Edge Artificial Intelligence: Real Time Non-Invasive Technique for vital signs Myocardial Infraction detection using Jetson NANO

Mohan H M[1], Anitha S[2], Sai Ho Ling[3], Rifai Chai[4]

[1]R/S,Dept. ECE, ACS College of Engineering, Visvesvaraya Technological University, Belagavi, India,

[2]Dept of ECE, ACS College of Engineering, Bangalore, India,

[3]School of Biomedical Engineering. University of Technology Sydney, Australia,

[4] School of Software and Electrical Engineering, Swinburne University of Technology

Correspondence should be addressed to Mohan H M[1]; mohanhm@gmail.com

Abstract

The medical history highlights that Myocardial Infarction is one of the leading factors of death in human beings. The rapid evolution in autonomous technologies, the rise of computer vision and edge computing offers intriguing possibilities in Health care Monitoring Systems. The present work focuses on the chest pain and fall posture based vital signs detection using intelligence surveillance camera to address emergency situation during Myocardial Infarction. A real-time embedded solution persuaded from “Edge AI” is implemented using state of art Convolution Neural Networks: Single Shot Detector, SSD Inception V2, SSD MobileNet V2, on Internet of Things embedded GPU platform NVIDIA’s Jetson Nano. Deep learning algorithmic approach is implemented for 3000 indoor color images datasets: Nanyang Technological University Red Blue Green and Depth dataset and private RMS dataset. The research mainly pivots on two key factors in creating and training a CNN model in detecting the vital signs and evaluating its performance metrics. Here the author proposes a model which is cost effective and low power consumption; onboard detection of vital signs of Myocardial infarction and evaluated the metrics to achieve mean Average Precision of 76.4% and Average Recall of 80%.

Keywords Health Care Monitoring System, Myocardial Infraction, Computer vision, Vital Signs, Intelligence surveillance, Convolution Neural Network.

1. Introduction

One of the prime factor for sudden death worldwide is Ischemic heart disease and Angina pectoris (chest pain) is its most common symptom [1]. Earliest detected information related to Myocardial Infarction (MI) signs and symptoms, and immediately calling emergency services are the main initiative steps need in preventing the life risks. From the onset of symptoms, first one hour is extremely crucial to reach the hospital. In medical history, few critical factors related to heart attack symptoms have been explored such as: People with stroke, heart attack history, diabetes mellitus, and high cholesterol, and high blood pressure etc, [2]. The study highlights common symptoms during Myocardial Infarction wherein chest pain factor is the highest with 84% and shortness of breath, Neck/Back pain, Arm pain, dizziness, sweating account for other factors [3]. Chest pain is the evident clinical marker of myocardial ischemia in the acute phase of a suspected acute myocardial infarction [4]. A person encountering early warnings and experiences more signs and symptoms of MI has to seek immediate investigation and treatment by the doctor to avoid life risks. Considering the relationship of duration of pain and mortality rate, patients with the pain longest and prolonged duration had a highest mortality rate [5]. In recent days, Computer Health Monitoring Systems investigations during cardiac arrest purely depend on investigation reports like: Electrocardiogram (ECG), Echocardiogram, blood tests such as creatine kinase, creatine kinase MB activity and mass concentration, myoglobin, and cardiac troponin T [6].

The patient critical report will be transferred to cardiologist for diagnosing and to provide treatment either with conservative or surgical management.

A person experiencing substernal chest pain (Angina pectoris) responds by holding his clenched fist over the sternum is termed as Levine's sign. The clenched fist sign or palm sign/Cossio-Levin Sign is dominantly found as symptom in patients experiencing myocardial infarction and angina pectoris [7]. A study conducted on body language of chest pain patients at Coronary Care Unit (CCU) suggests that majority of respondents had a clenched fist to the center of the sternum, flat hand to the center of the sternum, both flat hands drawn from the center of the chest outwards and 68% of the participants was considered to be cardiac. The hand movements of chest pain patients has a greater importance from clinical context [8]. In the initial diagnosis of myocardial ischemia patient's Levine sign is of at most important to medical practitioners. [7]. Broader area of chest pain and discomfort corresponds to greater prospect of cardiac ischemia or myocardial infarction [9]. The literature survey reports that the chances of heart attack cases increases with increase in the age of elderly people [10].

