International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Real-Time Neurological Disease Prediction with 3D Single Pose Estimation using MediaPipe

[1]Shanmuga Sundari M., [2]Dr. Vijaya Chandra Jadala

Submitted: 10/09/2023 **Revised**: 20/10/2023 **Accepted**: 08/11/2023

Abstract: For the diagnosis and analysis of chronic diseases including Cerebellar Ataxia (CA), Spinocerebellar Ataxia (SCA), and Parkinson's disease, artificial intelligence (AI) is an emerging field. Many doctors, diagnosis teams, and medical professionals benefit from the identification and analysis of neurological illnesses made possible by AI technologies like machine learning and deep learning. Our previous research was completed with the gait value analysis for foot position. This paper aims to conduct the research in real time prediction of neurological disease. For enhancing disease detection model performance and generalizability, diverse datasets from various populations and circumstances must be gathered and annotated. In this research, we applied MediaPipe for real-time disease recognition in videos. We gathered the gait values from the real time videos. The collected gait values are experimented with the target gait values of normal persons. It shows the difference of the gait analysis which helps to predict the disease. The single pose estimation is applied for the activities of human to analyse the gait values. It estimate the gait values in different ankles and compare with the target values. Our research exhibits the predicted result in the real time video of human actions. The experimental result provides the person is affected with neurological disease or a normal person with single pose estimation. Logistic regression algorithm is used to predict the accuracy of the disease. We got 98.53% accuracy for logistic regression algorithm.

Keywords: *regression, neurological, MediaPipe, estimate, circumstances, generalizability*

1. Introduction

One important area of medical specialisation is neurological specialisation [1]. The brain instructs the body on how to respond to events. With this work, we can pinpoint the activity issue and determine the neurological system's capacity. Neurodegenerative disorders may be brought on by a change in a person's circadian cycle. Neurosurgery focuses on injuries to the spine, nerves, and brain. Our doctors treat neurological conditions with neurosurgery. Machine learning has the potential to aid in the diagnosis and treatment of cerebellar ataxia [2]. Researchers have used machine learning algorithms to analyze brain images and identify patterns that can help distinguish cerebellar ataxia from other neurological conditions.

One study used machine learning algorithms [3] to analyze brain scans of individuals with cerebellar ataxia and healthy controls. The algorithm was able to accurately distinguish between the two groups with a high degree of accuracy [4], suggesting that machine learning could be a useful tool for diagnosing cerebellar ataxia. Another study used machine learning to analyze gait [5] data from individuals with cerebellar ataxia. The algorithm was able to identify specific gait patterns that were characteristic of cerebellar ataxia, which could help clinicians more accurately diagnose the condition and monitor its progression.

[1]*Research Scholar, Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswarm, Andhra Pradesh, India*
mshanmugasundari@gmail.com
https://orcid.org/0000-0001-5755-474X
[2]*Associate Professor, Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswarm, Andhra Pradesh, India.*
vijayachandra.phd@gmail.com

Machine learning can also be used to develop personalized treatment plans for individuals with cerebellar ataxia. By analyzing data on an individual's symptoms, genetics, and medical history, machine learning algorithms can help clinicians identify the most effective treatment strategies for each patient. Overall, machine learning has the potential to enhance our understanding of cerebellar ataxia and improve patient outcomes by aiding in diagnosis, monitoring progression [6], and developing personalized treatment plans.

In recent years, advances in computer vision have been made in a variety of applications such as image classification [7], object detection, and facial recognition. Disease detection [8] in the video is one application that has gained a lot of interest, and it has the potential to be used in a variety of real-world applications such as security, entertainment, and marketing. However, due to the dynamic nature of video data and the difficulty in deriving disease-related information from facial images alone, reliably determining disease in the video remains a difficult subject.

Disease detection can potentially redefine how we associate with technology and each other in disciplines ranging from marketing and entertainment to healthcare and security. Disease detection can assist in improving the quality of our lives and the services we receive by personalizing experiences and strengthening the accuracy of decision-making processes.

We designed a single-shot detector model that is specifically optimized for mobile real-time applications, similar to the face detection model utilized in MediaPipe [9] Face Mesh, for identifying the initial hand positions. Detecting hands is a challenging task, and both our lite model and full model must function across a wide range of hand sizes with a significant scale range (~20x) compared to the image frame, while also being capable of identifying obstructed and self-obstructed hands. Unlike faces, which have distinguishable contrast patterns in regions like the eyes and mouth, hands lack such features, making it comparatively difficult to locate them accurately solely based on their visual characteristics. Therefore, providing supplementary context, such as arm, body, or person features, can assist in precise hand localization.

A new object detection framework called MediaPipe has recently been released, and it has already revealed great potential as a cutting-edge technique for real-time object detection. In this paper, we propose a novel MediaPipe-based method for disease detection in video. Our approach leverages the capability of MediaPipe to detect objects in real-time to detect faces in video frames and then employs a disease classification model to predict the disease of each detected face.

The primary objective of MediaPipe's disease detection evaluation is to determine its accuracy and effectiveness in detecting disease from images and videos. This entails evaluating MediaPipe's performance to that of other well-known object detection algorithms, identifying its advantages and disadvantages, and considering enhancements.

The study intends to contribute to the field of disease detection by inspecting MediaPipe and exploring its potential for practical applications. This research aims to develop a deeper understanding of MediaPipe's capabilities and limitations and identify ways to improve its performance for the aim of disease detection.

MediaPipe builds upon the strengths of its predecessors by incorporating new techniques and improvements in deep learning and computer vision. One of the advancements in MediaPipe is the use of a more sophisticated network architecture, which includes several advanced techniques such as dense connectivity and feature pyramid networks. This enables MediaPipe to detect objects more precisely and robustly, even in challenging situations such as low-resolution images and cluttered backgrounds.

Due to its reputation as a precise and efficient object detection algorithm and its most recent advancements, MediaPipe was selected as the algorithm to be validated for disease detection. Furthermore, MediaPipe has been used successfully for a variety of object detection tasks, which signifies that it might also be a good candidate for the disease detection task.

2. Literature Survey

The authors proposed a portable multi-sensory module for monitoring air quality to prevent neurological diseases. The paper starts by highlighting the negative impact of air pollution on human health, including an increased risk of neurological diseases. The authors then present their proposed solution, a portable multi-sensory module that can monitor air quality in real-time. The module includes sensors for measuring temperature, humidity, carbon monoxide, nitrogen dioxide, and particulate matter.

