Design And Analysis Of A Soft Computing Algorithm For Blood Cells Analysis.

A

SYNOPSIS

SUBMITTED TO THE

**BABA MASTNATH UNIVERSITY**

FOR THE PARTIAL FULFILLMENT FOR THE REQUIREMENT

OF

**DOCTOR OF PHILOSOPHY**

IN

**COMPUTER SCIENCE & ENGINEERING**

****

BY

**VANDITA SHARMA**

UNDER THE GUIDANCE OF

**DR. TILAK RAJ ROHILLA**

**ASSISTANT PROFESSOR**

DEPARTMENT OF COMPUTER SCIENE & ENGINEERING

**FACULTY OF ENGINEERING**

BABA MASTNATH UNIVERSITY

ASTHAL BOHAR, ROHTAK-124021

**2023**

** BABA MASTNATH UNIVERSITY, ASTHAL BOHAR, ROHTAK**

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

**FACULTY OF ENGINEERING**

**CERTIFICATE FROM SUPERVISOR**

1. Title of Research work: Design And Analysis Of A Soft Computing Algorithm For Blood Cells Analysis.

2. Name of Scholar: Vandita Sharma

3. Subject: Computer Science and Engineering

4. Name of Supervisor: Dr. Tilak Raj Rohilla

5. Designation: Assistant Professor

**Signature of the Supervisor Signature of Scholar**

**INDEX**

|  |  |  |
| --- | --- | --- |
| **S.NO.** | **TOPIC** | **Page No.** |
| **1.** | **Introduction** |  |
| **2.** | **Literature Review** |  |
| **3.** | **Research Gap** |  |
| **4.** | **Objective** |  |
| **5.** | **Methodology** |  |
| **6.** |  |  |
| **7.** | **References** |  |

## INTRODUCTION

Soft Computing is a branch of artificial intelligence that focuses on the development of computational algorithms and models that emulate human-like reasoning, decision-making, and perception abilities. Unlike traditional computing techniques, which are based on hard, precise rules and logical algorithms, Soft Computing involves the use of heuristic algorithms and fuzzy logic to solve complex problems that are not easily solved by conventional methods.

The main feature of Soft Computing includes *Fuzzy logic*: Soft Computing uses fuzzy logic to represent and manipulate uncertain or imprecise data. Fuzzy logic enables the system to make decisions based on the degree of certainty or uncertainty, rather than binary decisions based on yes or no. *Neural Networks*: Soft Computing also utilizes neural networks to create algorithms that can learn and adapt to new situations, as well as to recognize patterns in data.

***1)Genetic Algorithms*:** Soft Computing uses genetic algorithms to optimize solutions to complex problems by using principles of natural selection and evolution.

***2)Probabilistic Reasoning*:** Soft Computing also employs probabilistic reasoning to deal with uncertainty and randomness in data.

Applications of Soft Computing include Pattern recognition, Image processing, Control systems, Data mining, Optimization, Forecasting etc.

Neural Network

A Neural Network is a computational model that is inspired by the structure and function of the human brain. It is a network of artificial neurons that are interconnected and communicate with each other through weighted connections. Neural networks are designed to recognize patterns, make decisions, and perform various other tasks based on the input they receive.

The basic building block of a neural network is the artificial neuron or node, which receives input from other neurons or from the external environment. Each neuron in the network computes a weighted sum of its inputs and applies a nonlinear activation function to the sum to produce its output. The output of one neuron is then passed on as input to other neurons in the network, and this process is repeated until the final output is produced.

The process of training a neural network involves adjusting the weights of the connections between the neurons to improve the accuracy of the network's predictions or decisions. This is typically done using a supervised learning approach, where the network is presented with a set of input-output pairs and is trained to learn the mapping between them. The network adjusts its weights to minimize the difference between its predicted output and the actual output.

There are several different types of neural networks, each with its own architecture and application domain. Some common types of neural networks include:

Feedforward neural networks: These networks have a simple, layered architecture, where the output of one layer serves as input to the next layer. They are typically used for classification and prediction tasks.

Recurrent neural networks: These networks have loops in their architecture, which allows them to process sequential data such as time-series or speech signals.

Convolutional neural networks: These networks are designed to process data with a grid-like topology, such as images or videos. They use convolutional layers to extract features from the input and pooling layers to reduce the dimensionality of the data.

Autoencoders: These networks are designed to learn a compressed representation of the input data, which can be used for tasks such as data compression or denoising.

Neural networks have a wide range of applications in fields such as computer vision, natural language processing, speech recognition, and robotics. They are also used in many industrial and commercial applications, such as fraud detection, predictive maintenance, and recommendation systems.