In public health monitoring system cardiac fall detection is a major challenge and crucial for addressing the emergency situation for improving the survival rate. A robust, reliable, secure and highly accurate automatic fall detection system can offer medical assistance to the older adults and cardiac patients. From the literature survey, there are variety of risk factors are highlighted in fall detection approaches, and some of risk factors considered in this fall detection systems mainly focused on environmental, physical, and psychological principles as shown in Figure 1. The obstacles disturbances in the paths of environments, stationary and dynamic behavioral body conditions of the patients and finally psychological effects that belongings to people mind behavior [11]. The other way of classification found from the survey based on the polypharmacy and environmental behavioral characteristics like: intrinsic and extrinsic [12]. The psychological factors associated with elderly cardiac patients such as Fear of Fall Syndrome (FoF) and Cardiophobia is often neglected [13] [14]. The prevalence of falls due to cardiovascular disorders remains largely unknown [15].This present work helps in assisting and/or monitoring the patient's physical and psychological health through camera vision based approach.

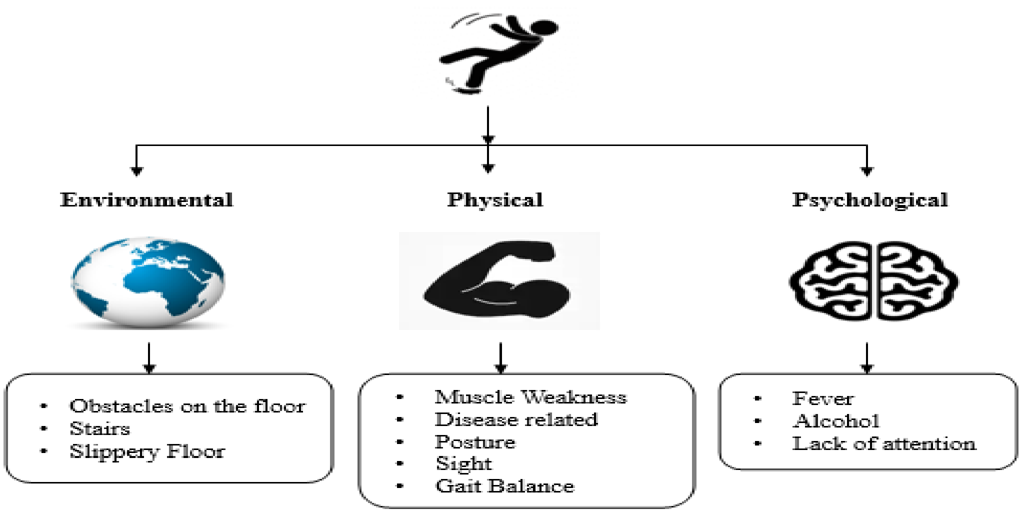


Figure 1: Human Falling Factors

In recent times conflation of Deep learning approaches, effective Internet of Things (IoT) architectures and Edge computing platform are exploited for solving real time remote monitoring applications in health care domain. AI researchers has paid significant attention on decreasing the number of parameters in the Deep Neural Networks thus reducing the computational burden, achieving low latency and memory thereby preserving maximum accuracy for Edge Artificial Intelligence applications[16].

Several noteworthy works have been carried out in medical care domain based on edge platform. The authors adapts a deep learning and edge-cloud computing framework for voice disorder detection and classification [17]. J Pena Queralta et al. implemented advanced architecture of Low Power Wide Area Network (LPWAN) technology along with IoT and deep learning algorithms to enhance the quality of remote health monitoring service. The effectiveness of this architecture is exploited by implementing a fall detection technique using recurrent neural networks (RNN) and the data processing and compression is performed via Edge- Fog computing platform [18]. The authors proposed a combination of Mobile health (mHealth) platform and machine learning approach to develop a detection and classification for skin cancer health issue. The on-device Inference app developed for medical application intends to lower the latency, improvises the privacy issue and saves bandwidth [19].

Recently researchers have developed various kinds of object detection algorithms in computer vision domain with real-time solutions implemented using embedded platforms such as: Raspberry Pi 4, Nvidia Jetson TX1, TX2, Nano and Jetson AGX Xavier. Considering Deep Neural Network based computer vision tasks, there has been specific hardware design concerned about energy conservation [20]. Vittorio Mazzia, et al. proposed a novel approach for apple detection with trained data set apple imaged using modified You Only Look Once (YOLOv3) - tiny algorithm with embedded platforms: NVIDIA's Jetson Nano, Jetson AGX Xavier and Raspberry Pi 3. The performance metrics have been evaluated for apple detection with various background factors and shown that the technique could be employed on unmanned ground vehicles to apple detection with physical aspects [21]. Luis Barba-Guaman et al. explained the reliable and more accurate measurement technique for pedestrians and vehicles detection with three pedestrian metrics like: accuracy, processing time and recall in various environmental conditions with NVIDIA Jetson hardware through convolution neural networks algorithm [22]. Victor Partel et al. the performance of two approaches are designed and simulated to detect the target vegetable object at desired location and simultaneously the spraying on the target object is being evaluated using two embedded GPUs [23]. From this survey the author utilizes a low cost and high computing facility of heterogeneous CPU + GPU SoC system to implement high quality CNNs SSD MobileNet V2, SSD Inception V2 CNNs. In the present work efforts were made to develop a chest pain posture based human fall model on NVIDIA Jetson Nano development board to detect the vital signs during myocardial Infarction.