In this study, the researcher employed eight different machine learning methods to analyze cancer occurrence using the WEKA tool. The methods used were Logistic Regression [10], Naive Bayes, Multilayer Perceptron, Stochastic Gradient Descent, Adaptive Boosting [12], and Bagging, Decision Tree, and Random Forest classifiers. The accuracy of each method was compared, and Random Forest was found to be the most accurate. The researchers evaluated the performance of the algorithms using the confusion matrix and other metrics. The study suggests that machine learning can be a useful tool for identifying patterns related to COVID-19 [13] severity prediction, such as C-reactive protein, platelets, and D-dimers.

The objective of the study was to utilize the Gini algorithm [14] and artificial neural network concepts to predict migraine and its severity by capturing features and using vector data to capture distributed delay values. The study included senior patients and compared the performance of these methods with previous research models. A machine learning approach was employed to predict the prognosis of head discomfort. However, the sensitivity of the model was observed to be lower for illnesses with low prevalence. The researchers suggested that the WEKA tool could be useful for feature selection and sampling in machine learning research. Additionally, the study explored several ensemble methods, including bagging, stacking, and voting concepts, using various machine learning algorithms. The researchers proposed several promising frameworks using ensemble learning methods, which involved prediction parameters [15] such as AUC, accuracy, F-measure, Recall, Precision, and MCC. The prediction parameters were presented in a tabular format for output.

1 These are the initial tests of gesture recognition on these targets using the BlazePose skeleton. In addition, we offer an improved BlazePose-based skeleton topology

[16] that can boost the performance of its forerunner. With the use of the Optimized topology described in this paper, we were able to increase the precision of the ST-GCN model's results on the UCF-101 benchmark by over 13%.

The algorithmic approach for estimating poses is thought to be resistant to jarring changes in the surrounding environment and motion. Pose estimation includes both 2D and 3D models [17]. The computing burden of 3D estimate is greater than that of 2D, although 2D has the drawback of errant movements that might lead to inaccuracies. OpenPose, PoseNet, AlphaPose, and BlazePose make up the 2D pose estimation algorithm in people. For feature extraction and real-time applications, BlazePose is deemed suitable. There are three different forms of BlazePose: Light, Full, and Heavy. Activity detection using BlazePose and LSTM [18] has been implemented, and it performs well. However, preprocessing is added in the BlazePose implementation to boost performance.

A computer vision job called "human pose estimation" forecasts the location of a person's body landmarks within a given image or video. By reviewing videos that were taken while the patient was outside of a clinical environment, this technology may be able to provide virtual motion assessments. This research compared the ability of an established solution (OpenPose) and a more recent pose estimation model (BlazePose) to generate clinically useful body feature points for virtual motion assessment. Keypoint coordinates were created from each model using ten videos of clinically pertinent motions (recorded by doctors). The Pearson correlation and root mean square error between the BlazePose and OpenPose keypoint trajectories were computed using OpenPose as the baseline.

BlazePose is a real-time pose estimation system [19] that can accurately estimate the 2D and 3D poses of the human body from a video stream or a series of images. It is based on a deep neural network architecture that uses convolutional layers to extract features from the input image and regression layers to estimate the pose. BlazePose can estimate the poses of multiple people in real-time with high accuracy, even when the people are occluded or overlapping. It achieves this by using a two-stage approach: first, it detects the keypoints of the body, such as the joints and landmarks, and then it estimates the 2D and 3D poses based on these keypoints.

The Rehabilitation Training Evaluation and Correction System based on BlazePose is a system that utilizes the BlazePose pose estimation model to provide real-time feedback and correction for patients undergoing rehabilitation training. The system works by capturing the patient's movements using a camera and then analyzing the movements using BlazePose to detect any deviations from the correct posture or movement patterns.

The system then provides real-time feedback to the patient through visual and audio cues, helping them to adjust their movements and maintain proper form. The system also tracks the patient's progress over time, allowing therapists to monitor their recovery and adjust their treatment plan as needed.

3D Human Pose Estimation using Blazepose and Direct Linear Transform (DLT) for Joint Angle Measurement is a method for estimating the 3D pose of a human body and measuring joint angles using the Blazepose model and Direct Linear Transform (DLT) algorithm [20]. This method is used in various applications such as biomechanics, sports analysis, and virtual reality.

The Blazepose model is used to estimate the 2D key points of the human body in real-time. These key points are then used as input for the DLT algorithm, which estimates the 3D position of the key points in space. The joint angles are calculated from the 3D positions of the key points using the standard inverse kinematics formulae.

3. Proposed Model:

Table 2 presents an overview of the various movements [21] that patients perform while attempting to complete an activity. The activity pattern determines the order in which activities occur over time. The flow of activity [22], which could indicate the severity of the disease, has been ignored during treatment. The study's dataset utilized the movement values shown in Table 2.

Table I. Activity movement

S.No	Movements	Direction
i	Stance	Capturing the upright standing position
ii	Swing	Rotational movement of the arm or hand
iii	Step length	Distance covered while walking
iv	Speed	Strength and velocity of walking
v	Hip	Determining the rotation of the hip
vi	Knee	Measuring the degree of movement in the knee joint
vii	Ankle	Measuring the degree of movement in the ankle joint
viii	Cadence	Continuous movement of the body while walking