Machine Learning is a subfield of artificial intelligence that involves the use of statistical algorithms and models to enable computer systems to learn from data, without being explicitly programmed. Machine learning algorithms are designed to automatically identify patterns and relationships in data and use these patterns to make predictions or decisions. There are several different types of machine learning algorithms, each with its own strengths and weaknesses. Some common types of machine learning algorithms include:

Supervised learning: In this type of learning, the algorithm is trained on a labeled dataset, where the correct output for each input is known. The algorithm learns to make predictions or decisions based on this labeled data.

Unsupervised learning: In this type of learning, the algorithm is trained on an unlabeled dataset, where the correct output for each input is unknown. The algorithm learns to identify patterns and relationships in the data without any explicit guidance.

Reinforcement learning: In this type of learning, the algorithm learns through trial and error, by receiving feedback in the form of rewards or penalties based on its actions.

Some common applications of machine learning include:

*Image and speech recognition*: Machine learning algorithms are used in image and speech recognition systems, such as those used in self-driving cars, virtual assistants, and security systems. *Natural language processing*: Machine learning algorithms are used in natural language processing systems, such as chatbots, language translation, and sentiment analysis. *Fraud detection*: Machine learning algorithms are used in fraud detection systems to identify unusual patterns or behavior in financial transactions. *Recommender systems*: Machine learning algorithms are used in recommender systems to make personalized recommendations based on a user's preferences and past behavior. *Medical diagnosis*: Machine learning algorithms are used in medical diagnosis systems to identify patterns in patient data and make predictions about disease risk or treatment outcomes.

Machine learning is a rapidly growing field, with new techniques and applications being developed all the time. As the amount of data being generated continues to increase, machine learning is becoming an increasingly important tool for analyzing and making sense of this data.

**Supervised Learning**

There are several types of machine learning algorithms, and supervised learning is one of them. In supervised learning, the algorithm is trained on a labeled dataset, where the correct output for each input is known. The algorithm learns to make predictions or decisions based on this labeled data.

Supervised learning algorithms can be further categorized into two types:

**Regression algorithms**: These algorithms are used for predicting a continuous value or a numerical output. For example, predicting house prices based on various features like location, number of rooms, and square footage.

**Classification algorithms**: These algorithms are used for predicting a categorical value or a discrete output. For example, predicting whether an email is spam or not based on various features like the sender's address, the subject line, and the contents of the email.

Supervised learning algorithms are widely used in various industries, including finance, healthcare, and marketing. For example, in finance, supervised learning algorithms can be used for fraud detection and credit risk assessment. In healthcare, supervised learning algorithms can be used for disease diagnosis and patient risk stratification.

Some common supervised learning algorithms include:

Linear regression: A regression algorithm that predicts a continuous output.

Logistic regression: A classification algorithm that predicts a binary output (e.g., yes or no).

Decision tree: A classification algorithm that creates a tree-like model of decisions based on various features.

Random forest: A classification algorithm that creates multiple decision trees and combines their results to improve accuracy.

Support vector machine (SVM): A classification algorithm that identifies the optimal boundary between different classes.

**Unsupervised learning**

Unsupervised learning is a type of machine learning where the algorithm is trained on an unlabeled dataset, meaning there are no pre-defined output values. Instead, the algorithm tries to find patterns or structure in the data on its own. The main goal of unsupervised learning is to discover hidden patterns, structures, or relationships in the data that can provide insights or help in making decisions. Unlike supervised learning, unsupervised learning does not require prior knowledge or labeled data.

Some common unsupervised learning algorithms include:

Clustering: A technique used to group similar data points together into clusters. For example, grouping customers into segments based on their purchase behavior.

Principal Component Analysis (PCA): A technique used to reduce the dimensionality of the data by identifying the most important features that explain most of the variance in the data.

Association Rule Learning: A technique used to find patterns or relationships between different variables in the data. For example, identifying the most purchased items together in a supermarket.

Anomaly Detection: A technique used to identify unusual or rare data points that deviate from the normal patterns in the data. For example, detecting fraudulent transactions in financial data.

Unsupervised learning algorithms are used in a wide range of applications, including customer segmentation, fraud detection, anomaly detection, and data compression. These algorithms can help in identifying patterns or relationships that may not be immediately apparent to humans and can provide insights or help in making data-driven decisions.

**Reinforcement Learning**

Reinforcement learning is a type of machine learning that involves an agent learning to make decisions in an environment by trial and error. In reinforcement learning, an agent interacts with an environment and receives feedback in the form of rewards or penalties for each action it takes. The goal of the agent is to learn to take actions that maximize its cumulative reward over time.

Reinforcement learning is commonly used in settings where the optimal decision-making strategy is not well-defined or when the environment is complex and dynamic. The agent must learn to balance the exploration of new actions and the exploitation of actions that have worked well in the past.

