EYSA and submitted on 13/12/2022 has been **accepted unanimously** for the degree of PhD in Electrical and Computer Engineering.

_____

Asst. Prof. Dr. Sefer KURNAZ

the Supervisor

Thesis Defense Committee Members:

Asst. Prof. Dr. Sefer KURNAZ	Department of Software Engineering, Altınbaş University	_____
Prof. Dr. Osman NURİ UÇAN	Department Of Electrical and Electronics Engineering, Altınbaş University	_____
Assoc. Prof. Dr. Adil Deniz DURU	Department of Coaching Training, Marmara University	_____
Asst. Prof. Dr. Oğuz KARAN	Department Of Computer Engineering, Altınbaş University	_____
Asst. Prof. Dr. Zeynep ALTAN	Department of Software Engineering, Beykent University	_____

I hereby declare that this dissertation meets all format and submission requirements of a PhD dissertation.

Submission date of the thesis to Institute of Graduate Studies: ___/___/___

I hereby declare that all information/data presented in this graduation project has been obtained in full accordance with academic rules and ethical conduct. I also declare all unoriginal materials and conclusions have been cited in the text and all references mentioned in the Reference List have been cited in the text, and vice versa as required by the abovementioned rules and conduct.

Alaa Badr EYSA

Signature

DEDICATION

I dedicate and promise this study effort to my supervisor, who has been instrumental in helping me through the entire research process, as well as my family, who has always been supportive during my difficult times.

PREFACE

First and foremost, I want to express my gratitude to my supervisor, Asst. Prof. Dr. Sefer KURNAZ, for guiding and assisting me during the composition of this dissertation. I met with my supervisor, Asst. Prof. Dr. Sefer KURNAZ weekly discusses my progress, challenges, and suggestions. This helped me enormously in grasping the reasoning behind the research. It helped me understand the technological requirements for this study project.

ABSTRACT

DIAGNOSE COLON DISEASE BY FEATURE SELECTION BASED ON ARTIFICIAL NEURAL NETWORK AND GROUP TEACHING OPTIMIZATION ALGORITHM

Eysa, Alaa Badr,

PhD, Electrical and Computer Engineering, Altınbaş University,

Supervisor: Asst. Prof. Dr. Sefer KURNAZ

Date: December / 2022

Pages: 54

Multiple imaging modalities have substantially enhanced the diagnostic precision of contemporary medicine. Medical imaging has progressed to the point where an accurate diagnosis is achievable, and therapy can be commenced quickly. This article does an excellent job describing how to detect colorectal cancer. The recommended method for diagnosing plant diseases use a group teaching optimization algorithm to identify the most pertinent elements of an image. Multiple-layered trained neural networks are used to identify pictures as benign or malignant. The suggested technique for colon cancer diagnosis achieves a mean accuracy of 92.72%, sensitivity of 93.14%, and F1-Score of 94.26% when tested on the color data set. Kvasir picture data were classified with 96.42 percent accuracy, precision, sensitivity, plant disease, and F1-Score using the proposed method. Experiments demonstrate that the proposed strategy is superior to 3Layer for classifying photos of plant diseases. CNN There are currently numerous imaging modalities available for use in medical diagnostics. The four algorithms (DF, RF, CNN, and TFL).

Keywords: Plants Disease, Group Teaching Optimization Algorithm, Feature Selection, Artificial Neural Network.

TABLE OF CONTENTS

Pages

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

ABBREVIATIONS

ANN :	Artificial Neural Network
GA :	Genetic Algorithm
MS :	Multiple Sclerosis
CT :	Computer Tomography
PSO :	Particle Swarm Optimization
PCA :	Principal Component Analysis
FFNN :	Feed Forward Neural Networks

1. INTRODUCTION

1.1. BACKGROUND

Cancers of the colon and rectal are commonly induced by polyps in these areas. Each year, around one million people are diagnosed with colon or rectal cancer. Due to the fact that it is one of the most severe forms of cancer, this illness merits considerable attention.

Cancer of the colon, rectum, or large intestine is commonly referred to as "colorectal cancer" (CRC). Men are slightly more likely than women to develop CRC during their lifetimes. This rate is 4.5 percent for men and 3.2 percent for women. Adenoma is the most prevalent CRC precursor (polyps).

Due of their lengthy silence, these benign tumors of the colon or rectal mucosa are notoriously difficult to diagnose. As individuals age, adenomas tend to multiply. The progression of cancer from an adenoma has been thoroughly described, but research into the environmental elements that influence this progression is ongoing.

1.2. COLON AND RECTUM CANCER (CRC) PREVENTION WAYS

The main condition for protection from colon and rectum cancers is not to neglect the necessary screening tests. It is the primary way to prevent colon cancer to have occult blood in the stool, colonoscopy or sigmoidoscopy tests as recommended by your physician who evaluated your family history and medical history.

On the other hand, avoiding cigarettes, tobacco products, and alcohol, adopting an active lifestyle in terms of sports and exercise, not overconsumption of fat and red meat, and adopting a diet rich in whole grains and fiber foods are also protective against colon cancer.

1.3. COLON CANCER

Colon cancer refers to malignancies that begin in the intestinal lining of the colon, a 1.5-meter-long organ. This cancer develops when the cell and cell community populations of the inner lining of the large intestine expand uncontrolled.

15 to 20 cm from the end of the rectum or large intestine are where rectal cancers develop. Commonly, the term "colorectal cancers" refers to both of these disorders.

According to the Ministry of Health, colorectal cancers are one of the five most prevalent malignancies in our nation.

1.4. RISK FACTORS

1.4.1. Gender

There is no difference in incidence between men and women.

1.4.2. Other Risk Factors

Although the exact cause of colorectal cancer is not known, the risk factors that increase the development of colorectal cancer are as follows:

advanced age,

Presence of polyps in the intestine (especially those with adenomatous pathology),

Presence of colorectal cancer patients in the family,

Presence of certain genetic disorders (patients with non-hereditary polyposis colon cancer) and/or familial polyposis syndromes characterized by hereditary polyps in the colon and rectum, which have caused significant changes in genes,

Having inflammatory bowel disease (Ulcerative colitis or Crohn's disease) that may cause cancer by disrupting the intra-intestinal cell type within a certain period of time,

Having a history of ovarian, breast and uterine cancer in women,

Excessive consumption of processed and animal foods, less consumption of fruits and vegetables and smoking,

People with these risk factors should be screened for bowel cancer from an earlier age.

1.5. COLON CANCER SYMPTOMS

Polyps, which are intestinal cell growths, are the earliest warning signs of colon cancer. On the other hand, while patients first feel well, polyps frequently go unreported. When polyps advance to malignancy, increase in size, or multiply, abnormalities in bowel habits occur.

Anemia is characterized by newly developed constipation or, alternatively, a change in stools odor or consistency (in favor of diarrhea).

Abdominal pain, appetitelessness, and uncontrollable weight loss. After using the restroom, if there is blood in the feces, blood mixed with the stool, or bleeding from the auricle, medical attention should be sought.

These symptoms and indicators are insufficient to diagnose colorectal cancer. Nonetheless, if you observe any of these symptoms, it is crucial that you consult a physician to determine what's

wrong and how to cure it. If you have a family history of colon, breast, ovarian, or cervical cancer with one of these symptoms, you should seek medical attention immediately.

Bleeding from the rectum is not always a sign of colorectal cancer; it may be due to hemorrhoids or anal fissures, particularly in younger individuals, those with chronic constipation, and those with no other worrying symptoms. Without further study, the real cause of these disorders cannot be discovered in many instances..

1.6. ULCERATIVE COLITIS FOR COLON CANCER

Ulcerative colitis and Crohn's disease inflame the large intestine. Each causes colorectal cancer. Long-term ulcerative colitis patients (even those who have responded to treatment) have a higher risk of colorectal cancer (even if they have responded to treatment). These patients should have routine screening colonoscopies, regardless of symptoms.

In Crohn's disease-affected areas, colorectal cancer risk rises. In these two situations, continuing to smoke raises colorectal cancer risk.

Early colon cancer detection includes colonoscopy, stool blood testing, CT colonography, flexible sigmoidoscopy, and stool DNA testing.

Colonoscopy involves introducing a light and camera via the anus into the large intestine to check its inside while slightly expanding it. 30-45 minutes is typical.

Colon mucosal lining and abnormal-appearing tissues may also be biopsied. In the same session, polyps can be removed. Colonoscopy can stop bleeding foci. Colonoscopies should be approached with caution.

An anesthesiologist can sedate or give pain medication for a colonoscopy. A colonoscopy is less uncomfortable than many think.

The day before the colonoscopy, the patient must consume drugs and watery foods to cleanse their intestines (preparation for colonoscopy). Before the procedure, the doctor, nurse, or other healthcare expert will describe the preliminary processes and treatment strategy to calm the patient.

People without a personal or family history of colon cancer start colonoscopy screening at 50 and repeat it every 10 years if no polyps are discovered.

When a first-degree relative has colon, breast, uterine, or ovarian cancer before 65, the screening age is 40. If the person's oldest first-degree relative is under 40, it should be screened young.

Patients with ulcerative colitis for more than ten years should get annual colonoscopies.

Genetic polypoisis syndromes should be screened between 15 and 18 years old.

Depending on the quantity and pathology of polyps discovered during the screening colonoscopy, control colonoscopies are conducted every one to five or 10 years. Flexible sigmoidoscopy can detect left colon cancer between full colonoscopic exams.

"Annual colonoscopy" is false unless specific conditions are met.

CT colonography can replace colonoscopy every five years if a patient is unwilling. CT colonograms are only utilized for diagnosis; a colonoscopy is needed to biopsy polyps or masses.

Blood in the stool should be tested annually before the first colonoscopy. Every two to three years, excrement DNA is tested.

A healthy diet and regular screenings are key to preventing colon cancer.

You can prevent colon cancer by eating fiber-rich vegetables, fruits, and avoiding processed meals and animal protein.

MS, colorectal cancer, Alzheimer's disease, and other illnesses are easily diagnosed with imaging [1–4]. Colon cancer [5], a human disease, kills millions annually. Using CT scans and MRIs, this illness can be diagnosed. Colon cancer is a tumor in the colon. Due to aberrant cell development, malignant cells target healthy tissues [6]. Lifestyle, aging, and genetics cause colon cancer. Food, obesity, smoking, and inactivity are risk factors. Studies link alcohol and processed meats to sickness and cancer [7]. Early colorectal cancer forms a benign tumor. Over time, these tumors become malignant. Regular colonoscopies detect colon cancer. Colonoscopies are limited. Medical imaging primarily determines illness severity. CT scans drastically reduce colon cancer deaths [8]. Doctors can track illness progression and recurrence this way. Surgery, radiotherapy, chemotherapy, and medicine can treat this disease. According to the study, industrialized nations have higher cases. Experiments show these countries cause 65% of incidents. Less women than men have this syndrome [9]. Medical picture analysis can detect disease. Digital images and image processing technologies help evaluate colon cancer. Several studies were evaluating CT scans and other medical imaging. Most sickness diagnosis techniques use machine learning and deep learning. Medical images can be studied using artificial neural networks [4, SVMs [14], fuzzy algorithms [15], and expert systems [16] to identify disease. Splitting images helps locate damaged areas. This work is innovative since it does not employ CT image feature selection or statistical methodologies like PCA to choose features. The article's strengths are its clever feature selection technique and focus on categorization learning. Feature selection helps diagnose colon cancer..

1.7. RESEARCH PROPOSED

It is recommended that a binary variant of the group teaching optimization algorithm [17] be used in the proposed methods for feature selection. In the second stage, images are classified as benign or malignant using a trained neural network and a small number of features. We successfully segmented images associated with colon diseases using a GTO strategy for feature selection, the neural network's ability to learn, and the classification of the diseases into distinct classes.

1.8. RESEARCH APPROACH

The diagnosis of colon cancer frequently involves the use of imaging tools, and it is crucial to be able to interpret these images accurately. Combining human learning and training with machine learning techniques, such as neural network learning, this thesis seeks to identify the areas affected by Cologne syndrome. In the proposed method, the PCA algorithm is used to improve the accuracy of the PCA technique for selecting the appropriate feature through the optimization teaching and learning algorithm prior to disease region separation using neural network-based learning.

1.9. RESEARCH MOTIVATION

Highlights of the thesis:
Use the component reduction method to improve the knowledge and information contained in the image
Using human intelligence to select important image pixels in combination with PCA
Use the ability to learn the neural network for teaching and learning
Review and analyze images and provide an accurate segmenting of Cologne disease images.

1.10. ORGANIZATION OF THE THESIS

This thesis is divided to various sections. The problem has been introduced in the first section. The second section of the research background focuses upon researches in Colon disease field. New training learning, which had been utilized in feature selection, and the suggested approach for colon disease diagnosis has been formulated in fourth section utilizing binary version of training learning. Implementations and experiments, as well as the suggested approach and its

comparison to existing approaches, and the research findings and future ideas for developing a suggested algorithm in diagnosing of the colon disease have been described in the fourth section and final section.

2. LITERATURE REVIEW

This section analyzes the paper's colon cancer diagnosis guidelines. The talk opens with an overview of colorectal cancer and ML/DL research on colon cancer diagnoses.