3.1 Topology

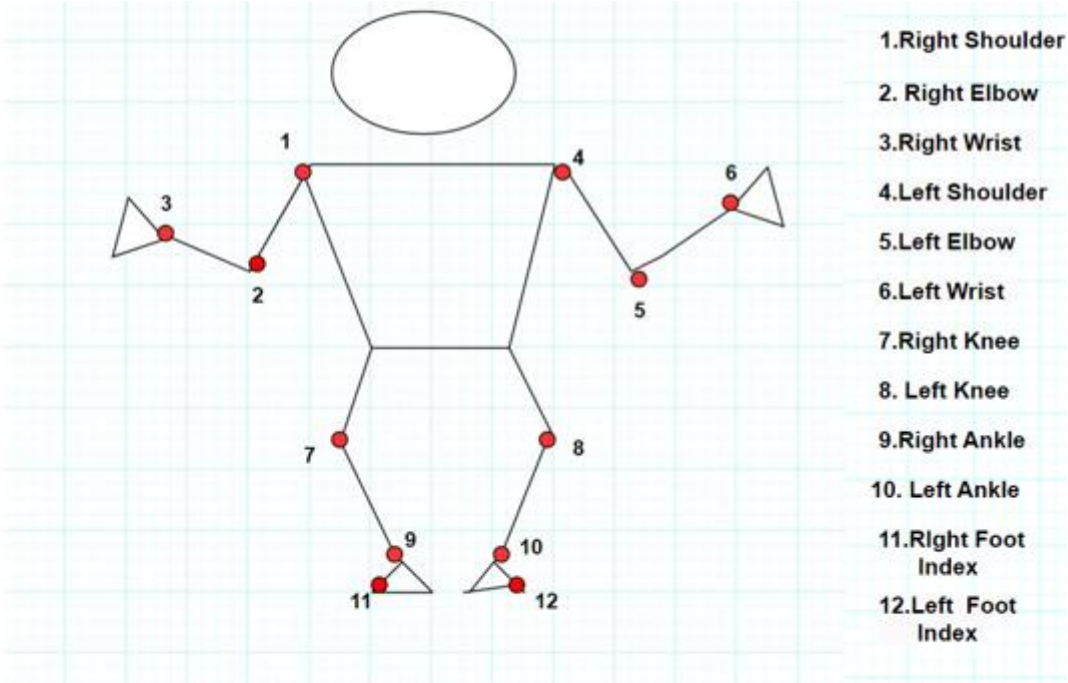


Fig.1 Posture Ankle estimation points

BlazePose is a human pose estimation model developed by Google, which can detect key points on the human body, including the ankles [23] in fig.1. Ankle movement refers to the movement of the ankle joint, which is responsible for dorsiflexion (lifting the foot upward) and plantarflexion (pointing the foot downward), as well as inversion (turning the foot inward) and eversion (turning the foot outward).

BlazePose can detect the position of the ankles in real-time and track their movement, which can be useful for a variety of applications such as fitness tracking, sports analysis, and rehabilitation. By analyzing ankle movement, BlazePose can provide insights into an individual's gait, balance, and overall movement patterns.

3.2 Flow Diagram:

Fig 2 explains the flow diagram of the research. In this the video, we explained the pose estimation of the person to ankle calculation using media pipe. This starts with video frame to calculation of each activity ankle of the person.

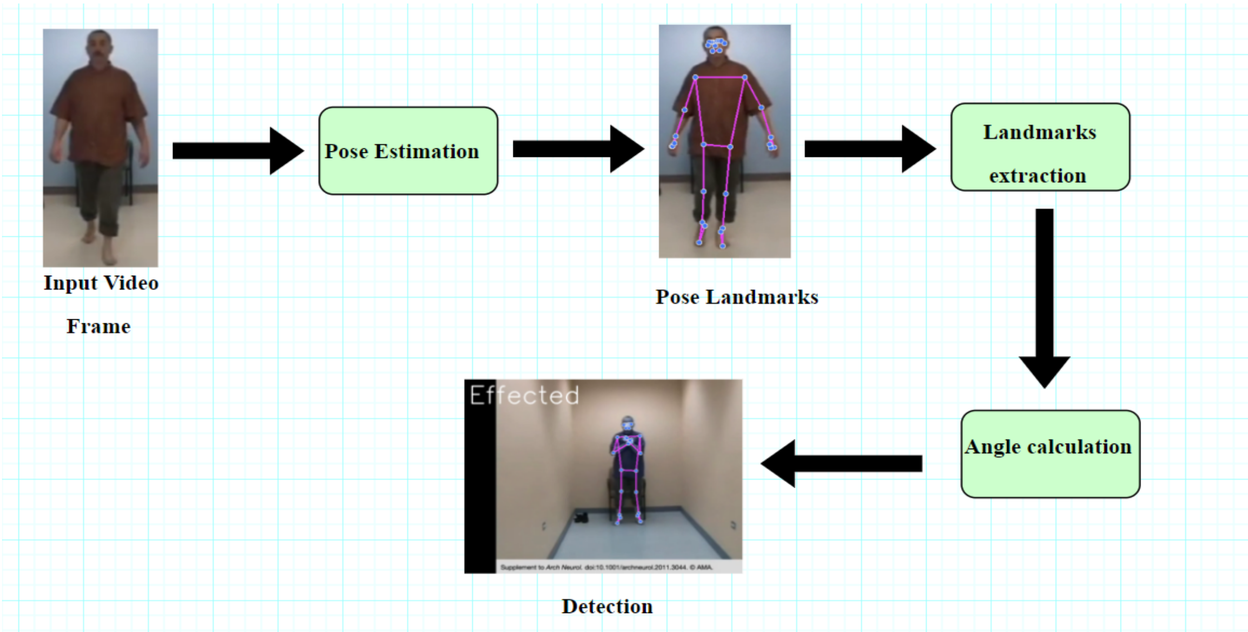


Fig.2 Workflow of Model

- To initiate our workflow, we begin by supplying our model with images and video frames [24] obtained from pre-recorded footage. These frames contain visual data in the form of pixels, which can be in the form of either video or image.
- To train our object detection model, we provide it with training data that includes a collection of labeled images. In these images, the desired objects are marked with manually created bounding boxes, indicating their precise location. Using this data, the model then learns to recognize and identify the targeted objects.
- The subsequent stage entails using our model to examine every frame of the video. We then identify whether any individuals are present and, if so, we ascertain their medical condition.
- At this point, the model will generate results indicating the presence of individuals and their medical conditions. These results will be presented as a list of bounding boxes surrounding the detected objects, along with their corresponding disease labels.
- Upon application, the model produces a confidence score indicating the level of certainty that the identified objects are indeed indicative of the specified disease. This score is utilized to eliminate false positives and modify the behaviour of the model. The identical steps are then repeated for each frame of the video, generating a continuous stream of real-time results.

3.3 Data Set:

We applied the real time video and identify the prediction in the human parts activity [25]. The gait values are collected from the joint of the human part. The values will store and check with the target gait values.

4. Methodology

MediaPipe is a framework useful for connecting many reusable components. The real work AI application are connecting with this vision technology and scoring the prediction analysis with machine learning concepts. Our 3D person pose reconstruction demo app was really simple to build thanks to MediaPipe, which also enabled rapid neural network interpretation on the go and synchronised our result visualisation with the video capture stream.

Multimodal graphs are supported by MediaPipe. Different calculators run in separate threads to speed up processing. Many built-in calculators have options for GPU [26] acceleration for speed optimisation. The system must be correctly synced while working with time series data. The graph makes sure that packet timestamps are taken into account when handling flow. The Framework manages CPU calculator synchronisation, context sharing, and inter-operations.

4.1 Architecture

BlazePose GHUM 3D is a convolutional neural network (CNN) [27] architecture designed for the human body's 3D pose estimation. It comprises three major parts: a body pose estimation network, a depth estimation network, and post-processing. The body pose estimation network receives as input an RGB image of a person and outputs a set of 2D keypoint coordinates corresponding to the individual's body joints. Utilising dilated convolutions, this network, which is based on the SPPE architecture, extracts features from the input image. As a result of the network, a series of 2D heatmaps are produced, each of which shows the probability that each joint will be found in various locations throughout the image.