The key components of a reinforcement learning system include:

Environment: The environment is the world in which the agent operates. It provides the agent with state information, such as the current position, and feedback in the form of rewards or penalties.

Agent: The agent is the decision-maker that interacts with the environment. It observes the current state, takes an action, and receives feedback from the environment.

Action: The action is the decision made by the agent based on the current state of the environment.

Reward: The reward is the feedback that the agent receives from the environment. It is a scalar value that indicates how well the agent's action achieved its objective.

Policy: The policy is the agent's strategy for selecting actions based on the current state of the environment.

Reinforcement learning has been used in a variety of applications, including robotics, game playing, and autonomous driving. One example of reinforcement learning in action is the AlphaGo program, which used reinforcement learning to beat world champions at the game

**Deep Learning**

Deep learning is a subset of machine learning that involves the use of artificial neural networks with multiple layers. Deep learning models are designed to learn representations of data through a hierarchical structure of layers that progressively extract more abstract and complex features from the input data.

Deep learning has become popular in recent years due to its ability to solve complex problems in a wide range of domains such as image recognition, natural language processing, speech recognition, and autonomous driving.

Some of the key features of deep learning include:

**Multiple layers:** Deep learning models consist of multiple layers of interconnected neurons. Each layer extracts features from the input data and passes them on to the next layer. This allows the model to learn complex patterns and relationships in the data.

**Large datasets:** Deep learning models require large datasets to learn from. The more data the model has access to, the better it can learn the underlying patterns and relationships in the data.

**Backpropagation:** Deep learning models use backpropagation to update the weights of the neurons in the network during training. Backpropagation is a technique for computing the gradient of the loss function with respect to the weights, which is used to update the weights and improve the accuracy of the model.

**Transfer learning:** Deep learning models can leverage pre-trained models to perform transfer learning. Transfer learning involves using a pre-trained model to extract features from a new dataset and fine-tuning the model on the new dataset to improve its accuracy.

Some of the popular deep learning architectures include Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Deep Belief Networks (DBNs). These architectures have been used in a variety of applications, such as image and speech recognition, natural language processing, and autonomous driving.

One of the key advantages of deep learning is its ability to learn complex patterns and relationships in data, which can lead to highly accurate predictions and better performance on tasks such as image classification and speech recognition. However, deep learning models can also be computationally expensive to train and require large amounts of data.

Clinical Data Analysis

Clinical data analysis involves using various analytical techniques to extract insights from clinical data to improve patient care, optimize clinical operations, and advance medical research. Clinical data includes a wide range of data types, such as patient demographics, electronic health records, medical imaging, clinical trials data, and genomics data.

Machine learning techniques are often used in clinical data analysis to extract meaningful insights from large and complex datasets. These techniques can help identify trends, predict outcomes, and optimize treatment plans. Some examples of machine learning applications in clinical data analysis include:

Predictive modeling: Using historical patient data to develop models that can predict patient outcomes or identify patients at high risk of developing certain conditions.

Natural Language Processing (NLP): Analyzing unstructured clinical data, such as doctor's notes and patient surveys, to extract insights and identify patterns that may not be easily identifiable through traditional methods.

Image analysis: Using machine learning algorithms to analyze medical images, such as X-rays and MRI scans, to detect abnormalities or identify specific features that may be indicative of certain conditions.

Electronic health record (EHR) analysis: Analyzing EHR data to identify trends and patterns in patient care, monitor patient outcomes, and optimize treatment plans.

Clinical trial analysis: Analyzing clinical trial data to identify factors that contribute to treatment efficacy or to improve the design and implementation of clinical trials.

The insights gained from clinical data analysis can help healthcare providers make more informed decisions, improve patient outcomes, and reduce healthcare costs. However, it is important to ensure that the privacy and security of patient data are maintained throughout the data analysis process.

Image Analysis.

Image analysis is the process of extracting meaningful information from images, such as identifying objects or patterns, measuring properties like size or color, and recognizing shapes or textures. There are various algorithms used in image analysis, depending on the specific task and type of data being analyzed. Here are some common algorithms used in image analysis:

Convolutional Neural Networks (CNNs): CNNs are a type of deep learning algorithm commonly used for image classification and object detection tasks. They are designed to automatically learn and extract relevant features from images, which makes them effective for tasks like identifying objects in images or recognizing faces.

K-Means Clustering: K-Means clustering is an unsupervised learning algorithm used to segment images into different regions based on similarities in pixel values. This algorithm is commonly used for tasks like image segmentation or color quantization.

Random Forest: Random Forest is a supervised learning algorithm commonly used for image classification tasks. It works by training an ensemble of decision trees on labeled image data to predict the class label of new images.