2.1. COLON CANCER

Colorectal cancer patients getting treatment should undergo imaging. After endoscopic and clinical evaluation, pelvic MRI [18] is needed. Infections, fevers, hemorrhoids, and IBD can mimic colon cancer. If you have any of these symptoms, see a doctor immediately [19]. Early rescue procedures can boost survival. Some cancer patients, such as the elderly, may have no symptoms. Screening can detect and treat colorectal cancer early. A screening method is medical imaging. Colon polyps can cause colon cancer (adenoma). Large adenomas may lead to cancer [20]. Medical imaging aids early colon cancer detection. Figure 2.1 depicts the histology of mucosal adenocarcinoma. The red arrow represents cancerous round tissues. The diagnosing process requires a microscopic exam. Colonoscopy and MRI are others. MR imaging can detect digestive tract malignancies (Figure 2.2). Figure 2.3 shows a colon cancer CT scan. [21] [22]:

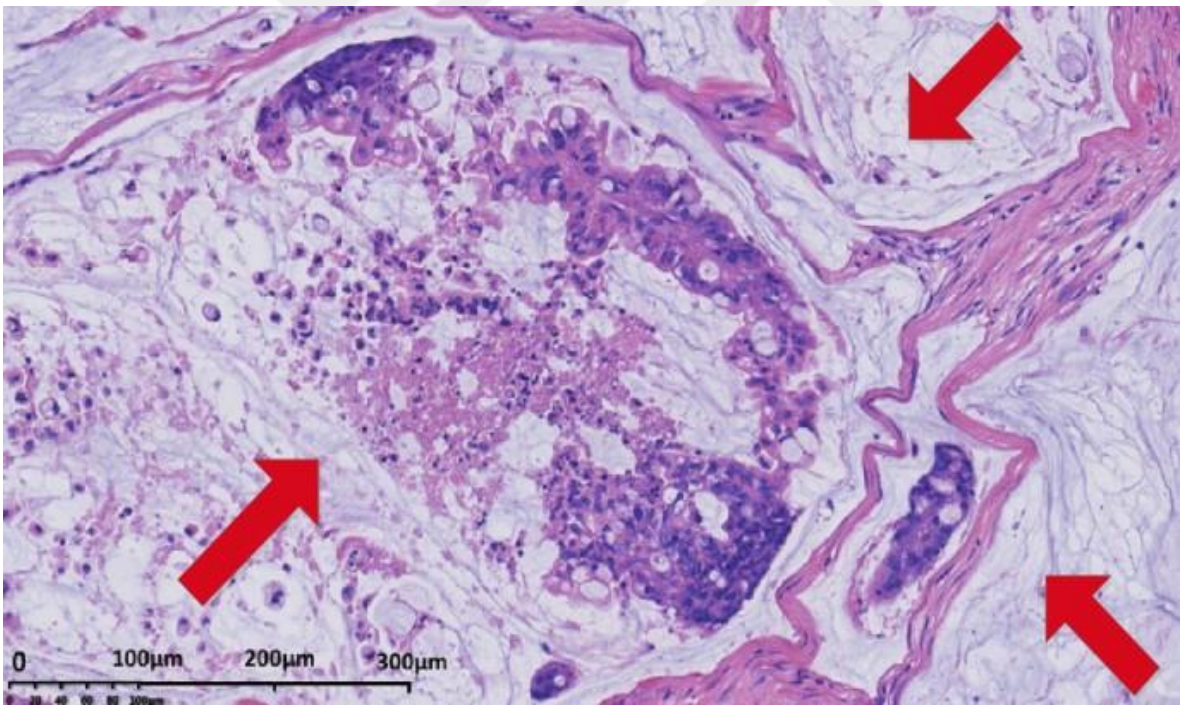


Figure 2.1: A picture of the mucosal area of colon tissue and cancer cells, the abundance of which shown in an area with a few red arrows

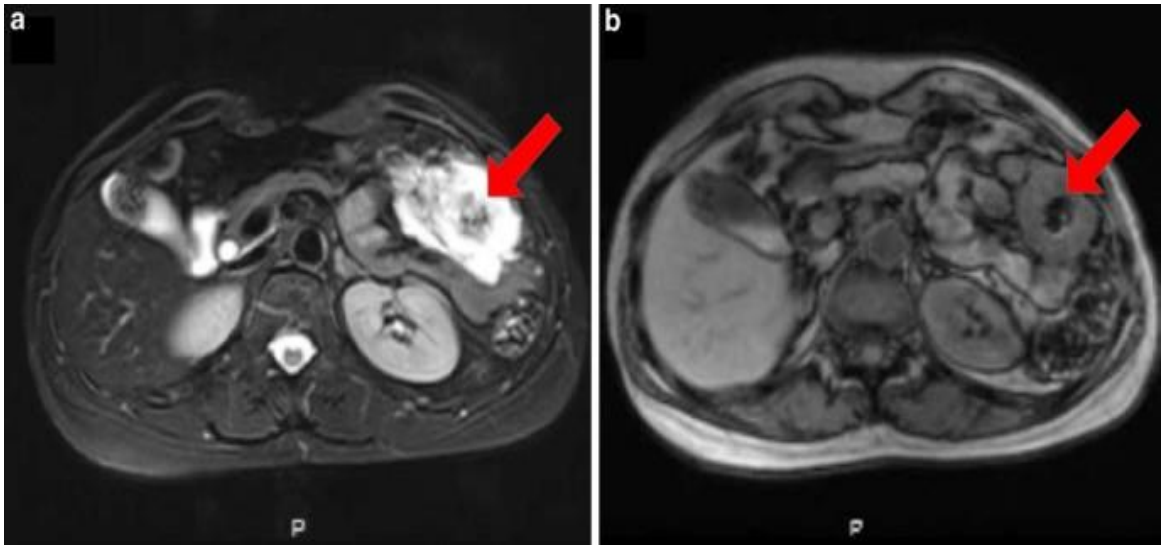


Figure 2.2: Left to right, T2 and T3 magnetic resonance images, respectively.

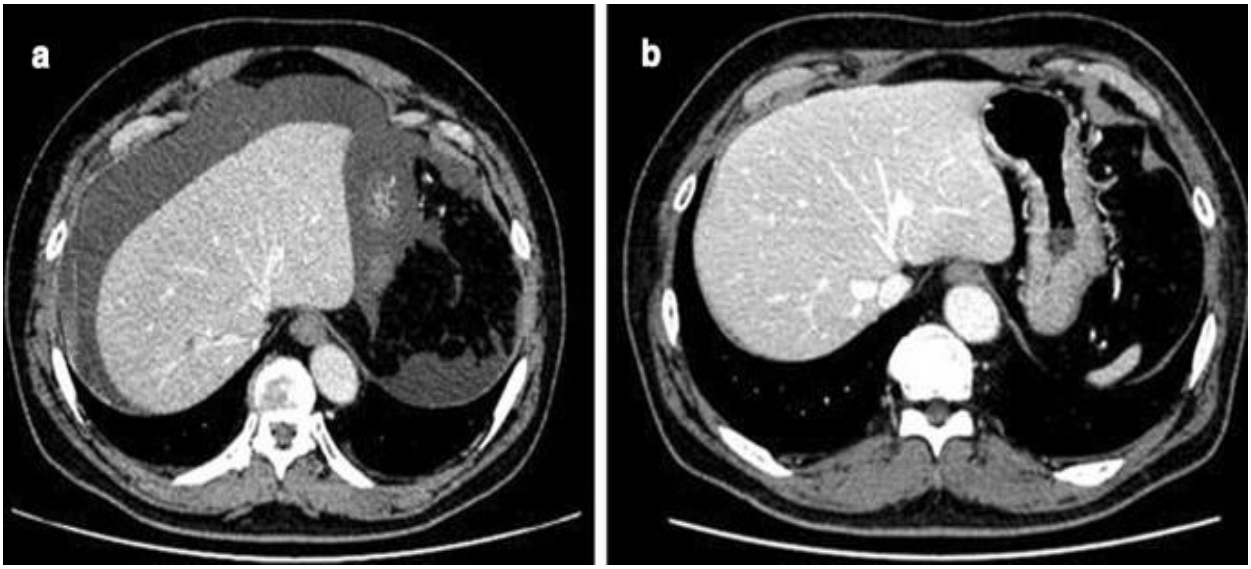


Figure 2.3: Two examples of colon CT scan images (tumor areas with bright spots)

2.2. DIAGNOSIS OF COLON DISEASE

The authors of the study claim [23] that deep learning will be able to accurately diagnose lung and colon cancer by 2021. According to their findings, the second largest cause of death worldwide is cancer. The mortality rate is roughly one in six. Lung and colon cancers are the most prevalent and deadly forms of the disease. Colon cancer accounts for one-fourth of all cancer diagnoses globally. This study proposes a framework for distinguishing between two benign and three malignant types of lung and colon tissue based on histological images by utilizing cutting-edge deep learning algorithms and digital image processing. The proposed framework has a maximum accuracy of 96.33 percent for recognizing malignant tissues, according to statistical research. Using this algorithm, physicians will be one step closer to creating an automated method for categorizing the many kinds

of colon and lung cancer. A paper [24] published in 2021 outlined a machine learning-based method for predicting colorectal cancer in patients. The ability of 110 created factors to predict patient survival is examined using univariate regression. With the use of a randomized forest-based machine learning technique, common prognostic characteristics, such as the characteristic vector, clinical factors, and pathological variables, have been explored. This collection of data is composed of 75% test data and 25% training data. Utilizing ROC and fuzzy matrix analysis, the performance of the categorization process was evaluated. Researchers improved the accuracy of their survival prediction model by using imaging-based heterogeneity factors. The median survival time for patients in the study was 39.3.9 months. In a 2021 study [23], histological classification techniques for colorectal cancer employing global labeling and deep learning were described. Colon cancer is the second most lethal type of cancer. Histological imaging is crucial for diagnosing colon cancer. For cancer classification, deep learning techniques are currently favored to a greater extent. This study employs deep learning to detect colorectal cancer. Fifty hospital samples of recent colon cancer were tested independently, and a 92% rate of accuracy was revealed. The 2021 study [25] advocated using deep learning on histopathology images to diagnose colon cancer. Their objective is to develop a synthetic framework capable of recognizing spatial prognostic patterns in histopathology pictures. To effectively classify tissue, they trained and evaluated a deep convolution neural network with 107,180 images. The study suggested that machine learning may be able to detect colon cancer by 2020. Machine learning will improve the diagnosis of colon cancer. For colon cancer screenings, the National Cancer Database is utilized. The 2020 study [27] explores the current state of knowledge about the detection of colon cancer. By analyzing colorectal cancer data with deep learning, they expect to progress the area. This examination focuses mostly on the deep learning frameworks that have been applied to the study of colon cancer. Colorectal cancer is a common issue for deep learning investigations. Several elements of cancer care have been examined, including detection, categorization, and prognosis. Machine learning and spatial attributes were utilized to detect colon polyps in endoscopic images in the year 2020. Polyps in the colon are an early indicator of colon cancer. It is crucial to prevent polyps from growing into cancer. It is famously difficult to detect polyps in their earliest stages using medical pictures due to sample characteristics and technological constraints. The classification of polyps automatically is the focus of this study. Using this method, pathologists and other medical professionals can identify intestinal polyps with greater ease. In this study, significant polyp properties are extracted using an endoscopic colonography-trained neural

network. The proposed method detects a wide variety of polyps with high accuracy (93.94 percent correct, 6.05 percent false positive rate).

2.3. ARTIFICIAL NEURAL NETWORK

2.3.1. Background

ANN mathematically represents a person's NNs. NNs consist of linked neurons and nerve cells. The average brain has 10-11 neurons. Momentarily stimulating a layer or group of cells' electrical impulses helps create neuronal connections. Neurotransmitters deliver information via electrochemical connections between neurons. Dendrites attach to the cell body through these junctions. Dendrites of neurons provide signals to the nerve cell body, which analyzes them to determine if an electrical signal will be formed. Neuron-to-neuron signals can be excitatory or inhibitory. Inhibitory influence stops neuron firing, while excitatory influence activates receiving neuron. The conductivity of electrochemical junctions that transfer information into the cell body limits how much any information source can affect the cell as a whole. Figure 2.4 displays the neuron's basic components.

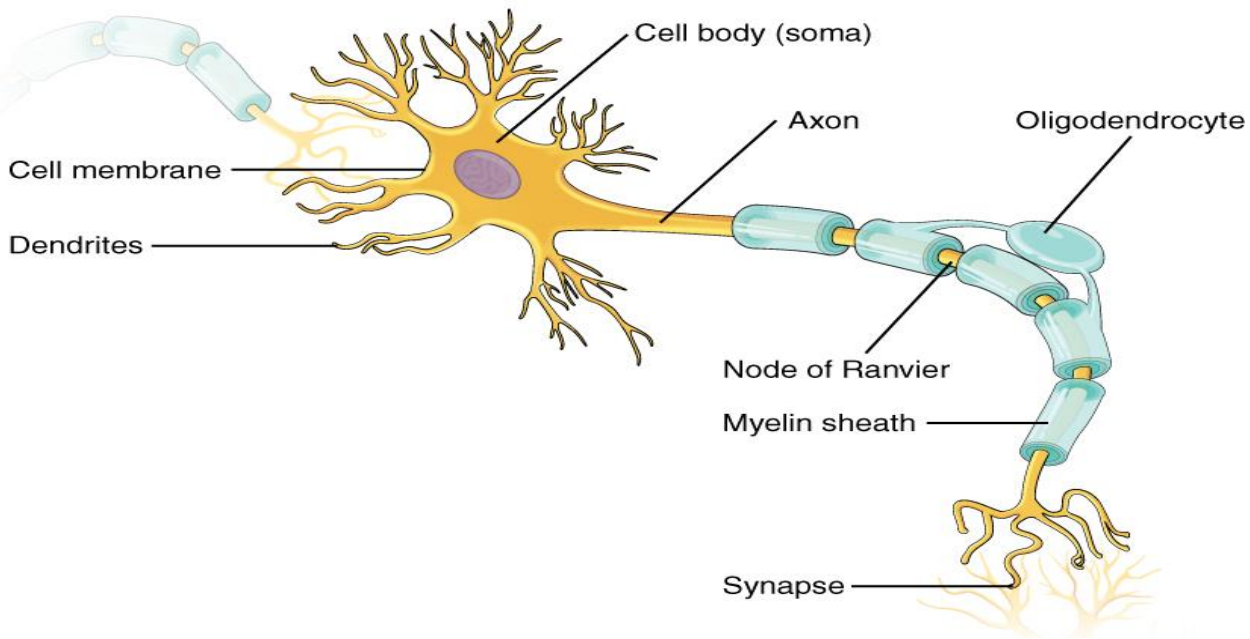


Figure 2.4: The Essential Components of the neuron's [29]

By employing mathematical principles called as loading, ANNs are able to modify the relevance of neuron inputs. Prior to transmitting any data to the neurons, it is replicated using the comparing weight. Then, independent of the value of any inherent bias, the neuron sums these weighted

characteristics and transmits the result via actuation. hence the output of the neuron is always 0. If a bias value is set, neurons can determine, based on the feature, whether or not they have generated the desired output. In Equation 2.1, both the bias of neuron j and the sum of its inputs, S, are displayed.

$$s_j = \sum_i x_i \, w_{ij} + b_j \tag{2.1}$$

xi represents a neuron's contributions, yi its loads, and bj its bias in this notation. Figure 2.5 depicts the neuron's visual representation of the factors that contribute to a request registering result.