The depth estimation network takes an RGB image [28] and two-dimensional keypoint coordinates as input and outputs a set of three-dimensional coordinates that represent the position of each joint in three-dimensional space. Another CNN architecture is used in this network to estimate the depth of each joint in relation to the camera. Post-processing entails filtering and smoothing the joint positions over time to reduce noise and jitter, resulting in improved accuracy and stability.

The BlazePose GHUM 3D model is especially useful for augmented reality, fitness tracking, and virtual try-on applications that require real-time 3D pose estimation on mobile devices. For research purposes, the model can be used to estimate the body pose of Parkinson's disease patients and compare it to that of healthy individuals, allowing for early detection and monitoring of the disease's progression.

Finally, BlazePose GHUM 3D is a cutting-edge neural network model which precisely predicts the 3D body pose of individuals in real time. Heatmaps and depth estimation networks [29] are used in the model's architecture, which is based on lightweight convolutional neural networks, to produce 3D joint positions. BlazePose GHUM 3D is especially helpful for mobile devices and has many uses, involving augmented reality, fitness tracking, and virtual try-on. Additionally, the model's ability to identify Parkinson's disease by means of body pose estimation gives medical [30] researchers a strong tool for the early identification and monitoring of the condition. BlazePose GHUM 3D, as a whole, indicates a significant development in the discipline of computer vision and has many uses across multiple sectors.

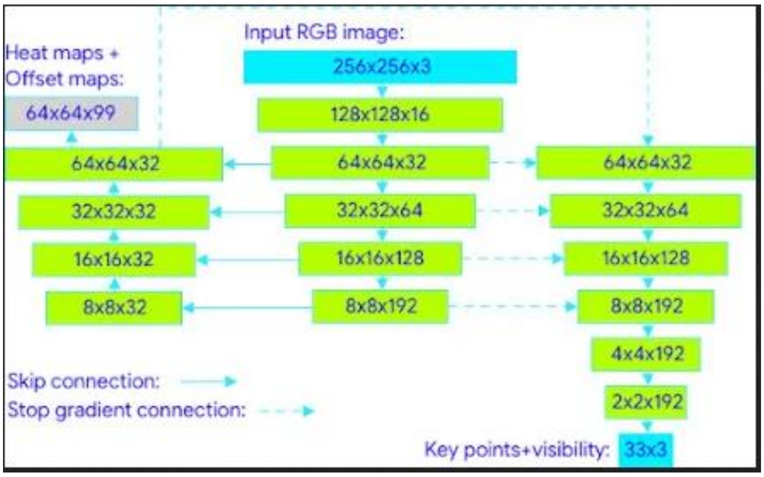


Fig. 3 Blazepose model architecture

4.2 Algorithm

The MediaPipe algorithm is used in object detection because of its speed in detecting the targets in real-time objects, its high accuracy and its learning capabilities. Fig 3 explains the blazepose architecture.

1. First, an input image is taken and resized to a recommended size of 416X416 pixels.
2. The image is then partitioned into a grid of cells, where each cell is employed for detecting an object.
3. Each cell predicts a specific number of bounding boxes with confidence scores and class probabilities.
4. Non-maximum suppression (NMS) is employed to remove intersecting bounding boxes with low confidence scores.
5. The remaining bounding boxes are then filtered based on their confidence scores, resulting in the final set of predictions.
6. Finally, the predictions are presented as bounding boxes, along with their corresponding class labels and confidence scores.

4.2.1 Process flow

The images or videos which are used for training are being annotated, which are converted to dataset format. These are divided into frames, which are passed to our Blazepose training model.

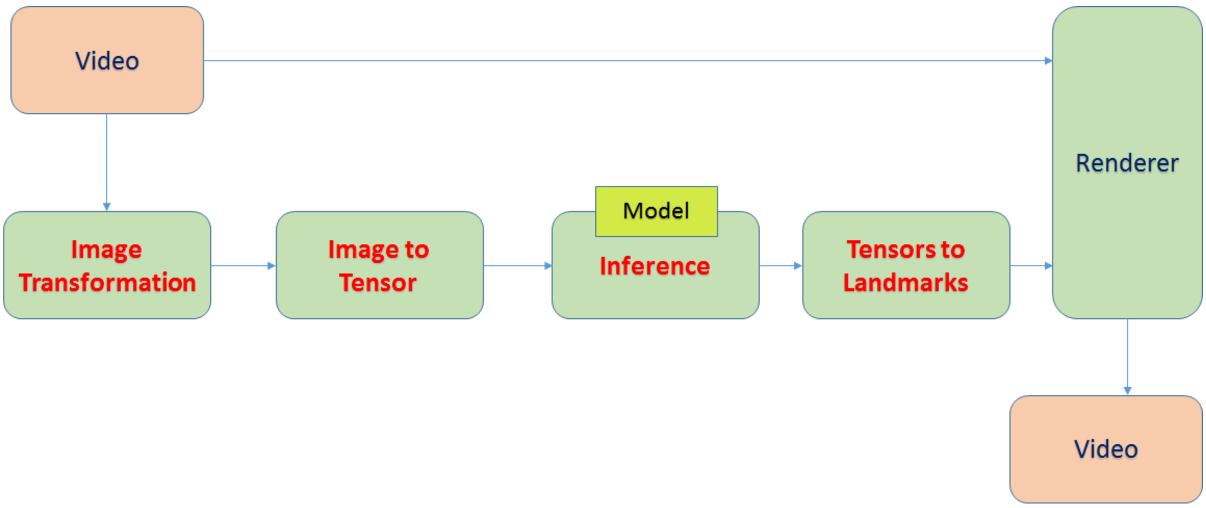


Fig.4 Flow of the model

The following process has categorized into different flow which explains in fig 4. The video has transformed into different levels and finally we predict the disease.

1. The collection of image and media processing calculators includes pre-processing calculators. This category includes the ImageTransform and ImageToTensors in the graph above.
2. For ML inference, inference calculators provide seamless integration with Tensorflow and Tensorflow Lite.
3. Calculators that do post-processing for machine learning (ML) carry out tasks including

classification, segmentation, and detection. A calculator for post-processing is TensorToLandmark.

4. A group of calculators known as utility calculators are used for finishing touches like image annotation.

4.2.2 Formula

The architecture of the Detector is built on the Single-Shot Detector (SSD) model. When an image of size (1,224,224,3) is fed into the model, it generates a bounding box of size (1,2254,12) along with a confidence score of size (1,2254,1). The 12 components of the bounding box consist of (x,y,w,h,kp1x,kp1y,...,kp4x,kp4y), where kp1x to kp4y are additional key points. Each element of the output, which includes 2254 components, has its own anchor, anchor scale, and offset that must be applied.