Support Vector Machines (SVMs): SVMs are a supervised learning algorithm used for both image classification and object detection tasks. SVMs work by finding the hyperplane that best separates the different classes of data in high-dimensional feature space.

Region-based Convolutional Neural Networks (R-CNNs): R-CNNs are a type of CNN algorithm that focuses on object detection in images. This algorithm works by first generating a set of region proposals that are likely to contain objects, and then using a CNN to classify each proposed region.

Fast Fourier Transform (FFT): FFT is a signal processing technique used to transform images from the spatial domain to the frequency domain. This algorithm is commonly used for tasks like image filtering, compression, and feature extraction.

These algorithms can be used for a wide range of image analysis tasks, including object detection, segmentation, classification, and feature extraction. Choosing the right algorithm depends on the specific task and type of data being analyzed.

Human Blood

Human blood is a specialized bodily fluid that plays a vital role in the transportation of oxygen, nutrients, hormones, and waste products throughout the body. It is composed of several different components, including red blood cells, white blood cells, platelets, and plasma.

Red blood cells, or erythrocytes, are the most abundant component of human blood and are responsible for carrying oxygen from the lungs to the tissues and organs throughout the body. They contain a protein called hemoglobin, which binds to oxygen and allows the cells to transport it.

White blood cells, or leukocytes, are a critical component of the immune system and play a key role in defending the body against infections and foreign substances. There are several different types of white blood cells, each with its unique function and role in the immune system.

Platelets are small, disk-shaped cells that play a critical role in blood clotting. They are responsible for forming clots at the site of an injury to prevent excessive bleeding and to promote the healing process.

Plasma is the liquid component of blood and makes up approximately 55% of the total blood volume. It is composed of water, proteins, and various other substances, including hormones, electrolytes, and waste products.

White Blood Cells

White blood cells, also called leukocytes, are a vital component of the human immune system. They play a critical role in defending the body against infections, diseases, and foreign substances.

There are several different types of white blood cells, each with its unique function and role in the immune system. These include:

Neutrophils: These are the most abundant type of white blood cells and are responsible for attacking and destroying bacteria and other pathogens.

Lymphocytes: These cells are responsible for producing antibodies and fighting viral infections. There are two main types of lymphocytes: B cells, which produce antibodies, and T cells, which directly attack infected cells.

Monocytes: These cells are responsible for engulfing and digesting foreign substances, such as bacteria and viruses.

Eosinophils: These cells are involved in fighting parasitic infections and are also implicated in allergic reactions.

Basophils: These cells play a role in allergic reactions and are involved in the inflammatory response.

**LITERATURE REVIEW**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **S.No.** | **Title/ Author** | **Objective** | **Methodology used** | **Result/ Conclusion** |
| **1** | White blood cell population dynamics for risk stratification of acute coronary syndrome |  |  |  |
| **2** |  |  |  |  |
| **3** |  |  |  |  |
| **4** |  |  |  |  |
| **5** |  |  |  |  |
| **6** |  |  |  |  |
| **7** |  |  |  |  |
| **8** |  |  |  |  |
| **9** |  |  |  |  |
| **10** |  |  |  |  |
| **11** |  |  |  |  |
| **12** |  |  |  |  |
| **13** |  |  |  |  |
| **14** |  |  |  |  |
| **15** |  |  |  |  |
| **16** |  |  |  |  |
| **17** |  |  |  |  |
| **18** |  |  |  |  |
| **19** |  |  |  |  |
| **20** |  |  |  |  |
| **21** |  |  |  |  |

The quantity of red cells, white blood cells, platelets, and other blood cells is crucial in pathological research for diagnosing illnesses such as anaemia, leukaemia, cancer, and other infectious diseases. The amount of white blood cells is one of those blood cell metrics that is crucial to the body's immune system. WBCs are also used to determine the risk of acute coronary syndrome and acute vascular events such as stroke [1].

Conventional WBC analyzers place a strong emphasis on cell size as well as cytoplasmic and nuclear characteristics [1]. The conclusions of systems that classified cells solely based on size, on the other hand, might be misleading. As a result, peripheral smears, which are still inspected manually today, are the most reliable method for identifying WBCs [2]. Yet, this technique is time consuming and leaves room for human error. As a result, it is possible to reduce the risk of human error and the time required for the manual operation by evaluating peripheral smear images using computer-aided automated detection and diagnostic devices.

Deep learning with Convolution Neural Networks (CNN)[4] is the best solution so far for medical imaging applications such as detection and classification[4,5] at the moment. While CNNs work best on large data sets, training them requires a large quantity of data and computing resources. The dataset is frequently too little and inadequate to train a CNN from start. In such a circumstance, harnessing the power of CNNs. Because of their tremendous natural structure and the availability of a large ImageNet dataset, deep convolutional neural networks (CNNs) have lately gained great success in image identification [6]. Transfer learning can be utilized to reduce computational cost [7,8].