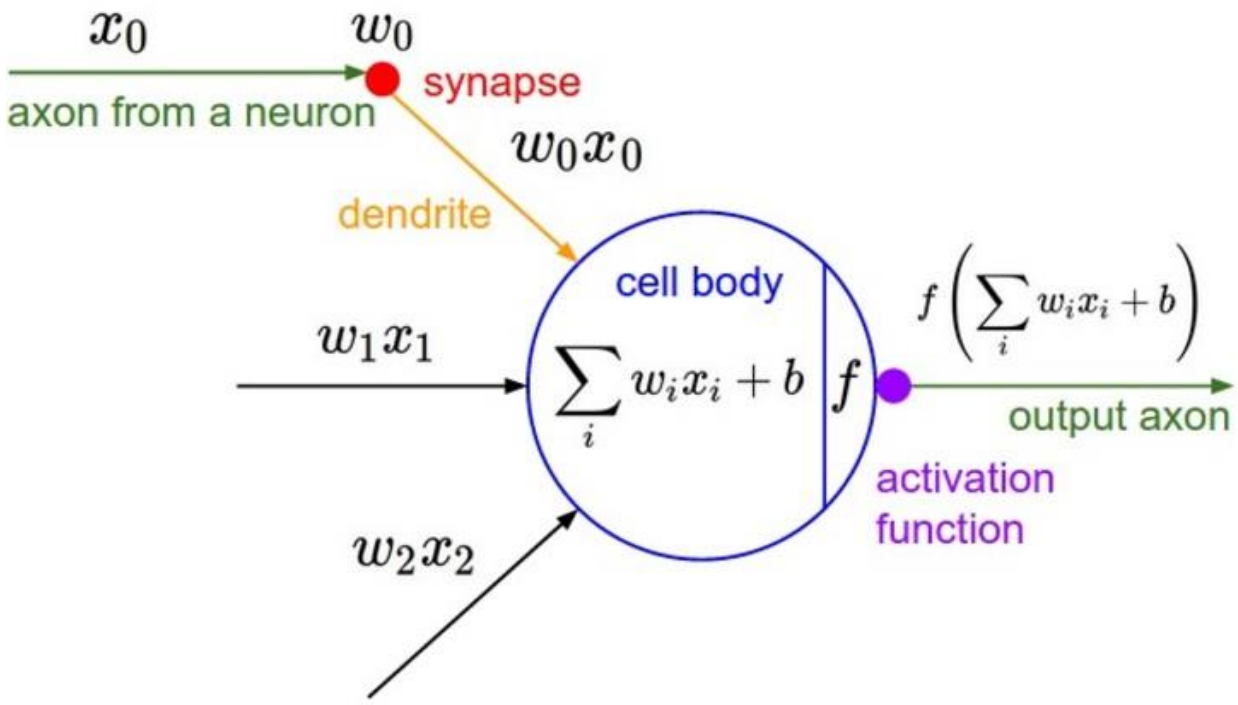


Figure 2.5: Visual representation of Neuron's Calculation operations [30]

The neuron gains the ability to set nonlinear limits for independent direction by converting the summation's results into an actuation effort. If an actuation work is removed from the neuron's calculation, the only possible limit a neuron can use to divide tuples in a data set into classes is a straight limit, which restricts the neuron's ability to make more precise predictions. In addition, neurons identified later in NN may be able to generate more complex limits for each class, thereby enhancing the accuracy of NN's predictions. In addition, the use of inclination values within each

neuron can facilitate the development of these complex limits by modifying the areas of each piece of mind-boggling limit, which is created by linking neuron limits before neuron. As shown in Figure 2.6, the Hyperbolic Tangent (TanH), Sigmoid, and Rectified Linear Unit (ReLU) are three commonly used initiation capacities. In terms of execution, NNs with ReLU initiation capacities have proven to be vastly superior to those with other enactment capacities. Due to the nonlinearity of the calculations, the outcome can be calculated from the contributions by identifying the necessary components. Due to the fact that a NN may take different paths to arrive at a particular result, and thus may be the result of a single element of a mixture of features, such networks are used as single-direction capacities for producing hash values that may be used to represent data, although hash esteem cannot be used for recovering initial information. [31][32][33].

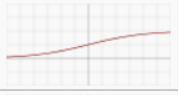
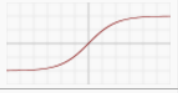
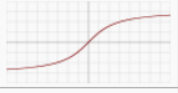
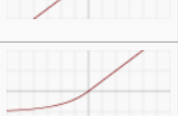
Name	Plot	Equation	Derivative
Identity		$f(x) = x$	$f'(x) = 1$
Binary step		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x \neq 0 \\ ? & \text{for } x = 0 \end{cases}$
Logistic (a.k.a Soft step)		$f(x) = \frac{1}{1+e^{-x}}$	$f'(x) = f(x)(1-f(x))$
TanH		$f(x) = \tanh(x) = \frac{2}{1+e^{-2x}} - 1$	$f'(x) = 1 - f(x)^2$
ArcTan		$f(x) = \tan^{-1}(x)$	$f'(x) = \frac{1}{x^2+1}$
Rectified Linear Unit (ReLU)		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Parameteric Rectified Linear Unit (PReLU) [2]		$f(x) = \begin{cases} \alpha x & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Exponential Linear Unit (ELU) [3]		$f(x) = \begin{cases} \alpha(e^x - 1) & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$	$f'(x) = \begin{cases} f(x) + \alpha & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
SoftPlus		$f(x) = \log_e(1+e^x)$	$f'(x) = \frac{1}{1+e^{-x}}$

Figure 2.6: Neurons' Activation Function [34][35]

2.3.2. The Feed Forward Neural Networks (FFNN)

The neurons in FFNN are disseminated in layers, with the results of neurons in each layer being linked to contributions to neurons in subsequent layers. The yield, information, and secret layers are the 3 layer types in the network. The information layers are the main layer with which the contributions from the outside world are associated, and the number of neurons in this layer corresponds to number of contributions that have been made by NN, whereas the number of neurons in result layer corresponds to number of results that are required by the framework. As a result, the number of neurons in such layers is limited by amount of execution that is required in order to complete the assignment for which an NN is being performed. Furthermore, in all cases, using only information and result layers is insufficient to complete the required task, due to the fact that it limits the amount and complexity of elements that the NN could recognize for the purpose of providing correct forecasts. Figure 2.7 shows an example of entirely associated feedforward deep neuronal architecture.

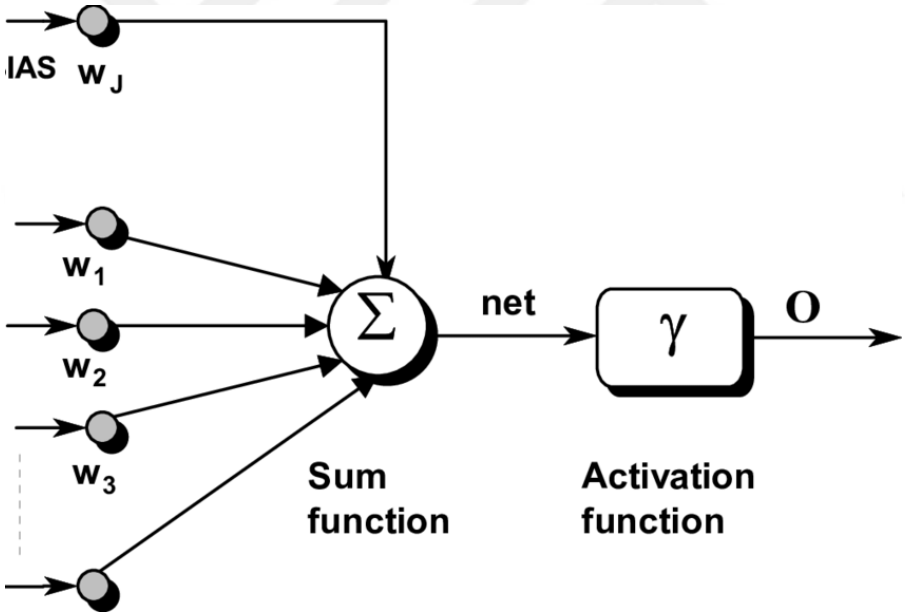


Figure 2.7: Examples of Artificial Neural Networks that Feed Forward.

Typically, the majority of NN display layers are inserted between the information and output levels. This layer is a covered layer since it is not directly accessible from the outside, unlike the information and result layers, which are connected to the framework's data sources and outputs. There are no hard limits on the number of neurons that can be placed in a hidden layer or storage layer; nevertheless, doing so increases the complexity of the computations required to forecast a class of information and,

consequently, the time required to do so. In addition to increasing the number of mixes that activate the neurons in that layer, an increase in the number of hidden layers also increases the number of visible highlights in that layer. For instance, as the network's depth increases, so does the complexity of the elements that might be located inside it. Networks containing several hidden nodes, sometimes known as "profound NNs," are able to detect more exceptional details. Overfitting is one of the core issues investigated in advanced NNs; it arises when predictions rely excessively on features that are only implicitly present in the NN. There is a possibility that other data sources that possibly belong to that class but do not stimulate the neurons associated with such things will be mislabeled. To solve this issue, a percentage of neurons from the hidden layer are deleted at random at the conclusion of each preprocessing cycle. This decreases the NN's reliance on explicit elements and forces it to generate the same forecast in innovative ways. Figure 2.8 illustrates how Dropout significantly improved NN predictions.

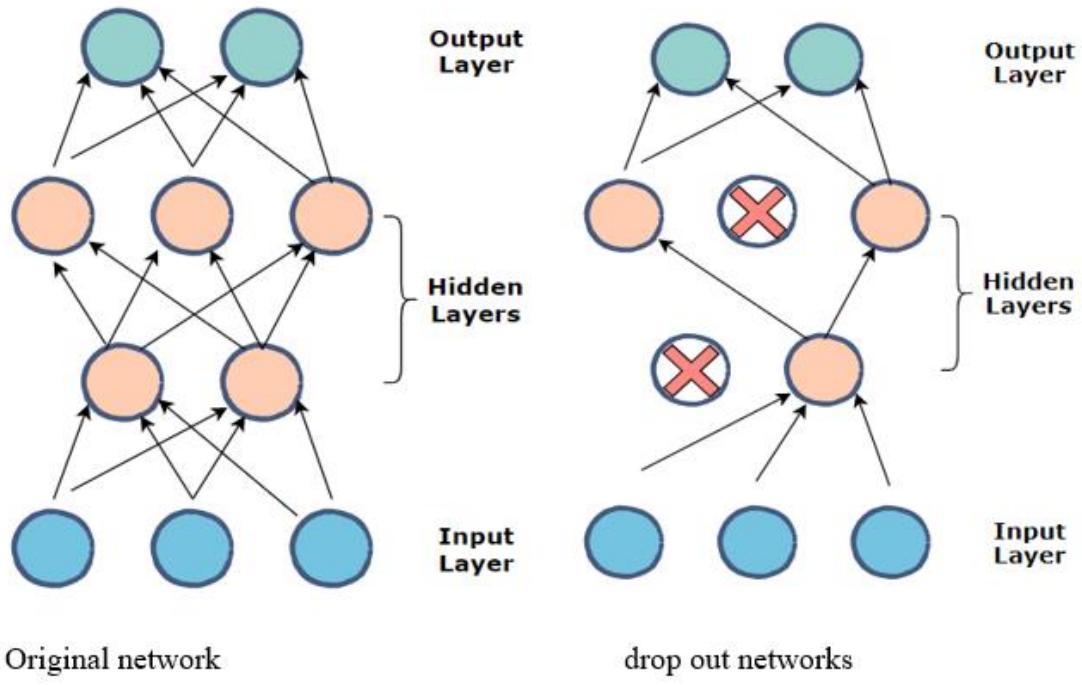


Figure 2.8: Dropout Neural Net Model [36]

The forecasts that are provided by am NN are results of calculations that have been carried out in every one of the neurons, from the information layer to result layer, which are based mostly on the inclinations and loads of the NN, notwithstanding the information upsides of the information layers.

The preparation stage entails adjusting the inclinations and loads of the NN for working on the precision of the forecasts provided by such networks, as the information values are behindhand the management of such networks. Using the Sum of Squared Errors (SSE) approach shown in Eq. 2.2, this update is completed through the estimation of the difference between necessary and anticipated quality, which is referred to as expense.

$$C = \sum_i \frac{1}{2}(y_i - \hat{y}_i)^2 \tag{2.2}$$

where y represents the anticipated value and ŷ, represents the actual obligatory incentive for i yields in result layer [37]. The predispositions and loads are refreshed by back-spreading across NN and updating their qualities based on slope drop computation in light of drop calculation. This computation deals with fractional first subsidiary regarding determined blunder in the previous layer for each weight and predisposition. The characteristics calculated from each subsidiary are after that deducted from the first value of that inclination or weight. The refreshed attributes ensure that the blunder in the result layer is reduced since the worth reported from the halfway subsidiary is positive in the case when the effect of that boundary creates the blunder. Furthermore, assuming the subsidiary's value is negative, extending the worth increases the blunder to the point where, when combined with deduction in the update condition, the new worth increases, lowering the error rate and further developing NN's expectations [38].To handle such contributions to a NN in previous and before NNs, two-layered sources of information, such as pictures, were converted into one-layered vectors. This modification results in a data deficiency in the two-layered information, in which highlights in one of the aspects are extended based on that aspect’s size. For example, in the case where pixels in an image are straightened equally and the columns of a picture are ordered close to one another, the elements created by vertical pixels are lost since distance between 2 adjoining pixels in an upward direction equals the picture's width. To overcome this problem, NNs are used to find highlights in two-layered data sources, allowing the neighborhood highlights to be found regardless of the direction of such elements in the data [39].

In the visual arrangement of humans, NNs are used for repeating the inferotemporal way. For distinguishing highlights, such networks use two-layered channels that are tangled throughout the entire image, with each channel identifying a specific component for each layer. The recognized

elements are also two-layered since the channels are 2-layered. This approach allows the separation of the highlights from 2-layered contribution without affecting input's state [40].

The channels in the layer are actually loads that are conveyed in 2-layered clusters and refreshed throughout dataset preparation, resulting in channels which have the ability of recognizing different highlights based on the sources of information and the assignment required by the NN, which is the primary difference from traditional PC vision approaches that detect explicit types of elements, such as corners and edges. Those are tangled throughout the entire picture, with the steps indicating the progression size of the development toward each path. Another exhibit is created by combining the load results with information values, with the size of new cluster that has been determined by the steps that have been defined for comparing layer [41].

Pooling layers that are divided to max-pooling and normal pooling layers, are one more important type of layer in the NNs. However, max-pooling layers have demonstrated a strong commitment to overfitting avoidance by reducing the size of one channel's results before transmitting the qualities to the next layer. Max-pooling layers have two-layered channels that are tangled as a result of the layer's contribution. Nonetheless, after the channels, max-pooling layers are used to create another cluster that has the most severe features identified in each maximum pooling channel. The steps that regulate the progression size that such channels move per convolution should be specified similarly to the size of the channels in these layers. Figure 2.9 shows an illustration of a maximum pooling layer with (2x2) channel size and (2x2) steps.