The CNN based detection technique highlights the simplicity of hardware utilization, optimal power utilization by targeting architecture and algorithmic approaches in low end edge device applications. The researchers from industries mainly targeted to enhance the performance of CNNs in object detection networks by incorporating new architectural design concepts and more accurate algorithmic approaches. CNN structures in object identification methods mainly organized two types: i) One stage learning based classification schemes ii) Two stage learning based classification. [24]. Design parameter exploration method increases the energy effectiveness of CNN based object detection solutions on mobile systems. For example, we can adjust the hyper- parameters of the used CNN model, sacrifice accuracy for speed up by means of dimension reduction techniques and other input level approximation methods. It helps to create a massive space of design parameters for the target CNN based object detection framework. With this sufficiently large design space, we can evaluate the accuracy and power consumption of different implementations, and search for a proper design point that yields the highest score measured in mAP/WH [25]. In recent years the researchers made tremendous efforts to tackle these challenges like improvement in performance metrics of object detection and providing optimal solution in CNN networks through which the performance metrics accuracy speed, mean average precision, average recall, power consumption etc., related to CNN architecture.

Here, the authors propose the vital signs myocardial infraction detection approach based on ConvNets SSD mobilenet V2 and SSD Inception V2 networks that can moderately improve the performance of CNN networks in a real time embedded environment with minimum computer resource utilization is a major contribution of this work. It adapts some of the following factors in this work (i) Create vital signs of myocardial infarction synthetic dataset with expert annotation. (ii) Identify, design and develop light weight

Convolution neural network approach for intelligent video surveillance system. (iii) A computer vision based on Levin's sign and fall detection system implemented using Jetson Nano and evaluating its run time performance.

There is a vast scope for exploring the opportunities in the area of non-invasive approach of detecting and predicting cardio vascular diseases and signs of heart attack. Despite the rapid advancements in low power edge devices, minimal works are carried out in health care segment using artificial intelligence and low-powered GPUs and less importance is provided in the related research works to find out the vital signs of myocardial infarction.

The reminder of the paper is presented as follows. Section 2 discusses about the related work. Section 3 explains the proposed methodology along with dataset being used and hardware description. Section 4 illustrates experimental results along with detailed discussion Section 5 draws conclusions along with future works.

2. Related Work

This section provides the basic information on various human fall detection approaches and edge based AI applications in recent years.

Recent days, many researchers have adapted lot of efforts to develop an accurate and efficient fall detection system with major safety factors in providing the safety measure to the elderly people. Here, fall detection systems classification has been discussed to detect fall incidents among elderly people based on several factors such as: Inertial based, context based and RF based systems as shown in Figure 2 [26]. Table 1 presents the summary of fall detection techniques organized with following criteria: Data sets used, Number of subjects considered as samples, Sensor modalities and algorithms carried out during their work.