Before being applied to a given job, the CNN is first pre-trained on a large and diverse generic image data set. Pre-trained neural networks for image categorization include VGGNet [9], Resnet [10], Nasnet, Mobilenet [12], Inception [13], and Xception [14].

Liang et al. [15] employed a mix of CNN and recurrent neural networks (RNN) models to identify WBC types by utilising the long-term dependent relationship between several fundamental elements of images and image labels, which was not comprehensively investigated in typical deep CNN techniques. In this study, transfer learning methodologies were applied, as well as multiple pre-trained CNN models. When compared to CNN architectures such as ResNet and GoogLeNet, this model achieved 90.70% recognition accuracy, which is still inadequate when compared to state-of-the-art techniques.

The performance of semi-supervised learning (SSL) approaches in blood cell identification was examined by Livieris et al. [16]. Semi-supervised learning methods are the best machine learning methodology for merging the hidden information in unlabeled data with the explicit categorization information in labelled data. The findings of the experiments suggest that adopting community-based tactics in a semi-controlled learning environment can increase performance. Using the SSL method with the KNN classifier, the study obtained 93.29% accuracy in recognising WBCs. Yet, while this technique improves the accuracy of WBC type identity detection, it suffers from interclass similarities.

Meanwhile, Bani et al. [18] described a method for CNN hyper-parameter tuning based on a genetic algorithm (GA). It identifies critical characteristics that contribute in the differentiation of WBC subtypes. This approach obtained 91% testing recognition accuracy [19] and 99.0% training recognition accuracy using the WBCs dataset. The downside of this strategy is that it is difficult to use and delivers low testing accuracy.

For categorizing WBC types, Banik et al. [35] presented a hybrid CNN architecture. To do this, they employed five deep convolutional layers, three max-pooling layers, and FC layers with one hidden layer. For the WBCs dataset, they achieved 90.79% accuracy using max-pooling to fuse two convolutional layer feature maps into an input for the FC layer, using test data [19]. It is faster than the CNN-RNN model, but it is less accurate due to intraclass and interclass similarities.

Numerous research has used microscopic hyperspectral imaging (HSI) technology to determine the kind of WBCs. The HSI combines spectral data. Xueqi Hu et al. [20] provide a particular methodology for healthy and diseased leukocytes using microscopic HSI technology. The nucleus and cytoplasm are obtained initially via the morphological watershed approach. The spectral properties are then extracted and combined with spatial data. The SVM is then utilized to differentiate between normal and diseased leukocytes. To improve leukocyte identification accuracy, spectral properties must be altered. Nevertheless, there are still certain restrictions, such as the method's need for additional tasks like segmentation and spectral feature extraction.

Similarly, D. Yifan et al. [20] introduce efficient leukocyte recognition method based on microscopic HSI technology. In order to obtain better segmentation result and further improve the representativeness of feature set, the ‘maximum angle convex cone algorithm’ and self-organizing iterative data analysis algorithm are combined to segment the leukocyte from microscopic hyperspectral images. Furthermore, the uniform and rotational variant local binary pattern is applied as texture features with shape and spectral features, SVM is used to classify the leukocytes into different classes. It achieves good segmentation and recognition results when spectral features are used. In contrast, method is complex and required pre-processing segmentation steps for extracting the traditional feature set for cell type identification.

Most prior techniques [22] use feature fusion procedures but ignore the feature selection (FS) method, which is one of the top indicators of identification accuracy e.g. Workload, classification error, and system resources are reduced using FS approaches [17,24]. Nonetheless, several of the algorithms used FS methodologies and achieved reasonable performance in WBC-type identification.

Shahin et al. [23] created a novel deep CNN-based approach for detecting WBC-types. In this method, the pre-trained deep properties of various CNN models at the FC layers are extracted and fed into SVM for WBC type categorization. With WBC datasets, it has a recognition accuracy of 96.10%. (2,551 images). The chi-squared FS technique is also used to select the most effective feature from a high-dimensional feature vector obtained from various CNN architectures.

Khagi et al. [24] developed an Alzheimer's disease classification approach employing FS methods based on the AlexNet architecture. The FS was calculated by employing feature-ranking algorithms such as Mutinffs, RefiefF, Laplacian, and UDFS. The benefit of this strategy is that it makes use of FS algorithms for feature selection. Unfortunately, performance was not particularly good since the intermediary layers of CNN were not used. Additionally, F. Ozyurt [15] offers a unique method that uses four pre-trained models for deep feature extraction: AlexNet [25], VGG-16 [26], GoogLeNet [27], and ResNet [10], followed by the minimal redundancy maximum relevance (MRMR) selection methodology to choose the effective features.