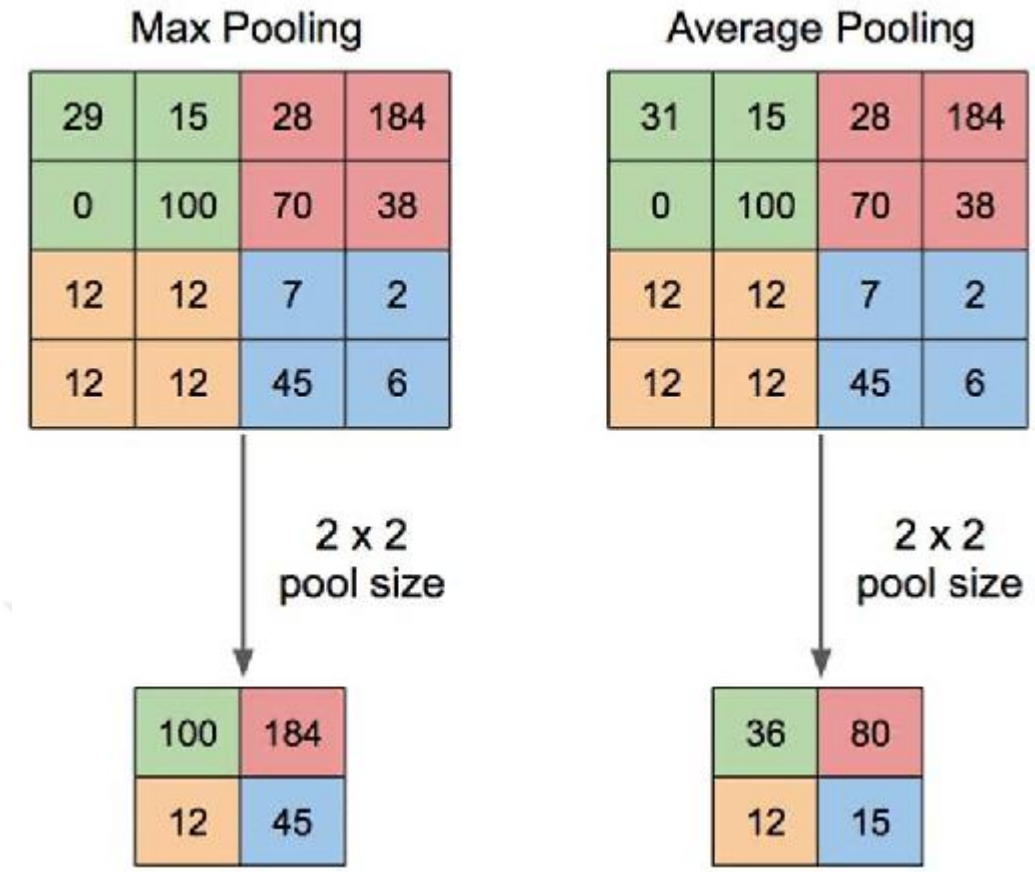


Figure 2.9: A example of the Output of a MaxPooling

The output of final layer, which could be a channel or a maximum pooling layer, is leveled to a single vector and completely that is associated with FFNN. In light of diverse blends of the channels in the layers, the neurons in the entirely linked layers may now distinguish advanced two-layered neighborhood highlights in the information. It's also possible to handle a three-layered contribution to NN, in which data is divided into layers of two-layered exhibits. The NN after that creates many channel layouts for each layer, with more profound levels gaining the ability to combine highlights from many layers [42].

2.3.3. Training Artificial Neural Networks

Like people's cerebrums, where geography of the organic NN and conductivities of the neuro-transmitters characterize the choices that have been made by the mind, the ANNs are additionally dependent upon the dispersion of the neurons and the loads among them for the purpose of settling on the required option. Two indistinguishable NNs may be used in entirely different tasks where different decisions are produced with the use of different load values amongst their neurons. The type of impact, the result of the neuron in the past layer over neuron in the following one, in addition to meaning of that impact on the result of neuron in the subsequent layer, is determined by the worth of a load between two neurons [43].Backpropagation is important in the ubiquity of NNs because it allows the network's exhibition to be substantially improved when this process has been utilized for the purpose of updating the value of the network's loads. For the purpose of refreshing loads w of an ANN, the back-propagation requires 3 characteristics, as can be seen from Eq. 2.3, which are the pace of progression of network's result, concerning misfortune being refreshed $\frac{\partial O}{\partial w}$, the mistake *E* between result of the network and the one that is really required from it and rate of learning L [44].

$$\widehat{w} = w - \frac{\partial O}{\partial w} \times E \times L \tag{2.3}$$

In spite of the type of error work that is used by an NN, like the cross-entropy and Mean Squared Error (MSE) capacities, such capacities compute one value that addresses the contrast between NN's result using current loading values and the qualities required from the network. The NN's result is acquired by using the forward pass of the NN to handle a cluster of the testing inputs from the preparation data-set, whilst the real results have been obtained directly from the preparation data-set or through handling the sources of information using pre-defined capacity. In backpropagation, the determined mistake esteem is after that used. However, because large blunder values could result in large delta values, a learning rate has been utilized for controlling delta values in lower ranges for loading refreshes. This delta values control ensures the avoidance of the detonating weight values, allowing the loads values which cause the basic mistake to be identified [45].

Through ascertaining pace of the progression of result blunder, concerning weight values, 3 possible characteristics may be produced, namely: A positive worth, demonstrating the fact that increasing one's weight self-esteem increases the error. To reduce the contrast between the NN and the necessary results, the weight esteem must be reduced by the work out delta esteem in this way.A negative worth, which is an indication of the fact that the error is lessened through the increase of weight's worth. To reduce the blunder worth and generate more accurate findings, the current weight esteem must be enlarged by the chosen delta esteem.A zero worth, which shows that no changes are needed to current weight value.The upsides of the loads in NN could be refreshed based on such potential qualities and using the recipe shown in Equation 2.3 to reduce the distinction between the NN's expectations and the actual result required to complete the ANN's assignment. However, in order to reduce the delta esteem employed for refreshing weight values, the optimal NN exhibition, which has been formed through limiting blunder via refreshing load values, working out optimal load values necessitates a number of emphases, such as ages [46].

2.3.4. Transfer Learning

The NN preparation is dependent upon misfortune that has been determined between qualities that are yielded by the neural network, using its present predispositions and loads, and results that are required for this information. Nevertheless, those boundaries cannot be quickly refreshed, to attempt at not missing world-wide least misfortune.in addition to that, as update begins from the yield layer to information layer, in other words, back-propagation, updates to boundaries of layers that are nearer to information layer are not significant. Which is why, the preparation of profound neural network necessitates massive assets and preparing tests for the purpose of accomplishing necessary presentations [47][48].

Layers that are closer to info layer are accountable for the identification of the crude high-lights, which will most likely be divided between different applications that NNs are used for. For instance, in a case where a Convolutional NN is required for perceiving face high-lights or distinctive mark includes, channels in layer which is closer to information layer could identify the comparative elements, instance upward, level and inclining lines. Blending those elements in after layer results in characterizing complicated high-lights which are more engaged with the accomplishment of necessary assignments. Anyway, throughout preparation, layers that are closer nearer to result layer

get quicker refreshes. As a result of that, the strategy of the preparation may be sped up essentially through starting with the lately perceived crude high-lights [49][50].nIn move learning, NN preparation does not start with no preparations, for instance, loads and inclinations aren't set in a random manner. In addition to that, NN construction, identified by number of the layers, neurons quantity in every one of the layers and the way through which layers are related to each other, could be differing, starting with one application then onto the following one, concerning the number of the elements that are to be recognized at every intricacy degree and the intricacy of elements that are to be identified to arrive at expected results. After that, a pre-trained neural network could be deployed for preparation, in spite of whether NN is prepared for comparative application, instead of indistinguishable application. Preparation uses crude highlights that neural network had definitely knew, just as making sure that pre-owned constructions are reasonable for applications. Thereby, move learning had resulted in permitting the critical decrease is preparing time at the same time as providing required presentations [51][52].

3. METHODOLOGY

3.1. PROPOPSED METHOD

Figure 3.1 displays the suggested procedure for identifying whether or not the condition of a patient is malignant. In this study, an iterative approach is provided. Collecting and preprocessing several photos from the colon disease dataset is the starting stage. In the pre-processing phase, image noise can be eliminated. Then, he enhanced the image using histogram balancing and other image processing techniques. This page contains only grayscale images with exposure values between 0 and 255. All fundamental elements of the suggested technique have been outlined. To correctly diagnose colon disease, it is vital to gather and analyze the pertinent information. Kvasir utilizes data collecting to amass images of situations relevant to his research. The Wavelet transform-spatial gray level dependence matrix (WT-SGLDM) [53] is utilized to begin extracting features from a CT scan image. The WT-SGLDM approach is applied in the suggested method in order to extract textural information from images. We extract the most pertinent information from the generated CT-Scan images to input into the multilayer artificial neural network. Because the feature selection problem is binary, only 0s and 1s can be used. A n-dimensional feature vector is composed of n attributes, each of which can take the values 0 or 1. In the suggested method, feature selection is represented by the number 1, while it is not represented by the number 0. Utilizing a feature vector and the group teaching optimization (GTO) algorithm, the proposed method is based on a feature vector. The GTO algorithm is incorporated into the recommended approach to feature selection because it is one of the few algorithms that can successfully imitate human learning behavior. Combining optimal global and local search, the GTO approach was first proposed in 2020 to improve the precision of meta-heuristic algorithms such as PSO and GA.

In contrast to other meta-heuristic approaches, this method actively seeks out both the optimal solution and promising candidates.

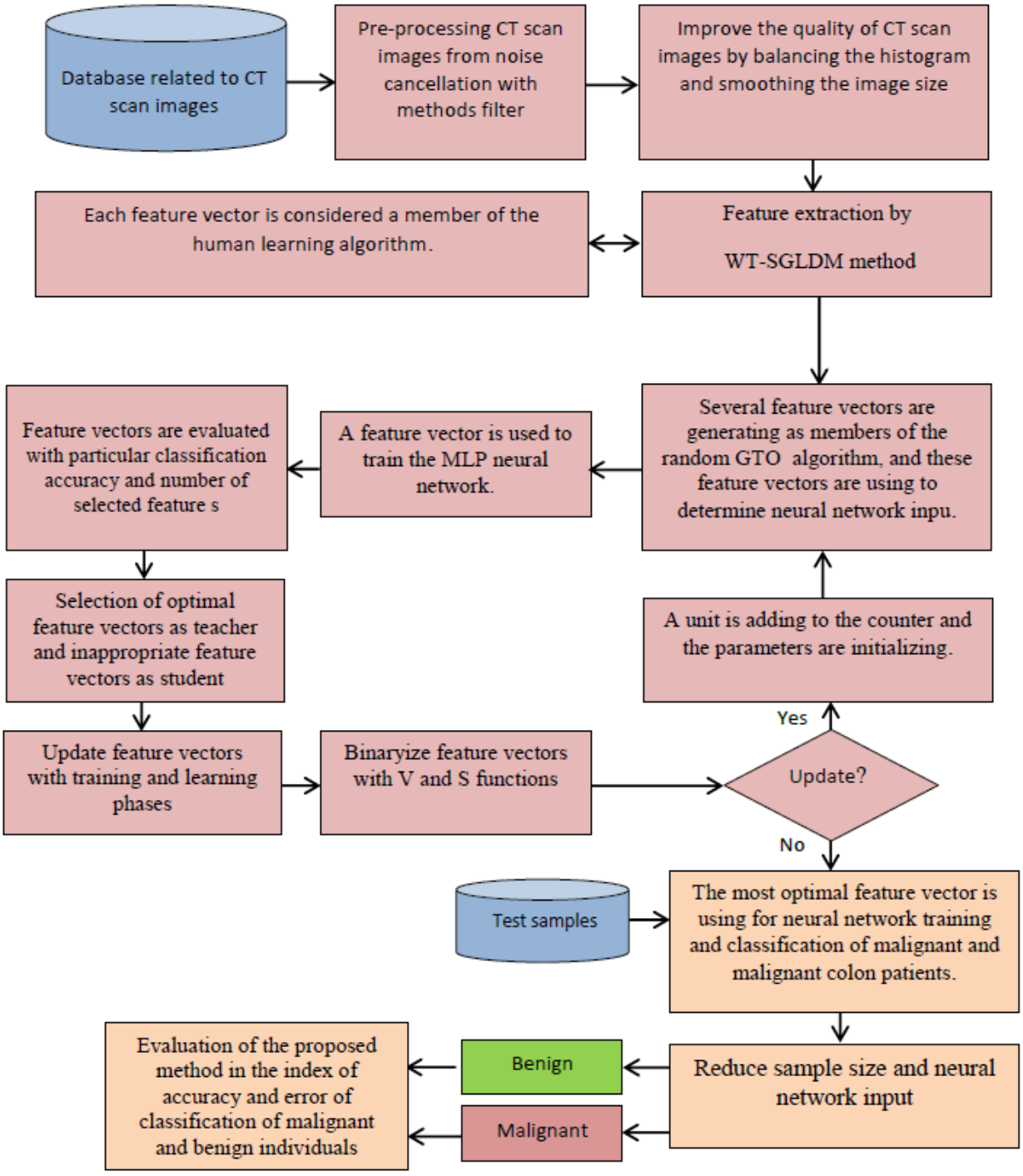


Figure 3.1: The structure of the suggested technique for diagnosing colon patients.

The subsequent step is to choose the most informative CT scan features and train an artificial neural network with multiple layers to differentiate between healthy and malignant tissue. A multilayer artificial neural network is used to minimize input dimensions and classify the data, so completing the suggested ideal feature vector method. In the third and final phase, test data

and samples are analyzed to evaluate the proposed approach. The following methods are advised for determining whether or not a picture from a colonoscopy depicts cancer:

i. The CT scan results for colon cancer are acquired After photos have been preprocessed, unique numbers are allocated to all differentiating features in preparation for machine learning .

ii. The third section examines the design and implementation of a multi-layered artificial neural network for classification problems.

iii. This initial evaluation of CT scan picture quality using the WT-SGLDM method for feature extraction.

iv. A single feature vector summarizes the most important characteristics of the dataset.

v. The GTO algorithm views each feature vector as a potential student, instructor, or peer.

vi. The initial random iteration of the GTO algorithm generates a set of feature vectors, which are subsequently utilized in subsequent iterations.

vii. Dimensionality reduction is applied to each feature vector linked with CT scan pictures of colon patients using the specified feature vector in order to train a neural network with fewer parameters. It is vital to establish a system for ranking each feature vector's quality. This method considers the normal imprecision of neural network output for a particular collection of attributes.

viii. At each tier, the best feature vector is chosen as the population teacher based on its capacity to distinguish between healthy and unhealthy individuals with high reliability.

ix. Joining the feature vectors and employing conversion functions to overcome this problem is part of the iterative improvement phase of the GTO approach.

x. At the conclusion of the training process, the optimal feature vector is used as input to the artificial neural network, and the network uses this information to distinguish between samples associated with people with colon cancer and healthy samples.

xi. The proposed method is evaluated and compared to competing methods using accuracy and classification error metrics in patients with benign and malignant colon conditions.