The Detector can be utilized in two ways. In box mode, the position (x,y) and size (w,h) determine the bounding box. In contrast, in alignment mode, the scale and angle are determined using (kp1x,kp1y) and (kp2x,kp2y), and a bounding box that accounts for rotation can be predicted.

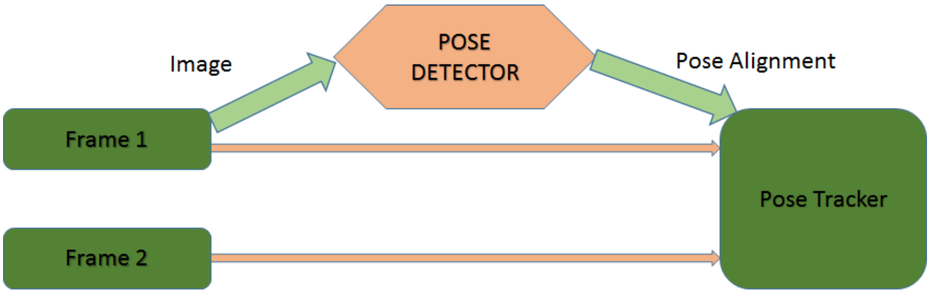


Fig.5 Blazepose estimation

Figure 5 illustrates the process of tracking pose estimation from frames. The BlazePose (Upper Body) model is capable of estimating only the upper body. MediaPipe initially introduced the upper body model and later released the full body model. The full body and upper body models have different specifications, such as the detector resolution of 128x128 for the upper body model.

Intersection Over Union (IOU)

$$IOU = \frac{Inter\ section\ area}{Total\ area} \quad (1)$$

The term "intersection area (1)" describes the portion of the predicted bounding box that intersects with the ground-truth bounding box. On the other hand, the total area refers to the combined area covered by both the predicted and ground-truth bounding boxes. Fig 6 explains the box bounding of the pic position.

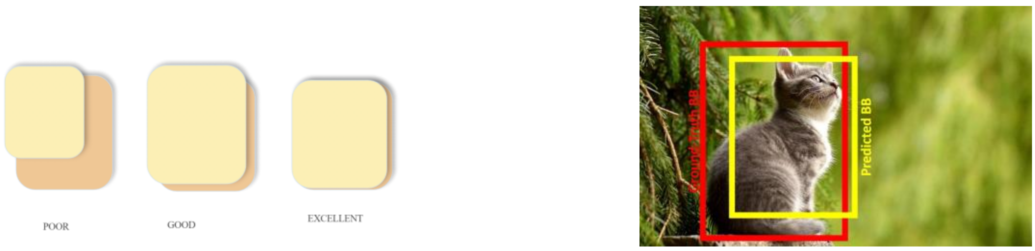


Fig.6 IOU for different bounding boxes

5. Experimental Setup:

In order to train the model, we took videos of normal positions and affected positions. It will sent to trained based on the values what is captured from the algorithms. Then it will be converted into dataset. The dataset will be trained under machine learning algorithm to get the accuracy.

6. Experimental Results:

The functions that measure angles and distances are capable of generating various angle lists, depending on the user's requirements, while the scoring systems can provide predictions based on different thresholds set by the user. If a higher penalty parameter is used, the pose similarity criteria becomes more stringent. In such cases, the suspicious/target behavior detector may fail to detect certain target poses. These functions can extract the essential information required for pose estimation.

The resultant video displayed FPS (Frames per second). It was observed that blazepose processed 12 FPS at a particular instant of the input video. FPS can be calculated using the below formula where. time () is a function in python's built-in library.

currentTime = time. time ()

fps = 1 / (currentTime - startTime)

startTime = currentTime

Fig. 7 FPS and Disease affected detection results

Fig 7 shows the FPS value and result of the person using ankle values. The blazepose is categorize the person is affected with cerebellar ataxia disease. Fig 8 depicts the FPS and normal person result. It is predicted the person is normal with frame per second.

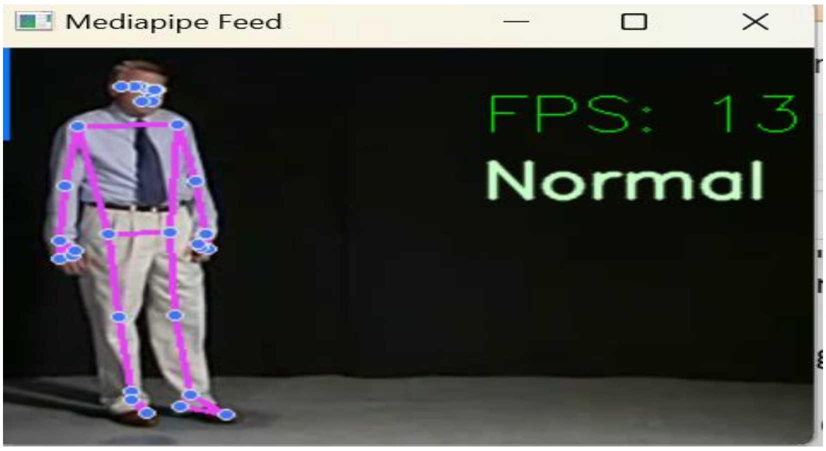


Fig. 8 FPS and Normal person detection results

7. Result and Analysis:

7.1 Logistic Regression Algorithm

We applied logistic regression to predict the neurological disease affected or not. This is the binary classification the logistic regression would be the best algorithm to predict.

Logistic regression is a statistical method used to model the relationship between a binary outcome variable (dependent variable) and one or more predictor variables (independent variables). It is commonly used in various fields, including healthcare, to predict the likelihood of an event occurring.

In logistic regression, the dependent variable is binary, meaning it can take only two possible values, such as "yes" or "no," "positive" or "negative," "1" or "0," etc. The independent variables can be continuous, categorical, or a combination of both. The logistic regression model estimates the probability of the dependent variable taking on the value of "1" (or the positive outcome) based on the values of the independent variables.

The logistic regression model works by applying the logistic function, also known as the sigmoid function, to the linear combination of the independent variables. The sigmoid function "squashes" the linear combination into a range of probabilities between 0 and 1, representing the likelihood of the positive outcome. The logistic function can be expressed as:

P(Y=1|X) = 1 / (1 + e^-(b0 + b1X1 + b2X2 + ... + bpxp))

where P(Y=1|X) is the predicted probability of the positive outcome, b0 is the intercept, b1, b2, ..., bp are the coefficients of the independent variables X1, X2, ..., Xp, and e is the natural logarithm base. The logistic regression model is trained by maximizing the likelihood function, which estimates the probability of observing the actual outcomes given the model parameters. The maximum likelihood estimation (MLE) method is used to find the values of the coefficients that maximize the likelihood function.