Sahlol, A.T. [17] A hybrid classification technique for White Blood Cell Leukaemia image categorization was presented. It is based on utilizing a deep convolutional neural network (VGGNet) to extract features from WBC images, followed by filtering the acquired features with a statistically enhanced Salp Swarm Algorithm (SESSA) to extract the significant characteristics and reject unnecessary features.

[29] Employ an automated novel deep neural network model for WBC recognition, particularly multi-layer convolutional features of the AlexNet architecture followed by FS methods (MLANet-FS). This model takes use of multi-layer convolutional features from different layers of a CNN architecture. After combining these convolutional features into a single vector, FS approaches are utilised to choose the most effective features for WBC-type classification. The ELM (MLANet-FS-ELM) is then utilized for WBC classification.

[30] The YOLOv5s, YOLOv5x, and Detectron2 R50-FPN pretrained object recognition approaches are used to classify and recognize White blood cells with achieved success rates of 83.3%, 94.66%, and 93.66%, respectively.

[31] investigated use of different feature extractors (Alexnet, VGG16, and Darknet19. According to the testing findings, the YOLOv3 using Alexnet as the feature extractor achieved the highest mean average accuracy (98%) along with lowest losses and requires the shortest time to test.

[32] Provided YOLO-based architecture for identifying microscopic WBC images, as well as a white blood cell detection vision system. Detection strategy is implemented using an enhanced residual convolution module spatial pyramid pooling, enhances the attention mechanism, and improves the loss and activation function of YOLOX-nano.

**RESEARCH GAP**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Reference** | **Objective** | **Date Set Used** | **Method** | **Disadvantages** |
| [29],2022 | WBC identification using multi-layer convolutional  features with an ELM. | Kaggle Data set.[21] | MLANet-FS-ELM (ReliefF) with 99.06% Accuracy. | Less effective on atypical cells. |
| [30],2022 | Detection and classification of white blood cells with an improved deep learning-based approach. | Raabin Health Database | Hybrid Model based on YOLOV5s,YOLOv5x.  With 98% Accuracy. | Small Data Set. |
| [28],2022 | Detection of different types of  blood cells. | Blood Cells Dataset of Microsoft | YOLO\_v3, SSD, faster  R-CNN, R-FCN with accuracy 99% | More localization errors. |
| [11],2022 | Detecting white blood cells,  red blood cells, and platelets. | BCCD | Tiny and efficient YOLOF  (TE-YOLOF) with accuracy 91.9% |  |
| [15],2020 | A fused CNN model for WBC detection with MRMR feature selection  and extreme learning machine. | BCCD Dataset, Kaggle Dataset [21] | CNN-MRM-ELM with 96% Accuracy. | Optimization Techniques can be applied. |
| [17],2020 | Efficient classification of white blood cell leukemia with improved swarm optimization of deep features. | ALL\_DB2,  (C-NMC) | FE: VGGNET,  FS: SESSA,  Classification: Evaluate the selected features.  High FS accuracy.96.11 and 87.9 respectively for data set. | Not Found. |
| [20],2020 | Spatial-spectral identification of abnormal leukocytes based on microscopic hyperspectral imaging technology. | Molecular Hyperspectral Data Acquisition (MHSI) | Data Acquisition, Preprocessing, Segmentation, Feature Extraction, Classification and Identification using SVM Classifier. With Sensitivity and Specificity, OA is above 99%. | Low accuracy of Lymphoblast detection.  Scarcity of abnormal leukocyte data |
| [31],2020 | White Blood Cells Detection using YOLOv3 with CNN Feature Extraction Models. | LISC | FE: Using AlexNet,VGG16,Darknet 19,53. Darknet 19 has highest accuracy. | Small Data Set. |
| [24],2019 | Comparative analysis of Alzheimer’s disease classification by CDR level using CNN, feature selection, and machine-learning techniques. | MRI Image. | Fine Tuned CNN from MRI image. | Single MRI Image. |
| [15],2019 | Fused CNN for WBC Image Classification. | Kaggle Dataset [21] | AlexNet and QDA with 90% Accuracy. | Misclassifies other WBC Images as Neutrophil. |
| [22], 2018 | Combining CNN With RNN for Blood Cell Image Classification. | BCCD dataset, Kaggle Dataset [21] | Xception and LSTM with 90.79% Accuracy. | Experiment based on Most single cell images.  Accuracy 90% |
| [23],2017 | White Blood Cells Identification System Based on Convolutional  Deep Neural Learning Networks. | ALL\_DB1 & ALL\_DB2 | Pre-Processing, CNN FS: Chi-Squared or PCA, Classification e.g. SVM. | Performance Variation is high with different FS methods.  Small Data Set. |

Table 1: Research Gaps

**METHODOLOGY**

The field of object detection and classification using Convolutional Neural Networks (CNNs) is an active area of research in computer vision. Here are some common research methodologies that are used in this field.