The remainder of this section introduces and describes in detail each step of the proposed method for distinguishing between benign and malignant individuals with colon disease.

3.2. PRE-PROCESSING IMAGES

The proposed method employs noise cancelling techniques to enhance the image quality of colon scans. In most research, noise is decreased using methods such as median and mean, however in this study, CT scan pictures are modified using interpolation. The proposed method utilizes an

interpolation mechanism to fill in image gaps in a variety of directions. By assessing surrounding information and examining the image from multiple viewpoints, this approach can estimate the number of central or noisy pixels. Figure 3.2 displays the general layout of the method for noise reduction that utilizes interpolation data from neighboring pixels. [54]:

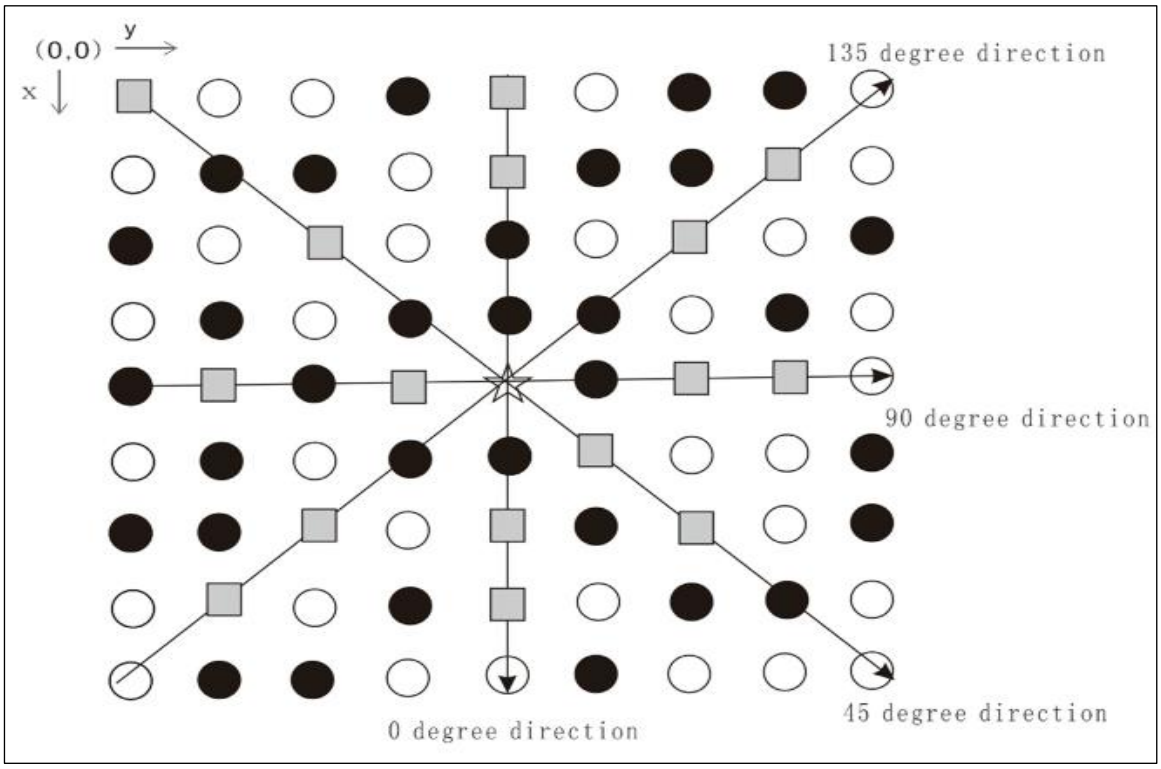


Figure 3.2: Noise removal with the concept of neighborhood in the method of interpolation of adjacent pixels

The proposed method locates pixels adjacent to a central pixel by looking in four directions at angles of 0, 45, 90, and 135 degrees. To insert the middle pixel, its four adjacent pixels on both sides must be utilized. Using their proposed method and the interpolation equation given in (3.1) [54], the value of the central pixel, p (x, y), is interpolated:

$$p(x,y) = c_0 R_0 + c_{45} R_{45} + c_{90} R_{09} + c_{135} R_{135} \tag{3.1}$$

$$c_i = \frac{der_i}{\sum_j \quad der_j}, \quad j = 0^{\circ}, 45^{\circ}, 90^{\circ}, 135^{\circ} \tag{3.2}$$

$$der_i = e^{-\left|\frac{r_l^i - r_r^i}{x_l^i - x_r^i}\right|}, \quad j = 0^{\circ}, 45^{\circ}, 90^{\circ}, 135^{\circ} \tag{3.3}$$

3.3. FEATURE EXTRACTION

By preprocessing colonoscopy images, it is possible to increase the accuracy of feature extraction. In this case, WT-SGLDM is used to extract features. The WT-SGLDM method can be used to identify photographic textures. Table 3.1 displays the results of a histological examination performed on a colon CT scan:

All of these characteristics can serve as inputs for more sophisticated neural networks. As part of the proposed method, a small subset of these 42 features is selected through group learning and then fed into a synthetic neural network.

3.4. FEATURE SELECTION

Table 3.1 depicts the returning features that were used to train the neural network; these features correspond closely to the image textures. Each of the 42 known characteristics of colonies is distinct from the others. During the process of feature selection, artificial neural networks must be trained on the image's essential properties in order to accurately identify between healthy and malignant images. This work provides a binary version of the GTO algorithm that may be used to identify the best suitable feature for the suggested approach. Consider the feature vector in Equation (which specifies the texture of the CT scan image) where each component has a value between 0 and 1 :

$$F_i^j = \{1\ F_i^j\ is\ selected\ \ 0\ F_i^j\ is\ deselect \tag{3.4}$$

As shown in Equation, the first step of the proposed method involves generating feature vectors, or the random population of the GTO algorithm:

$$Pop = \{[F_1^1, F_1^2, \dots, F_1^D], [F_2^1, F_2^2, \dots, F_2^D], \dots, [F_n^1, F_n^2, \dots, F_n^D]\} \tag{3.5}$$

N texture-related feature vectors were recovered from the scanned image, and D significant texture image properties were incorporated. nPop is the initial number of feature vectors before training begins. Figure 3.3 depicts the procedure and strategy used to train artificial neural networks to distinguish between normal and malignant tissue in photographs. Using a succession of feature vectors, the artificial neural network is trained. The suggested method uses feature vectors with any of the 42 features specified in Table 3.1 to determine the input type of the neural network. When training neural networks, only a subset of these characteristics is utilized. Each feature vector may have a distinct combination of ones and zeros. The artificial neural network

is taught how to behave using fragments of each attribute. The inputs of a multilayer neural network are updated during training by applying the selected features of a feature vector.

3.5. CLASSIFICATION

The "hidden layers" of the ANN contain several artificial neurons. Without intermediaries, the outputs of a synthetic neuron are transferred to the intended neuron. Multiplying the input's base values by the grid's weights, and then adding the product to the input. Weights are denoted by the symbol W, while biceps are represented by the letter B. By adjusting the weights and bays, the output of the neuron can be contained within a specified range, producing in an activation function such as a sigmoid, sigmoid tangent, ReLU, or Leaky ReLU activity function with deviation. Equations (3.6) and (3.7) depict the often employed sigmoid activity and sigmoid tangent functions in machine learning, as well as their respective application criteria:

$$f(x) = \frac{1}{1+exp^{-x}} \tag{3.6}$$

$$f(x) = \frac{exp^{x}-exp^{-x}}{exp^{x}+exp^{-x}} \tag{3.7}$$

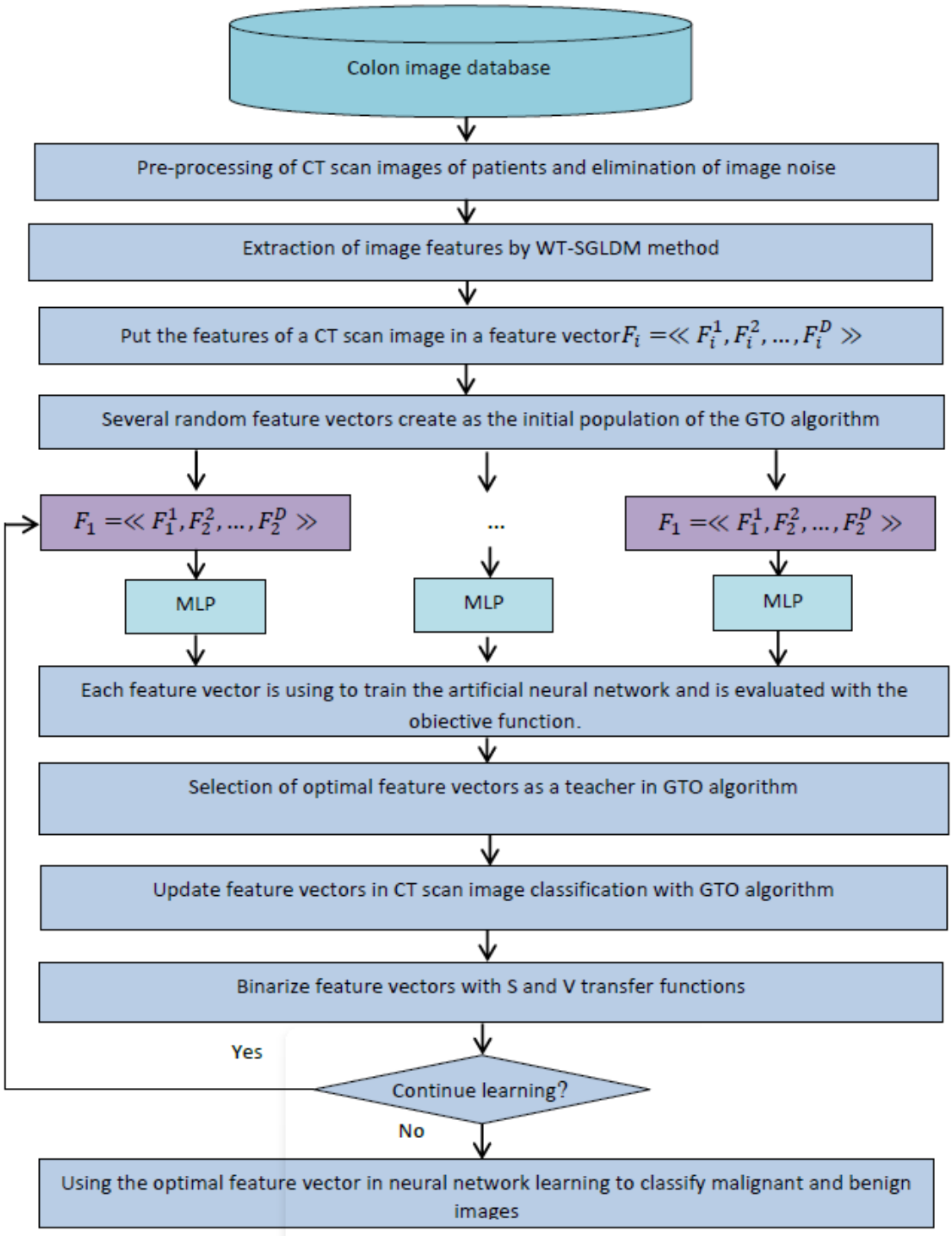


Figure 3.3: Steps for selecting features in malignant and benign images

Artificial neural network output can be affected by the activity function according to Equation (3.8):

$$\widetilde{f(X)} = f_{act}(W.X + b) \tag{3.8}$$

The error value of a sample in an artificial neural network calculates as Equation (3.9):

$$e_i = |\widetilde{f(X_i)} - f(X_i)| \tag{3.9}$$

The mean error in all tests or training samples consider as the objective function and formulate according to Equation (3.10).

$$e = \frac{\sum_{i=1}^{n} |\widetilde{f(X_i)} - f(X_i)|^2}{n} \tag{3.10}$$

3.6. OBJECTIVE FUNCTION

The number of features and the classification error are used to determine which feature vectors to employ in the MLP classification procedure. The proposed method for feature selection employs the appropriate objective function as equation (3.11):

$$Cost = \rho Error + \varphi \frac{|M|}{|N|} \tag{3.11}$$

| N | All possible features to classify malignant and benign images of colon patients. Simpler objective function express in equation (3.12):

$$Cost = \rho \frac{\sum_{i=1}^{N} (Class_i^{real} - Class_i^{est})^2}{N} + (1 - \rho) \frac{|M|}{|N|} \tag{3.12}$$

3.7. UPDATE FEATURE VECTORS

In this article, the GTO algorithm is used to update feature vectors. The group-training foundation of the GTO algorithm was inspired by Confucian thinking. According to Confucius, if you wish to educate someone effectively, you must first analyze their unique traits before you can provide them with the proper direction. As a result of the talents, they bring to the table, most nations now place a high value on youth education. This indicates that teachers should use real-world examples and data when designing classes and evaluating students' development. The GTO algorithm is based on ideas derived from human group learning. Examining the mean and standard deviation of a class or group's scores reveals that they are distributed normally. The normal distribution and equation can be used to characterize results on standardized testing (3.13):

$$f(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(X-\mu)^2}{2\sigma^2}} \tag{3.13}$$

The mean, standard deviation, and final grade for a student in a classroom context are correspondingly denoted by 2, and X. The TLBO algorithm only examines one population type, in which a teacher is the best option. Having a greater degree of education increases one's possibilities. Figure 3.4 is a graphical depiction of the GTO method. This program distinguishes between typical students and those who merit praise. People in these groups frequently consider their professors to be the most deserving members of society. Every educator can instruct a class successfully on their own. At each level, students can collaborate and learn from one another. Figure 4.5 and the graph to the left illustrate the statistical distribution of scores or skills derived from this method. The best performer has a favorable opinion of the instructor and the recommended solution.