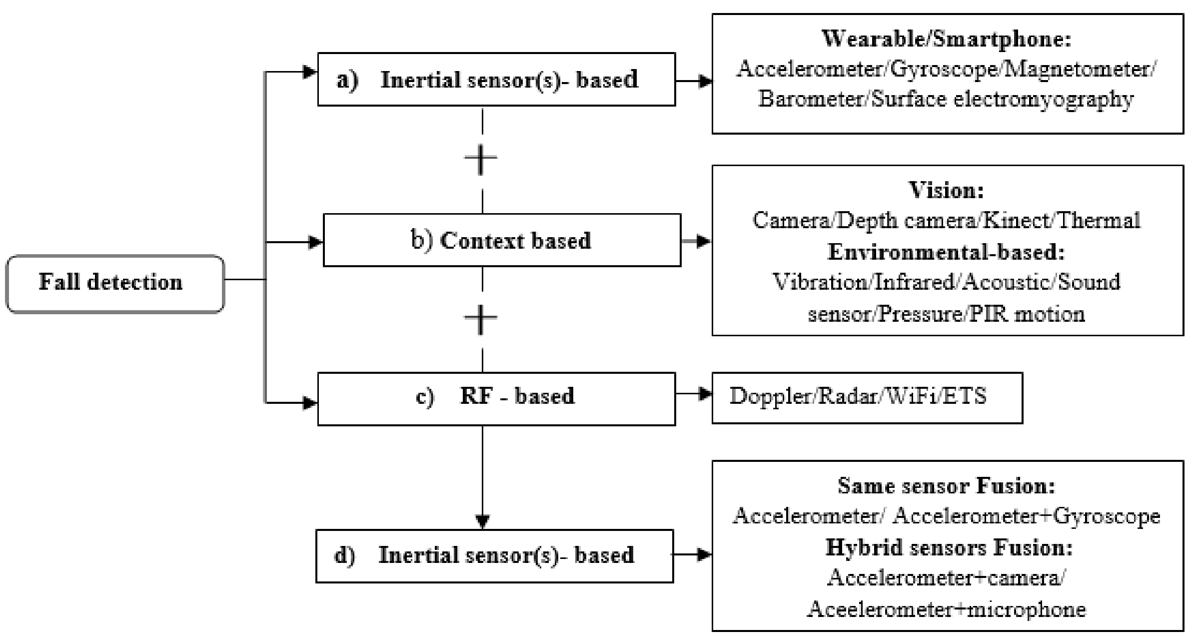


Figure 2: Fall Based methods

Table 1: Fall detection Techniques

Author/year	Data sets	No. subjects (age)	Sensor	Algorithms
Saleh and Jeann`es [27]	Simulated	23 (19-30), 15 (60-75)	Accelerometer (Waist)	SVM
Zitouni et al.[28]	Simulated	6 (N/A)	Accelerometer (Sole)	Threshold
Wu et al. [29]	Public (Simulated)	42 (N/A), 36 (N/A)	Accelerometer (Chest and thigh)	Decision tree
Huang et al.[30]	Simulated	12 (19-29)	Vibration	HMM
Tian et al. [31]	Simulated	140 (N/A)	FMCW radio	CNN
Wang et al.[32]	Simulated	N/A	WiFi	SVM, Random Forests
Kerdjidj et al. [33]	Simulated	17 (N/A)	Accelerometer, Gyroscope	Compressive sensing
Queralta et al.[34]	Public (Simulated)	57 (20-47)	Accelerometer, Gyroscope, Magnetometer	LTSM
Han et al. [35]	Simulated	N/A	Web camera	CNN
Kong et al. [36]	Public	Public	Camera (Surveillance)	CNN
Ko et al. [37]	Simulated	N/A	Camera (Smartphone)	Rao-Blackwellized Particle Filtering
Shojaei-Hashemi et al.[38]	Public (Simulated)	40 (10-15)	Kinect	LSTM
Min et al. [39]	Public (Simulated)	4(N/A), 11 (22-39)	Kinect	SVM
Ozcan et al.[40]	Simulated	10 (24 - 31)	Web camera	Relative-entropy-based

2.1 Wearable device based approaches

A fall detection based devices should adapt robustness and high reliability for real time scenarios. Wearable technology relies on embedded sensors that detect, analyze, and transmit information for monitoring human activities. To address the issue of human based fall approaches several innovative research works have been carried out involving wearable devices with inertial sensors such as accelerometers, Fusion of accelerometer and posture sensors, Tri-axial accelerometer, gyroscope etc.,

One of the challenging problems of wearable based fall approach is to design a low power highly accurate detector for both indoor and outdoor environment. The authors design low computational cost wearable fall detector based on two level Support vector machine online feature extraction method using a 3-axial accelerometer. The machine learning-based system works with multiple sampling frequencies with best accuracy/complexity tradeoff [27]. The authors developed a novel approach of wearable fall technique based on tri-axial accelerometer. A smart sole tracker was designed using the concept of differential acceleration and time threshold based on Low energy Bluetooth communication [28]. A Fall-detection Ensemble Decision Tree (FEDT) algorithm was proposed by Wu T et al. for a reliable fall detection in practical scenario utilizing mobile cloud computing resources [29] Huang, Y et al. implemented a novel idea of training free fall recognition based Hidden Markov Model (HMM) named as GFall, based on geophones. The model developed intended to reduce the false alarm rate using a reconfirmation mechanism called Energy of- Arrival (EoA) positioning for detecting human fall [30]. Tian Y et al. introduced Aryokee, and Frequency-Modulated Continuous-Wave radar (FMCW) based signal to overcome the limitations of other wearable fall based approaches. The work tries to address certain practical challenges such as tackling complex falls and sudden non-fall movements, Detect falls in the existence of other motion and Generalization to environment and people [31]. Wang, Y et al. designed a device for free fall detection through a combination of WiFall, Wireless network and ML approaches like: Support Vector Machine and Random Forest. The system leverages Channel State Information (CSI) as the criterion uses a technique of temporal stability and frequency diversity for human activity and fall detection [32]. The authors presented a novel compression sensing technique and devised a Shimmer device for fall and human activity detection mainly aimed for reducing the energy consumption. The method explores the advantage of combining two sensors: accelerometer and gyroscope and incorporate compression sensing capability and final classification is performed using ML algorithms such as Ensemble Classifier (EC), Support Vector Machine (SVM), Decision Tree (DT) and k-Nearest Neighbor (k-NN) [33]. A fall based remote healthcare monitoring was designed by employing IoT- architecture based systems, LPWAN technology and RNN deep learning algorithm to increase the effectiveness in detection and classification [34]. There are major drawbacks associated with wearable devices such as: Generation of more false alarms, devices getting disconnected easily, sensitive to external factors, person forgetting to wear, and inconvenience of wearing it all day long which makes the system inconsistent to provide highly accurate automatic fall detection.