After training the logistic regression model, it is possible to utilize it for forecasting the likelihood of the positive outcome for novel observations, taking into account their independent variable values. A threshold probability can be set to classify the observations into the positive or negative outcome. It is important to note that logistic regression assumes that the relationship between the independent variables and the dependent variable is linear on the logit scale. Nonlinear relationships can be modeled using polynomial or interaction terms, or by using more complex models such as generalized additive models (GAMs) or neural networks.

7.2 The CA affected person evaluation

The dataset is prepared using python code. The CSV file is prepared from the ankle movements of the person. The different ankle values are captured and converted into .csv file. There are 4 different ankles values and different frames. So the collected dataset is depending on the single activity and collection of the values will be stored in excel sheet. Fig 9 shows the pixel values of different ankle values.

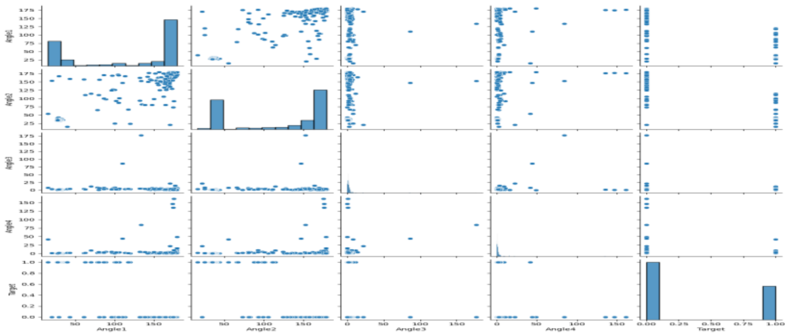


Fig. 9 Pixel values of ankle movements for affected person

Fig 10 explains the different histogram values of the affected person ankle values. The 4 ankle values and target values are represented into histogram model.

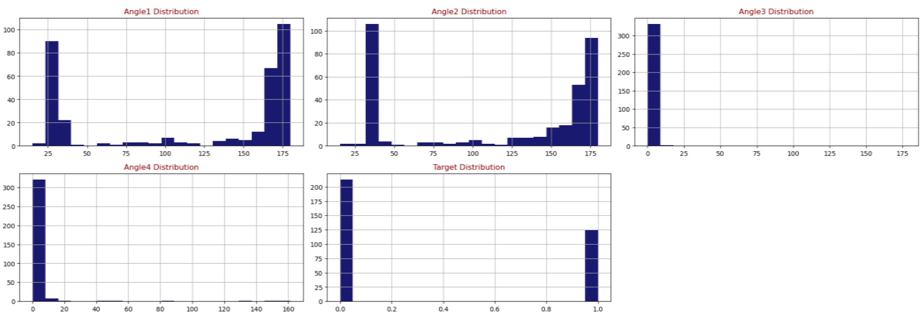
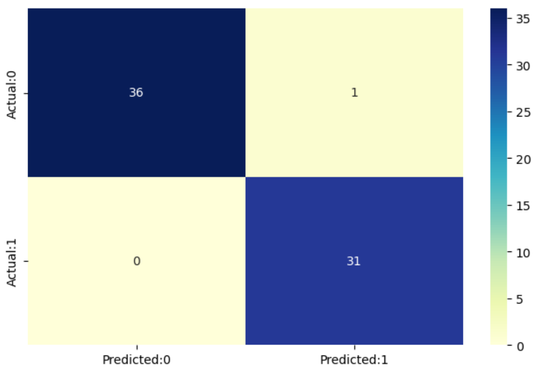


Fig 10. Histogram for Ankle prediction for affected person

Fig 11 depicts the confusion matrix of the affected person. There are 68 values are collected into excel sheet. The 68 values are formed into confusion matrix and finally shows that person is affected by CA disease.



c

Fig 11. Confusion matrix for Ankle prediction for affected person

To create a bar graph of the target class prediction for neurological disease prediction, follow these steps and explains in fig 12.

- Obtain the predicted probabilities of the positive and negative outcomes for each observation in the dataset from the machine learning model. These probabilities should sum to 1 for each observation.
- Determine the threshold probability for classifying an observation into the positive or negative outcome. This threshold can be set based on the desired level of sensitivity or specificity of the model.
- Count the number of observations that have a predicted probability greater than the threshold for the positive outcome. Count the number of observations that have a predicted probability less than or equal to the threshold for the negative outcome.
- Create a bar graph with two bars, one for the positive outcome and one for the negative outcome. The height of each bar represents the count of the observations that belong to that outcome.
- Label the x-axis with the two outcome categories, such as "Positive" and "Negative." Label the y-axis with the count of observations.
- Optionally, add color to the bars to make the graph more visually appealing and to distinguish between the two outcomes.

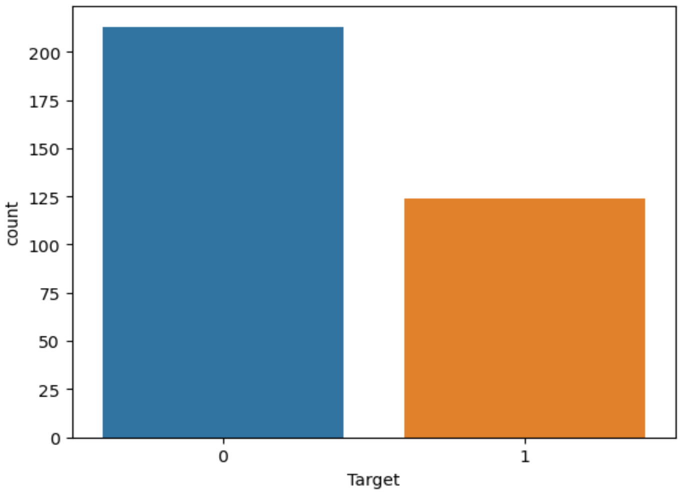


Fig 12. Count of target classes for prediction of affected person

7.3 The normal person evaluation

Fig 13 and 14 explains the normal person's ankle values and histograms representation also. The logistic regression algorithm used to predict that the person is normal.