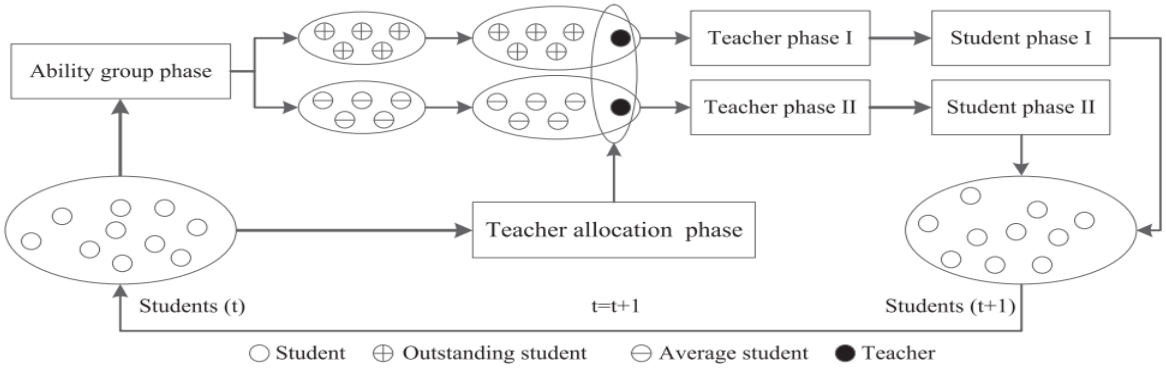


Figure 3.4: Teaching and learning mechanism in GTO algorithm

According to grade level, the GTO population can be categorized as either less qualified students or average kids. The picture illustrates the distribution of grades for two groups of students: those with average qualifications and those with higher qualifications. In general, the dispersion of grades among students with average talents is lower than that of those with superior abilities.

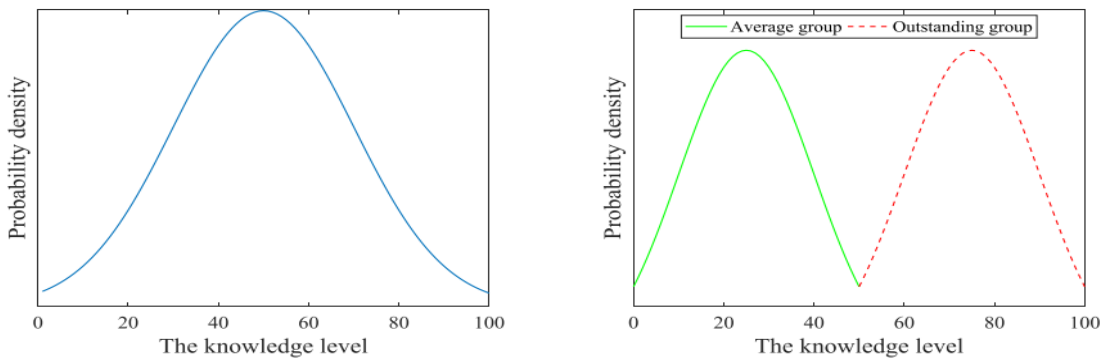


Figure 3.5: Left to right, score distribution in TLBO algorithm and GTO algorithm, respectively

A proposed approach allows a feature vector to represent either a professor or a student. Based on objective criteria, each attribute vector is rated as either intelligent or moderate. If an attribute vector can reduce the objective function further, it should be referred to as a learning instruction. Several ways of updating sample-defining feature vectors are given in order to improve the categorization of multilayer artificial neural networks in the diagnosis of human malignancy and health.

3.8. PRINCIPLE COMPONENT ANALYSIS

PCA may be applied to provide insight on a set of statistics utilizing contemporary, uncorrelated data ("additives"). This strategy is useful for reducing the dimensionality of a data set since the additives are rated based on how much genuine variance they describe. In technical terms, the PCA seeks a projection that results in the expression of the data with the fewest squares. This converts a specified collection of likely correlated observations into a defined set of nonlinearly correlated variables, known as dominating additives. Common uses of principal component analysis include predictive modeling and exploratory data analysis (PCA). After centering the data on the mean of each feature, the principal component analysis (PCA) entails calculating the decomposition of the covariance matrix into eigenvalues. Although it bears certain formal similarities with factor analysis and can be used as a proxy for factor extraction within that framework, the two should be maintained distinct.

Principle component analysis (PCA) creates a linear transformation that selects a new coordinate system for a particular data collection, with the biggest variance of the data set represented on the primary axis (the first principal component), and so on. Before generating the linear transformation, the covariance matrix, also known as the correlation matrix, must be constructed. Due to its symmetry, this matrix possesses an exhaustive collection of eigenvectors. In order to reduce the dimensionality of the statistics, there must be a linear transformation from the old base coordinates to the new base coordinates. In addition, the new base coordinates reflect the early statistics' important makeup.

PCA is especially beneficial for lowering the overall number of dimensions in a dataset. The first few main components explain a substantial proportion of the observed variation (more the more correlated the original variables were). These low-order components may contain the "most essential" information, while the others can be disregarded. There are a number of estimation methods available; the optimal method will rely on the correlation pattern of the original data.

3.8.1. PCA Mathematics

Consider the likelihood of recognizing a trend when n individuals are involved and m random variables are measured for each individual. Using PCA, we can identify a group of underlying variables, p m, that collectively provide an approximate explanation for the values of m variables. The existence of these p underlying elements might be regarded as a reduction in the dimension of the records, as we no longer require m values for each character. Each p in the data collection is handled as a first-order notion during the invocation of the process. PCA can be utilized in the following key ways:

i. When the data are not dimensionally homogeneous or when random variables are not quantified in the same order of importance, a correlation matrix-based method is employed.

ii. When statistical data has the same median value and is dimensionally homogeneous, a technique based on the covariance matrix is employed..

3.8.2. Correlation-Based Method

The method starts from the correlation matrix, let's consider the value of each of the m random variables F_j. For each of the n individuals, let's take the value of these variables and write the data set in matrix form:

$$(F_j^{\beta})_{j=1,\dots,m}^{\beta=1,\dots,n} \tag{3.21}$$

Note that each set

$$\mathcal{M}_j = \{F_j^{\beta} | \beta = 1, \dots, n\} \tag{3.22}$$

can be considered a random sample for the variable F_j. From the $m \times n$ data corresponding to the m random variables, the sample correlation matrix can be constructed, which is defined by:

$$\mathbf{R} = [r_{ij}] \in M_{m\times m}, \tag{3.23}$$

$$r_{ij} = \frac{\mathrm{cov}(F_i, F_j)}{\sqrt{\mathrm{var}(F_i)\mathrm{var}(F_j)}} \tag{3.24}$$

Since the correlation matrix is symmetric then it is diagonalizable and its eigenvalues λ_i, verify

$$\sum_{i=1}^{m} \lambda_i = m \tag{3.25}$$

Because of the previous belongings, these m eigenvalues are known as the weights of each of the m important components. the principle elements diagnosed mathematically are represented with the aid of the eigenvector base of the matrix **R**. it is clear that each of the variables can be expressed as a linear combination of the eigenvectors or principal components.

3.8.3. Method Based On Covariances

The objective is to map a set of nxm statistics X to a set of nxl statistics Y with minimal information loss using the covariance matrix.

Ideally, we would like to be able to characterize each of these samples with no more than I variables, where l is the maximum number of variables permitted. The starting point is a fixed set of n samples with m describing variables. In addition, l, the required component count, must be significantly smaller than the smallest X dimension.

$$l \leq min\{n, m\} \tag{3.26}$$

The data for the analysis have to be centered at mean 0 (subtracting the mean of each column) and / or autoscaled (centered at mean 0 and dividing each column by its standard deviation).

$$\mathbf{X} = \sum_{a=1}^{l} \mathbf{t}_a \mathbf{p}_a^T + \mathbf{E} \tag{3.27}$$

Vectors $\mathbf{t}_a$They are known as scores and contain the information on how the samples are related to each other, and they also have the property of being orthogonal. Vectors $\mathbf{p}_a$They are called loadings and they report the relationship between the variables and have the quality of being orthonormal. By taking fewer principal components than variables and due to the fit error of the model with the data, an error is produced that accumulates in the matrix **E**.

PCA is based on the eigenvector decomposition of the covariance matrix. Which is calculated with the following equation:

$$\text{cov}(X) = \frac{X^T X}{n-1} \tag{3.28}$$

$$cov(X)\boldsymbol{P}_a = \lambda_a \boldsymbol{P}_a \tag{3.29}$$

$$\sum_{a=1}^{m} \lambda_a = 1 \tag{3.30}$$

Where λ_a is the eigenvalue associated with the eigenvector $\mathbf{p}_a$. By last,

$$\boldsymbol{t}_a = X\boldsymbol{P}_a \tag{3.31}$$

We can understand this equation as $\mathbf{t}_a$ are the projections of X in $\mathbf{p}_a$, where the eigenvalues λ_a they measure the amount of variance captured, that is, the information represented by each of the principal components. The amount of information that each main component captures decreases according to its number, that is, main component number one represents more information than two and so on.

3.8.4. Limitations

The implementation of the PCA is limited by a number of assumptions. linearity assumption: it is assumed that the observed data are linear combination of a certain base

4. RESULTS

Colon cancer patients undergo CT scans. Using the indicators from the proposed method, the images are then classified as malignant or benign.

4.1. DATA SET

This study employs endoscopic pictures captured with a variety of lighting configurations and magnification settings. This analysis made use of two data sets. Gill Hospital, a research center in South Korea, has categorized its COLORECTAL DATA into five categories. This dataset has a total of 3,515 photos distributed over its five categories. 770 photos are classed as normal, 634 images as adenocarcinoma, 775 images as adenomas, 563 images as having Crohn's disease, and 773 images as having ulcerative colitis [55]. As a secondary source of data, KVASIR is indispensable to our investigation. Each of the eight categories within the KVASIR dataset consists of 500 photographs [56].

4.2. EVALUATION

In this analysis, we categorize photographs based on their precision, sensitivity, and F1 score. According to criteria (4.1), (4.2), (4.3), and (4.4), the indicators are evaluated [57][58][59][60][61]:

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (4.1)$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (4.2)$$

$$precision = \frac{TN}{TN+FP} \times 100\% \quad (4.3)$$

$$F1 = 2 \times \frac{Recall \times precision}{Recall+precision} \times 100\% \quad (4.4)$$

4.3. ANALYSIS

In this study, the proposed method is assessed using a multilayer neural network in the MLP style. When selecting five hidden levels, the following will occur: Each layer contains a total of 42 hidden neurons. The purpose of these settings is to improve the results of five layers. There are 42 unique features in this image, and the same number of secondary features are concealed in the neural

network. The approach under consideration employs a population size of 20 and 50 iterations. In this article, Matlab (version 2018) is used for examples.

4.3.1. Objective Function Analysis

The technique's effectiveness can be measured in part by calculating the objective function of the value of the feature selection. A typical diagram of an objective or feature selection function in digital pictures is depicted in Figure 4.1. A data study of the objective function for feature selection demonstrates that the GTO method is utilized less frequently than other approaches. The decrease in error rate indicates that GTO chose the ideal feature vector for the neural network, enabling more accurate picture recognition. Two analytic factors for the tendency of the objective function to worsen are identified. The first justification is that traits are shrinking over time. The second key component is a drop in incorrectly identified photos of colon illness.

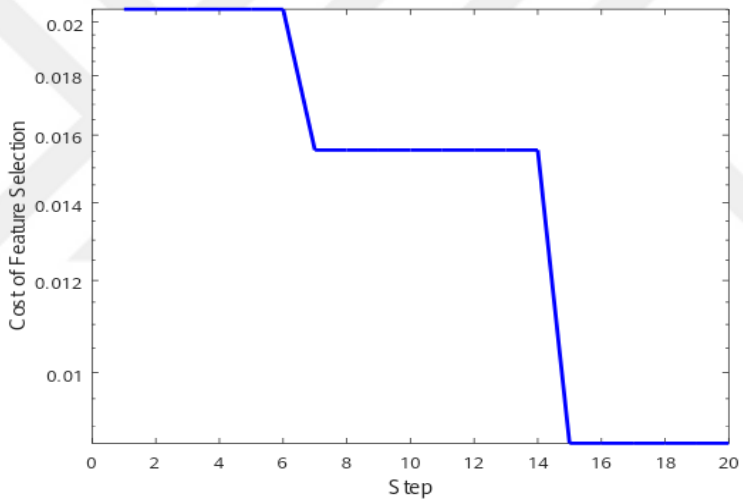


Figure 4.1: Reduction of the feature selection function in terms of GTO iterations

Figure 4.2 depicts the objective function value for feature vectors containing 10, 15, 20, and 25 attributes:

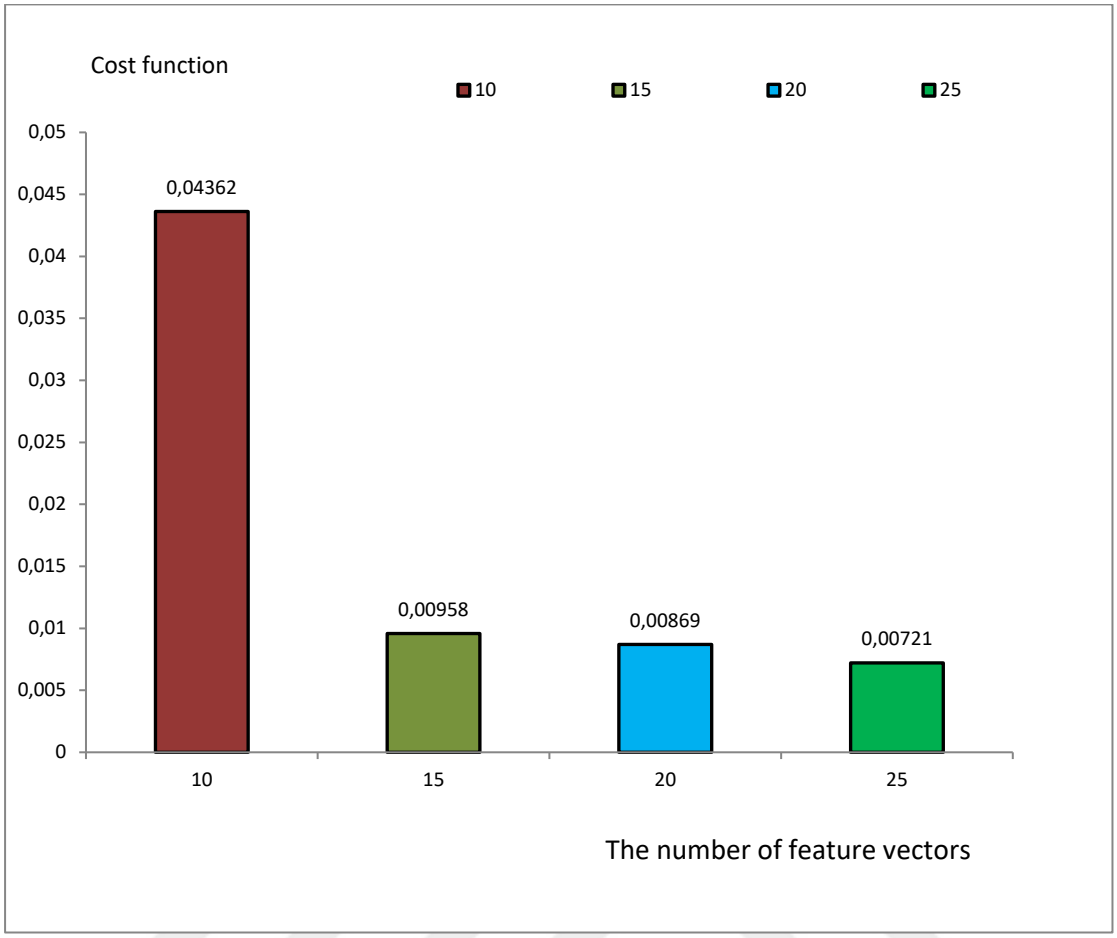


Figure 4.2: Reduction in the complexity of the GTO algorithm's feature selection function

The research demonstrates that the chance of locating the ideal feature vector rises as the number of feature vectors increases and the goal function falls from 0.04362% to 0.00721%. The drop is now 6,04 times smaller than it ever was.