2.2 Camera (Vision) based approaches

Vision based surveillance systems overcome the drawbacks of wearable fall approaches to impart practical and complex framework. Han Q et al. uniquely advocated a two-stream approach to process a video data for human fall detection and implemented using a lightweight CNN VGG network suitable for deploying on mobile phones [35]. Kong Y et al. put forward a computer vision based fall identification for Single and Multi-camera video surveillance system. An effective three stream CNN approach is presented wherein motion images are fed for silhouette feature extraction in first two stream and dynamic images with temporal information is fed to third stream [36]. The authors concentrated on dynamic and complex outdoor environment for solving universal human detection and fall. A Rao-Blackwellized Particle Filtering is utilized for feature extraction from RGB depth images [37]. A deep-learning technique of long short-term memory (LSTM) feed forward neural network for human fall detection is presented using a transfer learning approach and outperformed existing works based on hand-crafted features [38]. A privacy-preserving fall method was proposed wherein Kinect sensor 3D skeleton image input was utilized to train Support Vector Machine (SVM). The method achieved reduction in number of parameters compared to other deep learning approach with lower time cost [39]. The authors adapts an approach of vision-based activity monitoring with wearable camera worn by the subject compared to static camera installed at fixed locations. An improved varient of the histograms of oriented gradients (HOG) is implemented along with gradient local binary patterns (GLBP) an adequate threshold for fall prediction is estimated by Ali–Silvey distance measure [40]. Through extensive literature survey review it was found that very few works have been carried out based on noninvasive heart attack detection from color images. Gabriel Rojas Albarracín et al. proposed a heart attack detection approach from 1500 RGB color images using a Convolutional Neural Network [41].

As examined from the literature survey deep learning methods with inbuilt convolution networks based on advanced convolution networks are more predominantly applied to predict objects, classification and localization. CNN supports up to maximum extent in object detection techniques in evaluating automatically the objects based on their salient learning features. Deep learning techniques used automatic feature extraction to provide more accurate and efficient solutions to tackle the real time problems in computer vision domain.

Most popularly used human fall based publically available datasets have been summarized in Table 2. Information consists of number of subjects, age range of participants, total number of samples, sensors and its position used in dataset, scenario of the data collection zone has been summarized [42].

Table 2: Human fall based open datasets

Dataset/Year	Sensors	Number of Subjects (Age)	Total Samples	Position of Sensing points	Scenario
UP-Fall[2019]	A, C,E,L,IR,G	17 [18-24]	561	H,F,N,Wa,Wr,An	Lab
SisFall [2017]	A,G	38 [19-75]	4505	Wa	Gym,Hall
UniMiB SHAR[2017]	A	30[18-60]	7013	T	N/A
NTU[2016]	k	40[10-35]	56000	Ce	Lab
UMA Fall [2016]	A,G,M	17[18-35]	531	An,Ch,T,Wa,Wr	Home
MobiAct[2016]	A,G,O	57[22-47]	2526	T	Gym,Hall
MobiFall[2013]	A,G,O	24[22-47]	630	T	Gym,Hall

Note: N/A: Not Appropriately defined, C: RGB camera, A: Accelerometer, G: Gyroscope, O: Orientation Measurements, K: Kinect sensor, M: Magnetometer, IR: InfraRed sensor, L: Luminosity sensor, E: Electroencephalography (EEG) headset, Ce: Ceiling, T: Thigh (pocket), Wa: Waist, Wr: Wrist, An: Ankle, Ch: Chest, H: Head, N: Neck, F: Floor.