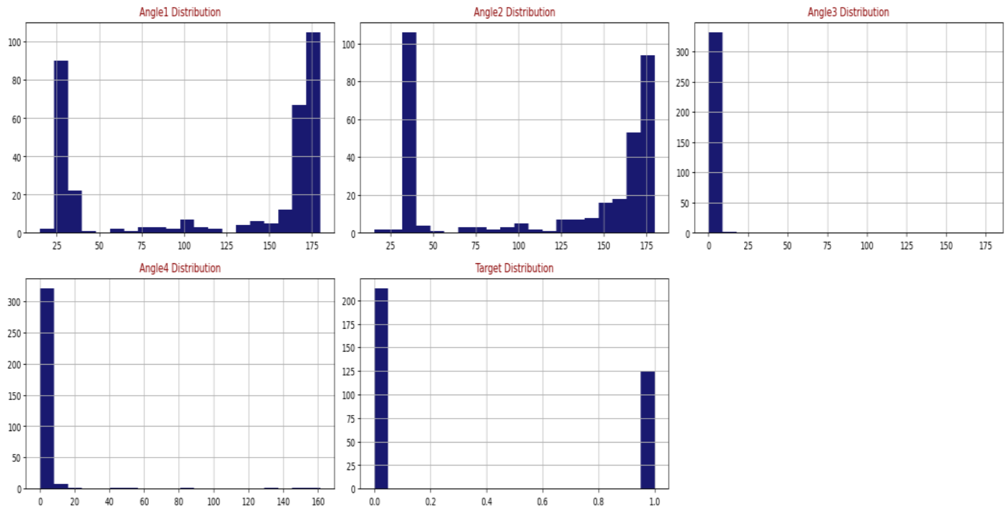


Fig 13. Histogram for Ankle prediction for affected person

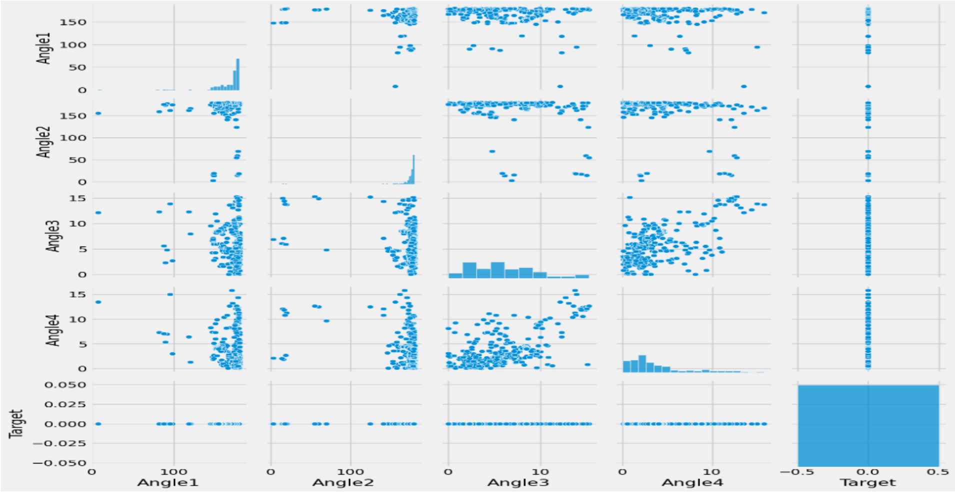


Fig. 14 Pixel values of ankle movements for normal person

Fig 15 is the final target class classification for normal patient and affected person prediction. So the target class prediction is shown as below. In this figure, the affected person accuracy is shown as 98.53% and normal person prediction is 100%.

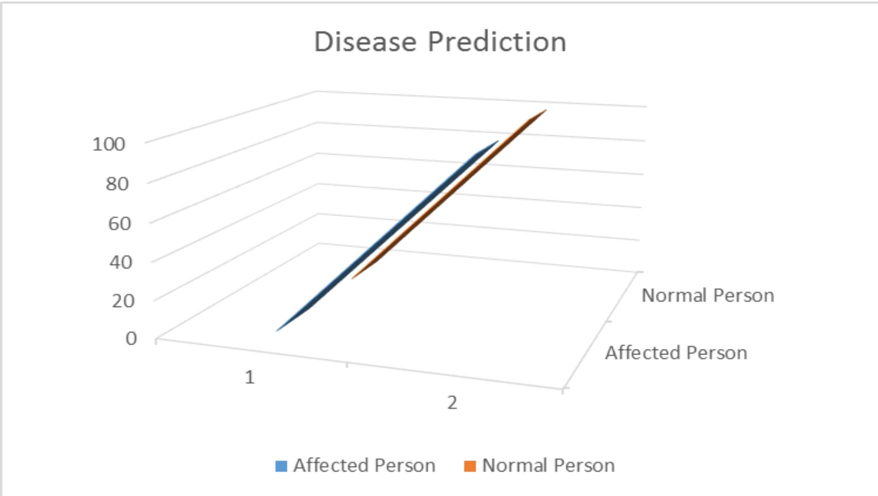


Fig 15. Graph for normal and affected person

8. Conclusion:

Sensory Ataxia and Cerebellar Ataxia are neurological disorders that significantly affect the patients' quality of life. Therefore, timely detection of these conditions is essential. This research proposes the video as a means to differentiate between the symptoms of these diseases. The advantages of this approach include greater comfort, reduced self-consciousness, and other benefits. The patient's activity issue is an indication of the severity of the condition. Deviations from the typical range of motion may suggest that the patient is suffering from CA illness. The data set is gathered using own video capturing method to record the patient's activity disorder data. The fundamental characteristics are identified first, and then the data is pre-processed by removing anomalies. We use Blazepose estimation to find the ankles in real time. Using blazepose, we collected the dataset, then we applied into logistic regression to identify the person is affected with CA disease or not. In this research, we experimented both normal and affected person and found the result for both category. We applied logistic regression algorithm to find the accuracy of the dataset. It is giving 98.53% of the accuracy for affected person and 100% accuracy for normal person in this research. In future, we will use multiple objects and movements for research and improve the accuracy for prediction.

Reference

[1] Shanmuga Sundari, M., & Jadala, V. C. (2023). Neurological disease prediction using impaired gait analysis for foot position in cerebellar ataxia by ensemble approach. Automatika, 64(3), 541-550.

[2] S. S. M, V. C. Jadala, S. K. Pasupuleti and P. Yellamma, "Deep Learning analysis using ResNet for Early Detection of Cerebellar Ataxia Disease," 2022 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC), Bhubaneswar, India, 2022, pp. 1-6, doi: 10.1109/ASSIC55218.2022.10088379.

[3] Sundari, M. S., Jadala, V. C., & Pasupuleti, S. K. (2022, June). Prediction of Activity pattern mining for Neurological disease using Convolution Neural Network. In 2022 7th International Conference on Communication and Electronics Systems (ICCES) (pp. 1319-1324). IEEE.

[4] Chandra, J. V., & Pasupuleti, S. K. (2022, March). Machine Learning Methodologies for predicting Neurological disease using Behavioral Activity Mining in Health Care. In 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 1035-1039). IEEE.