4.3.2. Classification Accuracy Analysis

The classification accuracy of colon patient images utilizing the GTO algorithm has been found to be consistent across studies (Figure 4.3). On the basis of the collected and analyzed data, it can be stated that the classification accuracy of photos improves with each cycle. The GTO algorithm's deliberative selection of the feature vector contributes to its enhanced accuracy. The GTO method selects the best feature vector for use in developing artificial neural networks. The primary objective of the feature selection technique is to select the most prominent characteristics from images of colon patients for the neural network to learn and train on. Between the first and last repetitions, the accuracy of this experiment ranges from 91.61 to 94.51%. As the number of feature vectors in the

colorectal dataset increases, the dataset's overall accuracy will increase by an average of 94.36 percent if the trials are repeated 50 times.

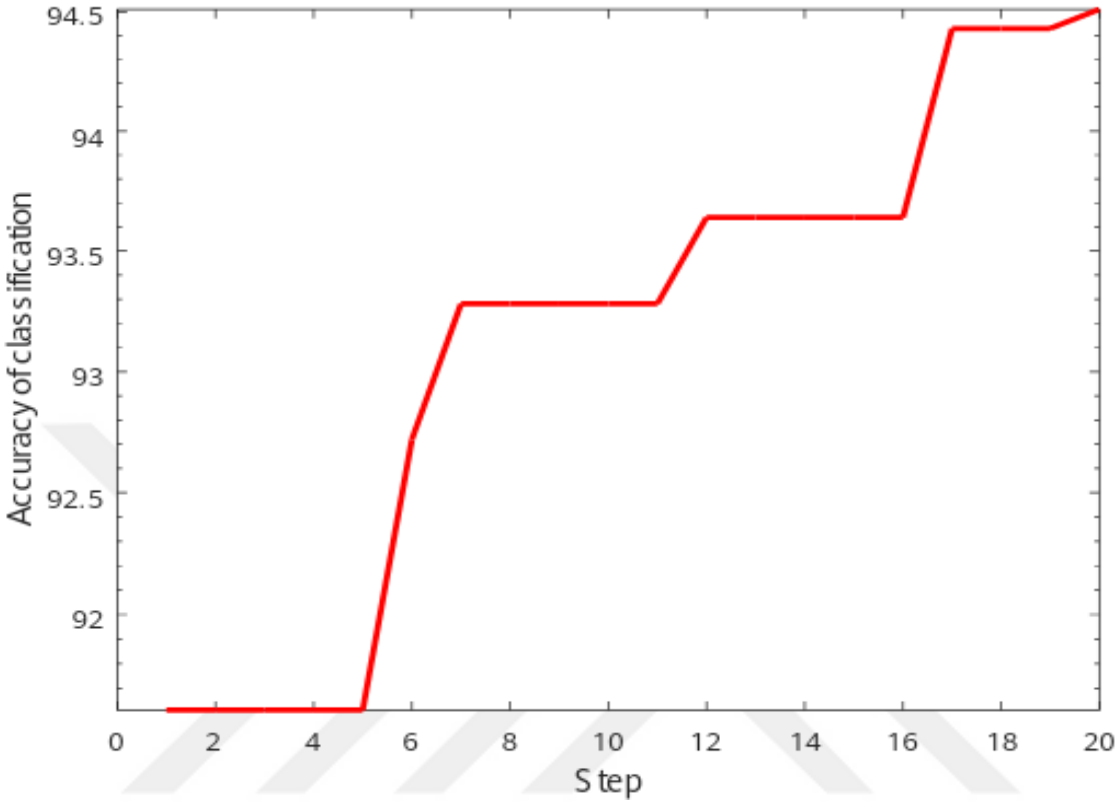


Figure 4.3: Increasing the GTO algorithm's iterative cycle accuracy for image classification.

4.3.3. Comparison

Following the selection of the colorectal dataset, the proposed method is applied and its efficacy is visually assessed. This data collection evaluates five distinct forms of colitis, including adenocarcinoma, adenoma, Crohn's disease, normal, and ulcerative. In terms of accuracy, sensitivity, and the F1 index, Table 4.1 illustrates the results of a comparison between the suggested method and other similar methods. In Figure 4.4, we can see how the proposed method compares to the other methods in terms of the mean values for precision, sensitivity, and the F1 index for five unique classes.

Table 4.1: Comparison between the proposed method's mean precision, sensitivity, and F1 indices and other methods.

method	precision	sensitivity	F1
R.Zhang et al.[62]	83	87	85
Y.Shin and I.Balansingham [63]	83	84	83
W.Ryan et al. [64]	89.4	89	89
Poudel, S et al. [55]	93.2	92.8	93
Proposed Method	92.72	93.14	94.26

According to the results, the F1 index, sensitivity, and average precision for the suggested method were 92.72 percent, 93.14 percent, and 94.2 percent, respectively. In terms of analyzing and classifying photographs of colon patients, the proposed technique surpasses the existing methods in the study [62], [63], [64], and [55] due to its greater sensitivity index.

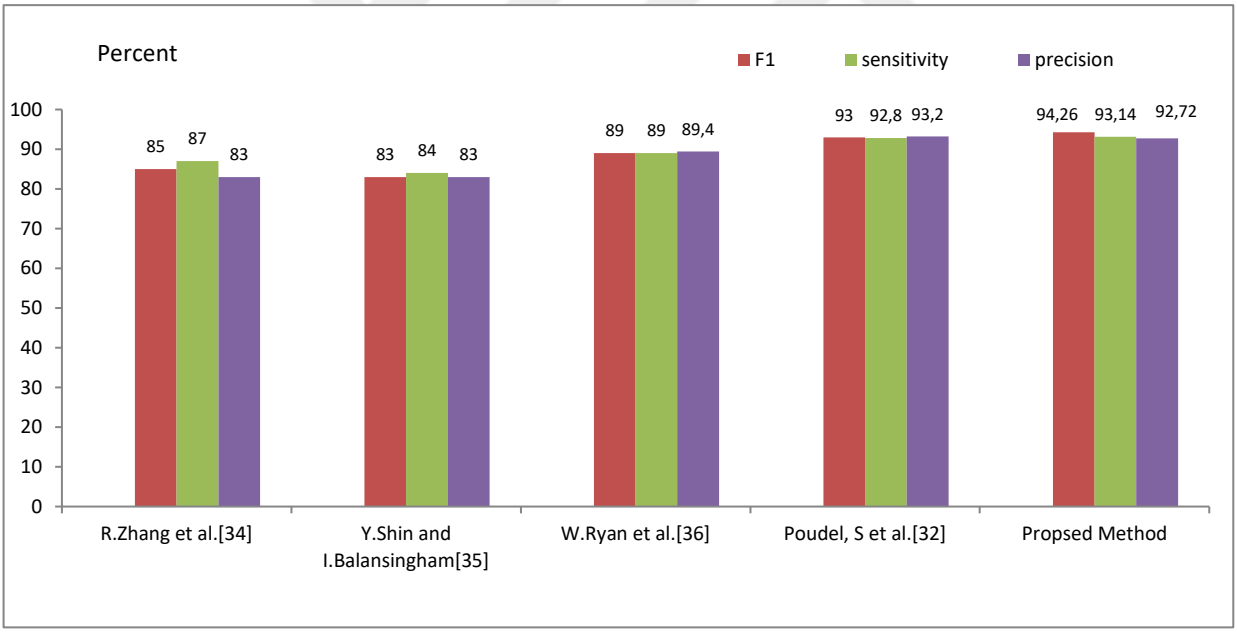


Figure 4.4: The proposed method's sensitivity, mean precision index, F1, and comparable techniques

Favorize the offered method due to its benefits. The reliability index for the article method is [55], or 93.2%. 92.72 percent, according to the specified approach, is the value of this indicator. The proposed technique was also evaluated using the Kvasir data set. The results of this study are compared to those of previous studies in the same field in Table 4.2 and Figure 4.5:

Table 4.2: Comparison between the proposed method's average accuracy, precision, sensitivity, and F1 index, as well as other approaches

method	accuracy	precision	sensitivity	F1
6Layer CNN [55]	91.4	66.1	64	65.1
3Layer CNN [55]	95.9	58.9	40.8	45.3
TFL [55]	92.4	69.8	68.9	69.3
Random Forest [55]	92.8	71.3	71.5	71.1
DropBlock CNN [55]	95.7	86.8	92.2	88
Proposed	96.42	88.62	92.69	89.54

In classifying images of colon patients, the proposed method takes into account the accuracy, precision, sensitivity, and F1 of methods such as 6Layer CNN, 3Layer CNN, TFL, Random Forest, and CNN based on a comparison with other methods. DropBlock labels photographs of colon illnesses more precisely.

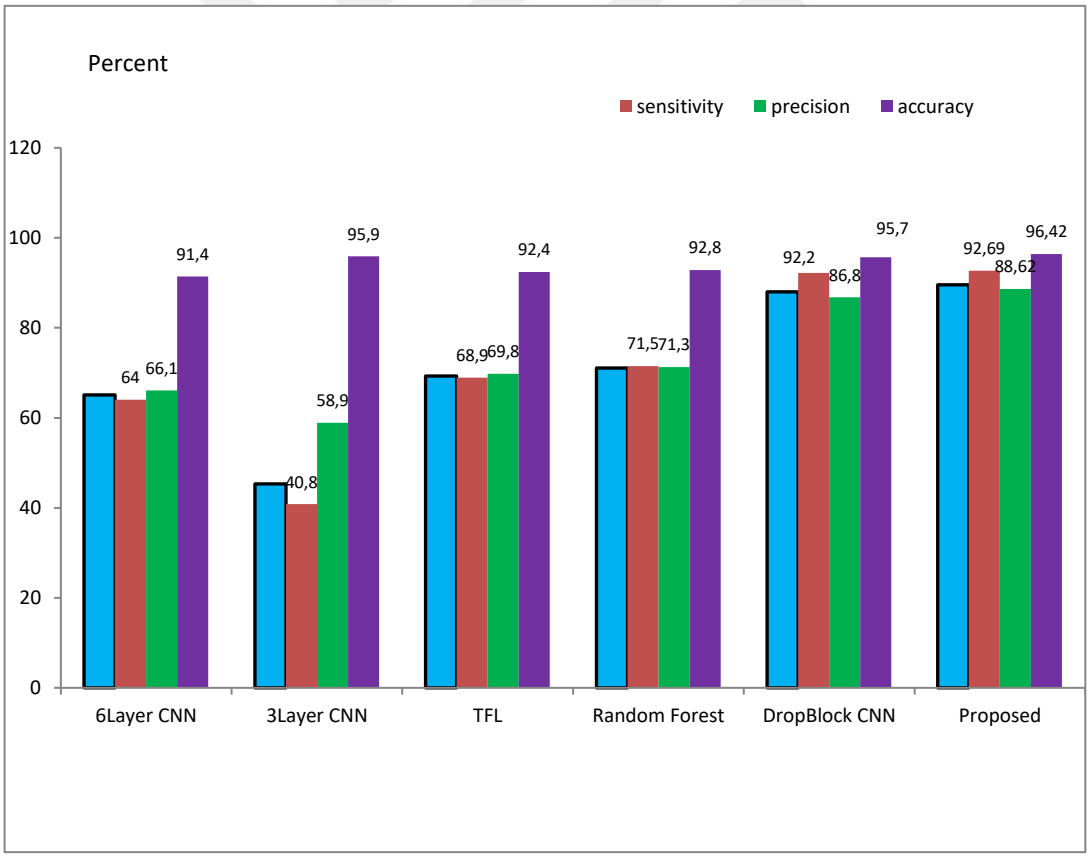


Figure 4.5: The proposed method's mean index of accuracy, precision, sensitivity, and F1 compared to competing methods

When classifying photos from the Kvasir data set, the suggested method was successful 96.42 percent of the time, 88.62 percent of the time, 92.6 percent of the time, and 89.5 percent of the time, respectively. If implemented as suggested, the CNN DropBlock approach delivers the highest classification accuracy, according to the research. The CNN 6Layer approach has the worst classification performance, according to the accuracy index, with an accuracy of approximately 91.4%. The proposed method ranks highest according to an accuracy index based on picture categorization. CNN 3Layer performs the poorest in this benchmark. The recommended approach outperforms CNN DropBlock, which has the highest sensitivity index value. CNN's 3Layer method has the smallest sensitivity index. On the F1 index, the suggested method surpasses all previous methods by a substantial margin, whereas the 3Nayer CNN method performs the poorest.