The sophisticated deep learning based architectures require heavy computational resources. Deep learning models can be deployed in a centralized cloud computing platforms which offer significant options for high performance computation. However there are certain challenges and constraints that make ML data services between devices and cloud environments impractical such as privacy, financial overheads, latency factor and energy that affects the performance of the system. Some of these problems can be majorly resolved through edge computing commonly known as “edge AI” in any computing field whose performance is evaluated locally obtained data from any sensing devices or database. Vittorio mazzia et al. implemented You Only Look Once, YOLO targets on classification and localization in object identification and proposed regression to recognize the objects in an image without Region Proposal Network (RPN) [21]. Four different models are deployed in an embedded system using deep learning algorithms for object identifications SSD Mobilenet V1, SSD Mobilenet V2, Penet and Multiped [22]. Recent developments in various deep learning neural network models and algorithmic achievements in deep learning techniques for robotic applications will able to achieve better metrics performance like: recognition rate and detection accuracy. The authors Shenshen Gu, et al. highlight two concepts for nonlinear models to improve the recognition and path planning of robots performance. In tennis ball collection robot using deep learning algorithms able to perform i) pointer network model to resolve travelling salesman problem of tennis ball and ii) YOLO model for real time object detection (tennis): and these two concepts are deployed on NVIDIA Jetson TX1 board for performance evaluation in tennis ball path finding. The experimental results highlight the path planning tennis ball recognition and also an optimal path planning solution very quickly [43]. Unmanned aerial vehicles importance and its deployment using NVIDIA Jetson TX board is capable to perform emergency activities where deployment of humans are crucial. The author achieved an optimal solution in object detection configurations with parametric resolved using mathematical equations. The experimental analysis is evaluated for detection accuracy, speed of detectors using CPU multiscale ACF detector as well as YOLO V2. YOLO V2 shows better performance than ACF in evaluating the performance metrics like frame rates and detection accuracy [44]. In real world scenario applications like drones, autonomous driving, robotics etc., there are certain constraints like computational resources to pursue high accuracy from a limited computational costs.

In this present work, we implement deep learning technique for feature extraction and detection for identifying the proposed vital signs of myocardial infarction through chest pain posture and fall posture image based fall detection application. Specifically, we adopt state-of-the-art CNN lightweight architectures SSD Inception V2 and Mobile Net SSD V2, a highly efficient, memory efficient network for low-powered GPU device. The present work exploits transfer learning technique using a state of art light weight pre-trained neural networks such as: ConvNets SSD Mobilenet V2 and SSD Inception V2. These lightweight CNN architectures can predict relatively faster than other algorithms due to its competitive performance by reducing the computational complexity and ease of implementation for enhanced performance on a low power embedded edge device. In the present research work, the main aim is to classify and identify the fall states considering as the vital sign during emergency situation of heart attacks such as: chest pain Levine's sign, partial fall and complete fall posture.

3. Proposed Methodology

3.1 Overview of the Proposed Model

In this proposed work, two state-of-art convolutional blocks are combined as a single lightweight advanced architecture to implement an object detection task into a computationally intensive GPU embedded device.

3.1.1 Single Shot Detector-SSD

Single Shot Multibox detector (SSD) architectural model adapts feed-forward convolutional network to achieve exemplary performance in object detection. The network efficiently performs localization and classification of objects in a single forward step. The entire SSD model consists of mainly two segments: Firstly the base network which performs high-quality image feature extraction and Secondly SSD evaluates the classification result. In case of MobileNetV2 SSD network shown in Figure 3, MobileNetV2 extracts the image features and subsequent convolution layers of SSD performs the classification task. SSD inherits the concept of anchor boxes strategy and feature pyramid structure from Faster RCNN algorithm to generate default boxes of various aspect ratios and scales followed by a non-maximum suppression technique to produce the final detections. The performance of a Single Shot multi-box Detector (SSD) is measured with scaling down the both model size and complexity using multiple feature maps in a network to enhance the metrics speed and removing the proposed regions; to predict large objects through deeper layers and smaller ones with shallow layers in applications such as: mobile and embedded devices. [46].

Every prediction of object in an image originates from a concept of boundary box in SSD. Feature maps of different resolutions are applied to pre-processed image in a convolutional manner to create overlapped bounding boxes. Several multi resolution boxes called as default boxes of different scales and sizes are generated relative to the input image as shown in Figure 4. Score values are evaluated for every box and highest score is selected finally as a class for the bounded box. The scale of every default boxes for every feature can be evaluated as in equation 1.

$$S_k = S_{min} + \frac{S_{max}-S_{min}}{m-1}(k-1), \quad (k \in [1,m]) \tag{1}$$

Where, m is the total number of feature maps S_{min} and S_{max} are the lowest and highest scaling factor to be set respectively.