**Data Collection and Preparation**

**Network Architecture**

**Data Augmentation**

**Hyperparameter Tuning**

**Evaluation Metrics**

**Transfer Learning**

**Ensemble Methods**

**Data Collection and Preparation**: Collecting and preparing a large dataset is the first step in building a robust object detection and classification model. The dataset should be representative of the problem being solved and should have a diverse range of images with a variety of objects and backgrounds. There are various data set available which can be utilized e.g. PASCAL VOC 2007+2012,MS COCO,Kaggle [19].

**Network Architecture**: Selecting an appropriate network architecture is crucial in developing an accurate object detection and classification model. There are several pre-trained networks such as : AlexNet [25], VGG-16 [26], GoogLeNet [27], ResNet [10],YOLO[32] and Inception available that can be used as a starting point. Researchers can also develop their own custom architectures.

**Data Augmentation**: Data augmentation techniques such as flipping, rotation, cropping, and zooming can be used to increase the size of the training dataset and improve the robustness of the model [37,20].

**Hyperparameter Tuning**: Hyperparameters such as learning rate, batch size, and number of epochs can significantly affect the performance of the model. Researchers can experiment with different values of these hyperparameters to achieve the best results.

**Evaluation Metrics**: Evaluation metrics such as mean average precision (mAP), accuracy [29,30,36,37], and F1 score are commonly used to evaluate the performance of object detection and classification models.

**Transfer Learning**: Transfer learning is a popular technique in object detection and classification that involves using a pre-trained network as a starting point and fine-tuning it for a specific task. This technique can significantly reduce the training time and improve the performance of the model [15].

**Ensemble Methods**: Ensemble methods involve combining multiple models to improve the overall performance. Researchers can experiment with different ensemble techniques such as bagging, boosting, and stacking to achieve better results [17].

Overall, the research methodology for object detection and classification using CNNs involves a combination of data collection, network architecture selection, hyperparameter tuning, and evaluation metrics. Researchers can experiment with different techniques to achieve the best results for a given problem.

## Objectives of the Study

This study will be directed to investigate for blood cell analysis in Human body with the use of lab explores different avenues regarding following concentration:

1. To study the existing algorithms used for clinical data analysis.
2. Performance analysis of existing soft computing algorithms for clinical data analysis.
3. Design and comparative analysis of the proposed algorithm for clinical data analysis.

**References**

[1] Chaudhury, A., Noiret, L., & Higgins, J. M “White blood cell population dynamics for risk stratification of acute coronary syndrome” Proceedings of the National Academy of Sciences of the United States of America.(2017), 114(46), 12344–12349.

[2] Bain B. J. “Diagnosis from the blood smear” The New England journal of medicine. (2005), 353(5), 498–507.

[3] J. Deng, W. Dong, R. Socher, L. -J. Li, Kai Li and Li Fei-Fei "ImageNet: A large-scale hierarchical image database" IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA. (2009), pp. 248-255.

[4] Y. Bengio, A. Courville and P. Vincent, "Representation Learning: A Review and New Perspectives" in IEEE Transactions on Pattern Analysis and Machine Intelligence. (2013), vol. 35, no. 8, pp. 1798-1828.

[5] Malon, C. D., & Cosatto, E. “Classification of mitotic figures with convolutional neural networks and seeded blob features” Journal of pathology informatics. (2013), 4, 9.

[6] LeCun, Y., Bengio, Y. & Hinton, G. “Deep learning. Nature” (2015), 521, 436–444

[7] A. S. Razavian, H. Azizpour, J. Sullivan and S. Carlsson, "CNN Features Off-the-Shelf: An Astounding Baseline for Recognition" IEEE Conference on Computer Vision and Pattern Recognition Workshops, Columbus, OH, USA. 2014, pp. 512-519

[8] Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E. &amp; Darrell, TDeCAF: “A Deep Convolutional Activation Feature for Generic Visual Recognition” Proceedings of the 31st International Conference on Machine Learning, in Proceedings of Machine Learning Research. (2014), 32(1):647-655

[9] Simonyan, K., & Zisserman, A. “Very Deep Convolutional Networks for Large-Scale Image ” Recognition CoRR.(2014), abs/1409.1556.

[10] K. He, X. Zhang, S. Ren and J. Sun "Deep Residual Learning for Image Recognition," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA. (2016), pp. 770-778.

[11] F. Xu, X. Li, H. Yang, Y. Wang, and W. Xiang “TE-YOLOF: Tiny and efficient YOLOF for blood cell detection” Biomedical Signal Processing and Control. vol. (2022), 73, p. 103416.