5. CONCLUSION

5.1. CONCLUSIONS

Colon cancer is a global killer. Diet and activity changes are increasing colon disease deaths. Blood tests can detect colon cancer, while colonoscopies and CT scans can detect colon illness. Doctors must examine many photos to comprehend colon illness. Individual treatment techniques for colon disease patients may be guided by classifications that consider etiology and severity. In this research, we classify colon illness images. Method involves steps. The suggested solution employs neighboring pixels to decrease photo noise. WT-SGLDM detects colon disease via image feature extraction. The GTO algorithm then chooses relevant features. GTO's photo selection criteria are used to train multilayer neural networks. MATLAB experiments with Kvasir and o color data. The proposed method achieves 92.72 percent accuracy, 93.14 percent sensitivity, and 94.26% F1 index when applied to color data. The suggested technique achieved 96.42% accuracy, 88.62% precision, 92.69% sensitivity, and 89.54% F1 index on the Kvasir data set. The recommended strategy beats recent machine learning-based experiments in precision, sensitivity, and F1 index. The proposed technique outperforms 3Nayer CNN, TFL, Random Forest, and DropBlock. CNN. The proposed method uses visual attributes for categorization and doesn't require time-consuming training like deep learning. The approach reliably sorts colon pictures into benign and cancerous categories. This method has flaws, like others. Inaccurate classification of some samples and feature selection time are some limitations of the proposed method. In a future study, we'll use video processing on colonoscopic pictures to identify damaged colon tissues.

5.2. FUTURE WORK

For Future work we can work on the metaheuristic optimization method such as Harris hawks optimization, Genetic Algorithm and gray wolf optimization method to improve the percentage of the accuracy.

REFERENCES

[1] D. Zhang, G. Huang, Q. Zhang, J. Han, J. Han, and Y. Yu, "Cross-modality deep feature learning for brain tumor segmentation," *Pattern Recognit.*, vol. 110, p. 107562, 2021.

[2] A. Ashraf, S. Naz, S. H. Shirazi, I. Razzak, and M. Parsad, "Deep transfer learning for alzheimer neurological disorder detection," *Multimed. Tools Appl.*, vol. 80, no. 20, pp. 30117–30142, 2021.

[3] J. Burggraaff *et al.*, "Manual and automated tissue segmentation confirm the impact of thalamus atrophy on cognition in multiple sclerosis: A multicenter study," *NeuroImage Clin.*, vol. 29, p. 102549, 2021.

[4] Y. Gonzalez *et al.*, "Semi-automatic sigmoid colon segmentation in CT for radiation therapy treatment planning via an iterative 2.5-D deep learning approach," *Med. Image Anal.*, vol. 68, p. 101896, 2021.

[5] Y. Shang *et al.*, "Pharmaceutical immunoglobulin G impairs anti-carcinoma activity of oxaliplatin in colon cancer cells," *Br. J. Cancer*, vol. 124, no. 8, pp. 1411–1420, 2021.

[6] L. Zhang *et al.*, "Single-cell analyses inform mechanisms of myeloid-targeted therapies in colon cancer," *Cell*, vol. 181, no. 2, pp. 442–459, 2020.

[7] F. Grass *et al.*, "Impact of delay to surgery on survival in stage I-III colon cancer," *Eur. J. Surg. Oncol.*, vol. 46, no. 3, pp. 455–461, 2020.

[8] N. A. Javan, A. Jebreili, B. Mozafari, and M. Hosseinioun, "Classification and Segmentation of Pulmonary Lesions in CT images using a combined VGG-XGBoost method, and an integrated Fuzzy Clustering-Level Set technique," *arXiv Prepr. arXiv2101.00948*, 2021.

[9] P. Achilli *et al.*, "Survival impact of adjuvant chemotherapy in patients with stage IIA colon cancer: analysis of the National Cancer Database," *Int. J. Cancer*, vol. 148, no. 1, pp. 161–169, 2021.

[10] G. Tulum, O. Osman, B. Bolat, Ö. Dandin, T. Ergin, and F. Cüce, "Colonic Polyp Classification Using Projection Image and Convolutional Neural Network," in *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*, 2019, pp. 1–4.

[11] X. Jia, X. Xing, Y. Yuan, L. Xing, and M. Q.-H. Meng, "Wireless capsule endoscopy: A new tool for cancer screening in the colon with deep-learning-based polyp recognition," *Proc. IEEE*, vol. 108, no. 1, pp. 178–197, 2019.

[12] S. Ahmed, M. Frikha, T. D. H. Hussein, and J. Rahebi, “Face Recognition System using Histograms of Oriented Gradients and Convolutional Neural Network based on with Particle Swarm Optimization,” in *2021 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, 2021, pp. 1–5.

[13] S. Ahmed, M. Frikha, T. D. H. Hussein, and J. Rahebi, “Optimum Feature Selection with Particle Swarm Optimization to Face Recognition System Using Gabor Wavelet Transform and Deep Learning,” *Biomed Res. Int.*, vol. 2021, 2021.

[14] A. K. Jerebko, J. D. Malley, M. Franaszek, and R. M. Summers, “Support vector machines committee classification method for computer-aided polyp detection in CT colonography1,” *Acad. Radiol.*, vol. 12, no. 4, pp. 479–486, 2005.

[15] C. Viji, N. Rajkumar, S. T. Suganthi, K. Venkatachalam, and S. Pandiyan, “An improved approach for automatic spine canal segmentation using probabilistic boosting tree (PBT) with fuzzy support vector machine,” *J. Ambient Intell. Humaniz. Comput.*, vol. 12, no. 6, pp. 6527–6536, 2021.

[16] F. Pagès *et al.*, “International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study,” *Lancet*, vol. 391, no. 10135, pp. 2128–2139, 2018.

[17] Y. Zhang and Z. Jin, “Group teaching optimization algorithm: A novel metaheuristic method for solving global optimization problems,” *Expert Syst. Appl.*, vol. 148, p. 113246, 2020.

[18] N. D’Souza *et al.*, “Ex vivo specimen MRI and pathology confirm a rectosigmoid mesenteric waist at the junction of the mesorectum and mesocolon,” *Color. Dis.*, vol. 22, no. 2, pp. 212–218, 2020.

[19] O. A. Catalano *et al.*, “Improving staging of rectal cancer in the pelvis: the role of PET/MRI,” *Eur. J. Nucl. Med. Mol. Imaging*, vol. 48, no. 4, pp. 1235–1245, 2021.

[20] S. Richards *et al.*, “Haploinsufficiency of Casitas B-Lineage Lymphoma Augments the Progression of Colon Cancer in the Background of Adenomatous Polyposis Coli Inactivation,” *Am. J. Pathol.*, vol. 190, no. 3, pp. 602–613, 2020.

[21] C. Luo, S. Cen, G. Ding, and W. Wu, “Mucinous colorectal adenocarcinoma: clinical pathology and treatment options,” *Cancer Commun.*, vol. 39, no. 1, pp. 1–13, 2019.

[22] R. Ataka *et al.*, “Eosinophilic peritonitis with colon cancer: a case report,” *BMC Gastroenterol.*, vol. 20, no. 1, pp. 1–7, 2020.

[23] M. Masud, N. Sikder, A.-A. Nahid, A. K. Bairagi, and M. A. AlZain, “A machine learning approach to diagnosing lung and colon cancer using a deep learning-based classification framework,” *Sensors*, vol. 21, no. 3, p. 748, 2021.

[24] D. Daye *et al.*, “Quantitative tumor heterogeneity MRI profiling improves machine learning–based prognostication in patients with metastatic colon cancer,” *Eur. Radiol.*, vol. 31, no. 8, pp. 5759–5767, 2021.

[25] L. Qi *et al.*, “Identification of prognostic spatial organization features in colorectal cancer microenvironment using deep learning on histopathology images,” *Med. Omi.*, vol. 2, p. 100008, 2021.

[26] B. Zhao *et al.*, “Using machine learning to construct nomograms for patients with metastatic colon cancer,” *Color. Dis.*, vol. 22, no. 8, pp. 914–922, 2020.

[27] I. Pacal, D. Karaboga, A. Basturk, B. Akay, and U. Nalbantoglu, “A comprehensive review of deep learning in colon cancer,” *Comput. Biol. Med.*, vol. 126, p. 104003, 2020.

[28] M. Arif Wirawan, A. Ummi, W. A. A. Wahyu Andi Saputra, F. M. Wibowo, S. Kom, and F. M. W. M Eng, “Maturity Differentiation of Colon Polyp Based on Endoscopic Images Using Machine Learning and Spatial Feature,” 2021.

[29] J. Fernández-Montoya, C. Avendaño, and P. Negredo, “The glutamatergic system in primary somatosensory neurons and its involvement in sensory input-dependent plasticity,” *Int. J. Mol. Sci.*, vol. 19, no. 1, p. 69, 2018.

[30] S. K. Esser, R. Appuswamy, P. Merolla, J. V Arthur, and D. S. Modha, “Backpropagation for energy-efficient neuromorphic computing,” *Adv. Neural Inf. Process. Syst.*, vol. 28, pp. 1117–1125, 2015.

[31] M. Turčaník and M. Javurek, “Hash function generation by neural network,” in *2016 New Trends in Signal Processing (NTSP)*, 2016, pp. 1–5.

[32] M. Turčaník, “Hash function generation based on neural networks and chaotic maps,” in *2017 Communication and Information Technologies (KIT)*, 2017, pp. 1–5.

[33] N. Abdoun, S. El Assad, R. Assaf, O. Déforges, M. Khalil, and S. Belghith, “Design and implementation of robust Keyed Hash functions based on Chaotic Neural Network,” 2018.

[34] B. Karlik and A. V. Olgac, “Performance analysis of various activation functions in generalized MLP architectures of neural networks,” *Int. J. Artif. Intell. Expert Syst.*, vol. 1, no. 4, pp. 111–122, 2011.

[35] B. Xu, N. Wang, T. Chen, and M. Li, “Empirical evaluation of rectified activations in convolutional network,” *arXiv Prepr. arXiv1505.00853*, 2015.

[36] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: a simple way to prevent neural networks from overfitting,” *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.

[37] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv Prepr. arXiv1412.6980*, 2014.

[38] X. Glorot and Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks,” in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, 2010, pp. 249–256.

[39] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[40] H.-C. Shin *et al.*, “Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning,” *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1285–1298, 2016.

[41] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[42] K. R. Konda, “Unsupervised Relational Feature Learning for Vision.” Univ.-Bibliothek Frankfurt am Main, 2016.

[43] I. K. Sethi and A. K. Jain, *Artificial neural networks and statistical pattern recognition: old and new connections*. Elsevier, 2014.

[44] R. Hecht-Nielsen, “Theory of the backpropagation neural network,” in *Neural networks for perception*, Elsevier, 1992, pp. 65–93.

[45] R. Miikkulainen *et al.*, “Evolving deep neural networks,” in *Artificial intelligence in the age of neural networks and brain computing*, Elsevier, 2019, pp. 293–312.

[46] J. Schmidhuber, “Deep learning in neural networks: An overview,” *Neural networks*, vol. 61, pp. 85–117, 2015.

[47] L. Torrey and J. Shavlik, “Transfer learning,” in *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*, IGI global, 2010, pp. 242–264.

[48] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, “A survey on deep transfer learning,” in *International conference on artificial neural networks*, 2018, pp. 270–279.

[49] H. Ravishankar *et al.*, “Understanding the mechanisms of deep transfer learning for medical images,” in *Deep learning and data labeling for medical applications*, Springer, 2016, pp. 188–196.

[50] X. Yin, X. Yu, K. Sohn, X. Liu, and M. Chandraker, “Feature transfer learning for face recognition with under-represented data,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 5704–5713.

[51] K. Gopalakrishnan, S. K. Khaitan, A. Choudhary, and A. Agrawal, “Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection,” *Constr. Build. Mater.*, vol. 157, pp. 322–330, 2017.

[52] Y. Yu, H. Lin, J. Meng, X. Wei, H. Guo, and Z. Zhao, “Deep transfer learning for modality classification of medical images,” *Information*, vol. 8, no. 3, p. 91, 2017.

[53] A. Kharrat and M. Neji, “A hybrid feature selection for MRI brain tumor classification,” in *International Conference on Innovations in Bio-Inspired Computing and Applications*, 2017, pp. 329–338.

[54] T. Bai, J. Tan, M. Hu, and Y. Wang, “A novel algorithm for removal of salt and pepper noise using continued fractions interpolation,” *Signal Processing*, vol. 102, pp. 247–255, 2014.

[55] S. Poudel, Y. J. Kim, D. M. Vo, and S.-W. Lee, “Colorectal disease classification using efficiently scaled dilation in convolutional neural network,” *IEEE Access*, vol. 8, pp. 99227–99238, 2020.

[56] D. Jha *et al.*, “Kvasir-seg: A segmented polyp dataset,” in *International Conference on Multimedia Modeling*, 2020, pp. 451–462.

[57] N. F. A. Al Shalchi and J. Rahebi, “Human retinal optic disc detection with grasshopper optimization algorithm,” *Multimed. Tools Appl.*, pp. 1–19, 2022.

[58] H. Al-Safi, J. Munilla, and J. Rahebi, “Patient privacy in smart cities by blockchain technology and feature selection with Harris Hawks Optimization (HHO) algorithm and machine learning,” *Multimed. Tools Appl.*, pp. 1–25, 2022.

[59] A. F. A. Iswisi, O. Karan, and J. Rahebi, “Diagnosis of Multiple Sclerosis Disease in Brain Magnetic Resonance Imaging Based on the Harris Hawks Optimization Algorithm,” *Biomed Res. Int.*, vol. 2021, 2021.

[60] H. Al-Safi, J. Munilla, and J. Rahebi, “Harris Hawks Optimization (HHO) Algorithm based on Artificial Neural Network for Heart Disease Diagnosis,” in *2021 IEEE International Conference on Mobile Networks and Wireless Communications (ICMNWC)*, 2021, pp. 1–5.

[61] A. T. H. Al-Rahlawee and J. Rahebi, “Multilevel thresholding of images with improved Otsu thresholding by black widow optimization algorithm,” *Multimed. Tools Appl.*, 2021, doi: 10.1007/s11042-021-10860-w.

[62] R. Zhang *et al.*, “Automatic detection and classification of colorectal polyps by transferring low-level CNN features from nonmedical domain,” *IEEE J. Biomed. Heal. informatics*, vol. 21, no. 1, pp. 41–47, 2016.

[63] Y. Shin and I. Balasingham, “Comparison of hand-craft feature based SVM and CNN based deep learning framework for automatic polyp classification,” in *2017 39th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, 2017, pp. 3277–3280.

[64] R. W. Stidham *et al.*, “Performance of a deep learning model vs human reviewers in grading endoscopic disease severity of patients with ulcerative colitis,” *JAMA Netw. open*, vol. 2, no. 5, pp. e193963–e193963, 2019.