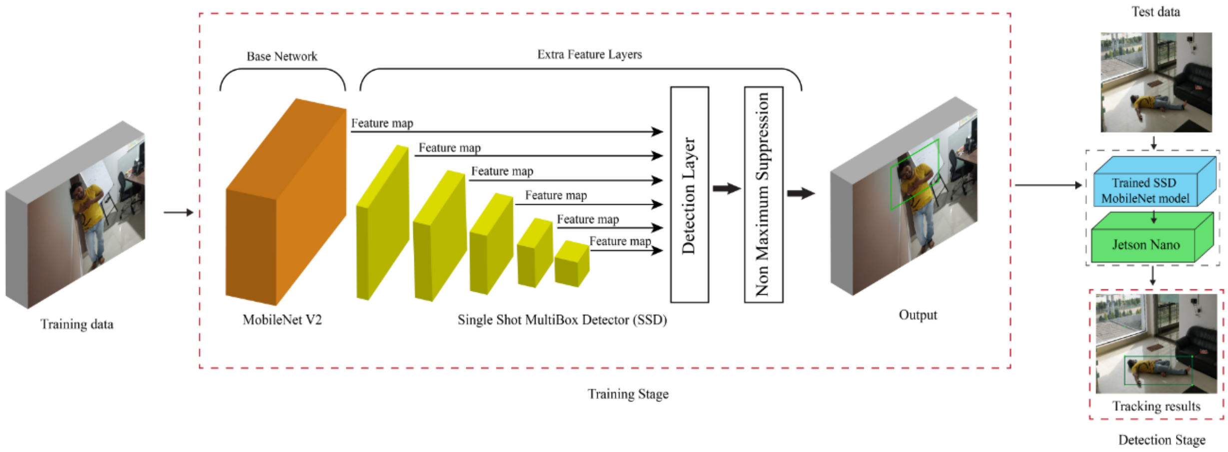


Figure 3: SSD MobileNet V2 architecture

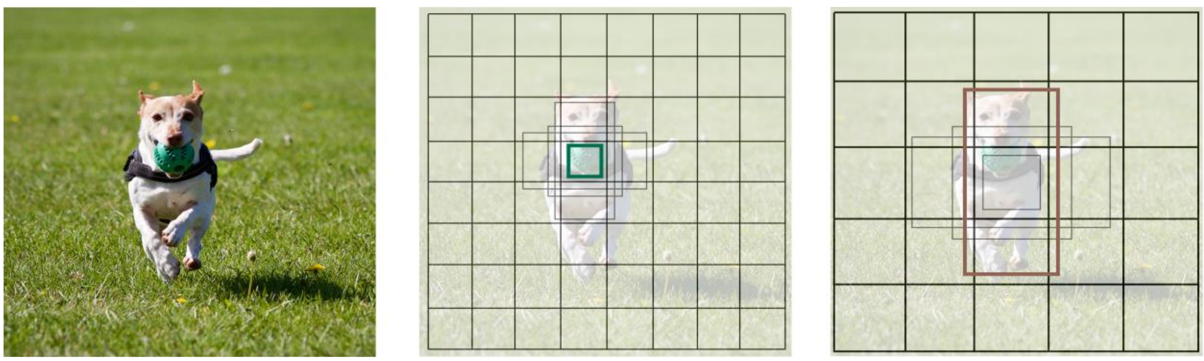


Figure 4: Generated default boxes of SSD Model

3.2 The Proposed Method

Figure 5 shows the methodology of the proposed work. The entire work is divided into two stages as follows; Training stage and Detection stage. The steps for the training and detection stage are structured as follows:

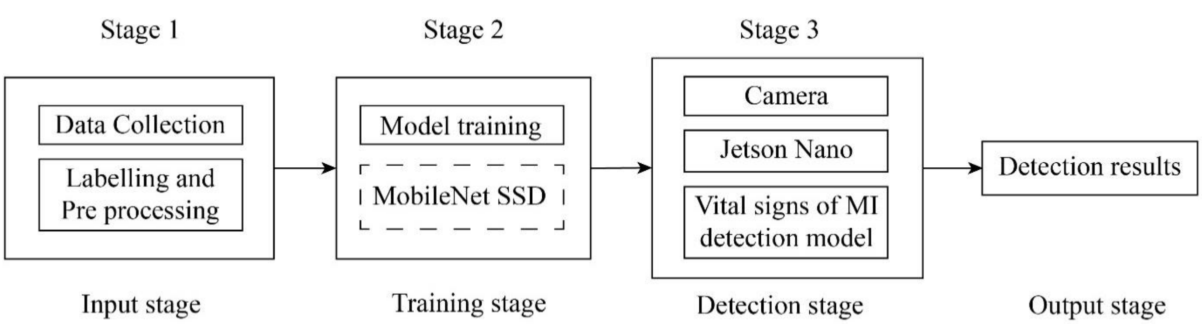


Figure 5: Methodology for training and detecting real time Vital Signs Myocardial Infarction.

The methodology includes three main stages: i) input stage ii) Training stage iii) Detection or Output stage. The input stage incorporates raw input images from two datasets: Custom synthetic dataset (RMS) and public benchmark dataset (NTU RGB+D). Subsequently preprocessing technique is carried out to downsize the images for improving the training speed and to avoid overfitting problem. Region of Interest (Chest pain posture, Partial and Complete fall postures) of every image is marked using the Labellmg software. During the training stage the custom dataset is trained using the COCO pre-trained (TensorFlow Zoo) Deep Neural Network CNN models SSD Inception V2 and SSD mobileNet V2. The training of the CNN model is performed on the workstation and the model has been deployed on Jetson hardware platform. During the final detection stage real-time detection is carried out on Jetson Nano board by connecting a single camera and trained CNN model is used to detect the vital sign postures of Myocardial Infarction as shown in Figure 5.