[12] Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications”. (2017), ArXiv, abs/1704.04861.

[13] C. Szegedy et al. "Going deeper with convolutions" IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA. (2015), pp. 1-9.

[14] Chollet, F. “Xception: Deep learning with depthwise separable convolutions”. Proceedings of the IEEE conference on computer vision and pattern recognition. (2017), p./pp. 1251--1258.

[15] F. Özyurt “A fused CNN model for WBC detection with MRMR feature selection and extreme learning machine” Soft Computing. (2019), vol. 24, no. 11, pp. 8163–8172.

[16] I. E Livieris “Identification of Blood Cell Subtypes from Images Using an Improved SSL Algorithm” Biomedical Journal of Scientific & Technical Research. (2018), vol. 9, no. 1.

[17] A.T. Sahlol, P. Kollmannsberger A.A. Ewees,” Efficient classification of white blood cell leukemia with improved swarm optimization of deep features” (2020), Sci. Rep. 10 2536.

[18] Bani-Hani D Khan N, Alsultan F, Karanjkar S, Nagarur N “Classification of leucocytes using convolutional neural network optimized through genetic algorithm”(2018)

[19] P. Mooney. “Blood Cell Images” (2020).

[20] X. Hu et al. “Spatial-spectral identification of abnormal leukocytes based on microscopic hyperspectral imaging technology” Journal of Innovative Optical Health Sciences.(2020), vol. 13, no. 02.

[21] P.P. Banik, R. Saha, K. Kim “Fused Convolutional Neural Network for White Blood Cell Image Classification” in International Conference on Artificial Intelligence in Information and Communication (ICAIIC). (2019), pp. 238-240.

[22] Y. Duan et al. “Leukocyte classification based on spatial and spectral features of microscopic hyperspectral images” Optics & Laser Technology. (2019) vol. 112, pp. 530–538,

[23] G. Liang, H. Hong, W. Xie, and L. Zheng “Combining Convolutional Neural Network With Recursive Neural Network for Blood Cell Image Classification” IEEE Access.(2018),vol. 6, pp. 36188–36197.

[24] A. I. Shahin, Y. Guo, K. M. Amin, and A. A. Sharawi “White blood cells identification system based on convolutional deep neural learning networks,” Computer Methods and Programs in Biomedicine. (2019), vol. 168, pp. 69–80.

[25] B. Khagi, G. Kwon, and R. Lama “Comparative analysis of Alzheimer’s disease classification by CDR level using CNN, feature selection, and machine‐learning techniques” International Journal of Imaging Systems and Technology.(2019), vol. 29, no. 3, pp. 297–310.

[26] A. Krizhevsky, I. Sutskever, and G. E. Hinton “ImageNet classification with deep convolutional neural networks” Communications of the ACM.(2017) vol. 60, no. 6, pp. 84–90.

[26] K. Simonyan, A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition”. (2015) arXiv:1409.1556.

[27] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. E. Reed, D. Anguelov, et al., “Going deeper with convolutions” (2015).

[28] Shah R, Shastri J, Bohara MH, Panchal BY, Goel P. “Detection of Different Types of Blood Cells: A Comparative Analysis” IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics(ICDCECE).(2022),1–5.

[29] A. Khan, A. Eker, A. Chefranov, and H. Demirel, “White blood cell type identification using multi-layer convolutional features with an extreme-learning machine” Biomedical Signal Processing and Control. (2021), vol. 69, p. 102932.

[30] F. AKALIN and N. YUMUŞAK “Detection and classification of white blood cells with an improved deep learning-based approach” Turkish Journal of Electrical Engineering and Computer Sciences,.(2022), vol. 30, no. 7, pp. 2725–2739.

[31] N. Rohaziat, M. Razali, W. Nurshazwani, and N. Othman “White Blood Cells Detection using YOLOv3 with CNN Feature Extraction Models” International Journal of Advanced Computer Science and Applications. (2020), vol. 11, no. 10.

[32] Z. Han et al. “One-stage and lightweight CNN detection approach with attention: Application to WBC detection of microscopic images” Computers in Biology and Medicine. (2023), vol. 154, p. 106606.

[33] T. Diwan, G. Anirudh, and J. V. Tembhurne “Object detection using YOLO: challenges, architectural successors, datasets and applications” Multimedia Tools and Applications. (2022), vol. 82, no. 6, pp. 9243–927.

[34] L. Liu et al. “Deep Learning for Generic Object Detection: A Survey” International Journal of Computer Vision.(2019), vol. 128, no. 2, pp. 261–318.

[35] P.P. Banik, R. Saha, K. Kim “Fused Convolutional Neural Network for White Blood Cell Image Classification” in International Conference on Artificial Intelligence in Information and Communication (ICAIIC).(2019), pp. 238-240.