[5] Raju, C. S. K., Pranitha, K., Samyuktha, P., & Madhumathi, J. (2022, March). Prediction of COVID 19-Chest Image Classification and Detection using RELM Classifier in Machine Learning. In 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 1184-1188). IEEE

[6] Alsawadi, M. S., & Rio, M. (2022, August). Human Action Recognition using BlazePose Skeleton on Spatial Temporal Graph Convolutional Neural Networks. In 2022 9th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE) (pp. 206-211). IEEE.

[7] Setiyadi, S., Mukhtar, H., Cahyadi, W. A., Lee, C. C., & Hong, W. T. (2022, December). Human Activity Detection Employing Full-Type 2D Blazepose Estimation with LSTM. In 2022 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob) (pp. 1-7). IEEE.

[8] Mroz, S., Baddour, N., McGuirk, C., Juneau, P., Tu, A., Cheung, K., & Lemaire, E. (2021, December). Comparing the quality of human pose estimation with blazepose or openpose. In 2021 4th International Conference on Bio-Engineering for Smart Technologies (BioSMART) (pp. 1-4). IEEE.

[9] Min, Z. (2022, February). Human Body Pose Intelligent Estimation Based on BlazePose. In 2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA) (pp. 150-153). IEEE.

[10] Yang, H., Wang, Y., & Shi, Y. Rehabilitation Training Evaluation and Correction System Based on BlazePose.

[11] Hakim, I. M., Zakaria, H., Muslim, K., & Ihsani, S. I. (2023, February). 3D Human Pose Estimation Using Blazepose and Direct Linear Transform (DLT) for Joint Angle Measurement. In 2023 International Conference on Artificial Intelligence in Information and Communication (ICAIIC) (pp. 236-241). IEEE.

[12] T. Ngo et al., “Balance deficits due to cerebellar ataxia: A machine learning and cloud-based approach,” IEEE Trans. Biomed. Eng., vol. 68,no. 5, pp. 1507–1517, May 2021.

[13] A. Esteva et al., “A guide to deep learning in healthcare,” Nature Med., vol. 25, no. 1, pp. 24–29, Jan. 2019.

[14] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, “Deep learning for healthcare: Review, opportunities and challenges,” Briefings Bioinf., vol. 19, no. 6, pp. 1236–1246, Nov. 2018.

[15] M. Adnan, S. Kalra, J. C. Cresswell, G. W. Taylor, and H. R. Tizhoosh, “Federated learning and differential privacy for medical image analysis,” Sci. Rep., vol. 12, no. 1, pp. 1–10, Dec. 2022.

[16] M. Hao, H. Li, X. Luo, G. Xu, H. Yang, and S. Liu, “Efficient and privacy-enhanced federated learning for industrial artificial intelligence,” IEEE Trans. Ind. Informat., vol. 16, no. 10, pp. 6532–6542, Oct. 2020.

[17] Y. Liu, J. J. Q. Yu, J. Kang, D. Niyato, and S. Zhang, “Privacy-preserving traffic flow prediction: A federated learning approach,” IEEE Internet Things J., vol. 7, no. 8, pp. 7751–7763, Aug. 2020.

[18] Q. Wu, K. He, and X. Chen, “Personalized federated learning for intelligent IoT applications: A cloud-edge based framework,” IEEE Open J. Comput. Soc., vol. 1, pp. 35–44, 2020.

[19] Y. Zhao, H. Liu, H. Li, P. Barnaghi, and H. Haddadi, “Semi-supervised federated learning for activity recognition,” 2020, arXiv:2011.00851.

[20] Shanmuga Sundari, M., Sudha Rani, M., & Ram, K. B. (2023). Acute Leukemia Classification and Prediction in Blood Cells Using Convolution Neural Network. In International Conference on Innovative Computing and Communications (pp. 129-137). Springer, Singapore.

[21] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, and H. V. Poor, “Federated learning for Internet of Things: A comprehensive survey,” IEEE Commun. Surveys Tuts., vol. 23, no. 3, pp. 1622–1658, 3rd Quart., 2021.

[22] Q. Wu, X. Chen, Z. Zhou, and J. Zhang, “FedHome: Cloud-edge based personalized federated learning for in-home health monitoring,” IEEE Trans. Mobile Comput., early access, Dec. 16, 2020, doi:10.1109/TMC.2020.3045266.

[23] N. Rieke et al., “The future of digital health with

federated learning," npj Digit. Med., vol. 3, no. 1, pp. 1–7, Dec. 2020.

[24] R. Kumar et al., "Blockchain-federated-learning and deep learning models for COVID-19 detection using CT imaging," IEEE Sensors J., vol. 21, no. 14, pp. 16301–16314, Jul. 2021.

[25] S. Warnat-Herresthal et al., "Swarm learning for decentralized and confidential clinical machine learning," Nature, vol. 594, no. 7862, pp. 265–270, 2021.

[26] Sundari, M. S., & Nayak, R. K. (2020). Process mining in healthcare systems: a critical review and its future. International Journal of Emerging Trends in Engineering Research, 8(9).

[27] Z. Yang, S. Zhong, A. Carass, S. H. Ying, and J. L. Prince, "Deep learning for cerebellar ataxia classification and functional score regression," in Machine Learning in Medical Imaging, G. Wu, D. Zhang, and L. Zhou, Eds. Cham, Switzerland: Springer, 2014, pp. 68–76.

[28] Z. Chang et al., "Accurate detection of cerebellar smooth pursuit eye movement abnormalities via mobile phone video and machine learning," Sci. Rep., vol. 10, no. 1, pp. 1–10, Dec. 2020.

[29] B. Kashyap, P. N. Pathirana, M. Horne, L. Power, and D. Szmulewicz, "Quantitative assessment of speech in cerebellar ataxia using magnitude and phase based cepstrum," Ann. Biomed. Eng., vol. 48, no. 4, pp. 1322–1336, Apr. 2020.

[30] H. Tran, P. N. Pathirana, M. Horne, L. Power, and D. J. Szmulewicz, "Quantitative evaluation of cerebellar ataxia through automated assessment of upper limb movements," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 27, no. 5, pp. 1081–1091, May 2019.

[31] Prof. Naveen Jain. (2013). FPGA Implementation of Hardware Architecture for H264/AV Codec Standards. International Journal of New Practices in Management and Engineering, 2(01), 01 - 07. Retrieved from http://ijnpme.org/index.php/IJNPME/article/view/11

[32] Ahammad, D. S. K. H. (2022). Microarray Cancer Classification with Stacked Classifier in Machine Learning Integrated Grid L1-Regulated Feature Selection. Machine Learning Applications in Engineering Education and Management, 2(1), 01–10. Retrieved from http://yashikajournals.com/index.php/mlaeem/article/view/18