3.3 Data Set Collection and Preprocessing

During deep learning algorithm implementation, input data set quality and total number of images play significant role on the final performance of the network. A benchmark standard is followed while collecting data samples from two sources classified as follows:

3.3.1 Action Recognition Dataset (NTU RGB+D)

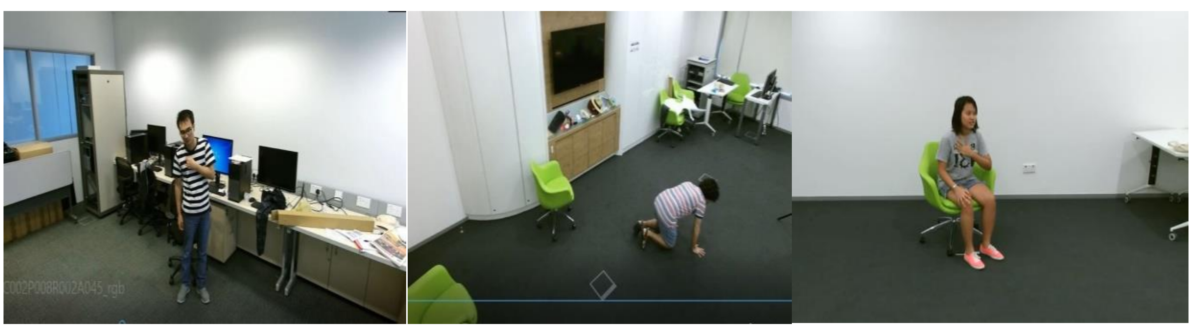
The benchmark dataset NTU RGB+D [20] contains about 56,000 video samples and 4 million frames with 60 action classes. Amir Shahroudy et al. highlight the limitations of most of currently available RGB+D based action recognition benchmarks like clear distinction of class labels, lack of training samples, variety of subjects and proper placement of cameras. Henceforth the NTU RGB+D database has tried to overcome the limitations. Every frame is captured by highly variant camera setting located at horizontal angles -45, 0, +45 fixed at same height. The video frames were captured with three cameras simultaneously. In RGB videos, every frame has been captured with pixel size of 1920 x 1080. The dataset consists of three categories based on: i) Mutual actions, ii) Medical conditions, iii) Daily routine actions. The present work incorporates the Medical condition action category 3D RGB images of chest pain and falling down classes as A45 and A43 respectively shown in Figure 6.

3.3.2 Private Dataset- RMS

A self-made dataset – RMS consists of 3D depth RGB Images captured from CCTV and One Plus 5 smart phone camera by simulating the real life heart attack chest pain and fall scenarios indoors. The dataset encloses images and videos captured under different scenarios of chest pain and fall such as: fall from standing position, sitting on chair posture, walking, sitting on bed position considering the routine human activities. The dataset consists of 1500 images of resolution 4608x3456 captured under different lighting conditions and recording angles. Figure 7 shows the RGB images of our private RMS dataset. Table 3 highlights the description about the dataset used in our work.

Table 3: Complete dataset consisting of Chest pain posture, Partial Fall and Complete fall

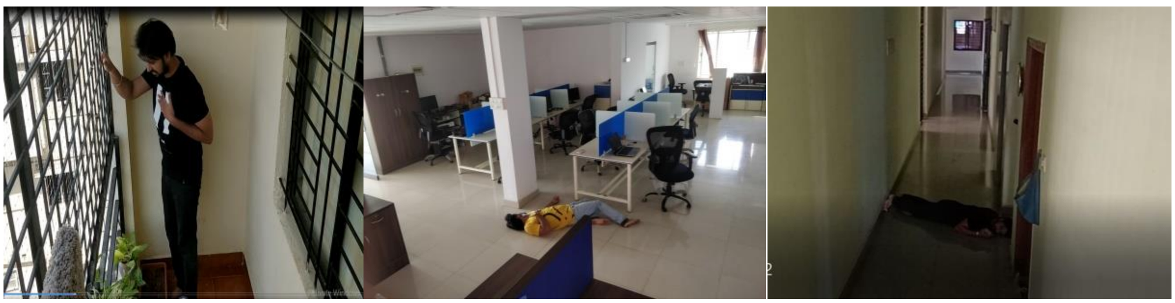
Dataset	Scenario	Number of images	Resolution after preprocessing
NTU RGB+D	Lab	1500	1240x600
Custom RMS	Home, Office, Lab	1500	1067x800



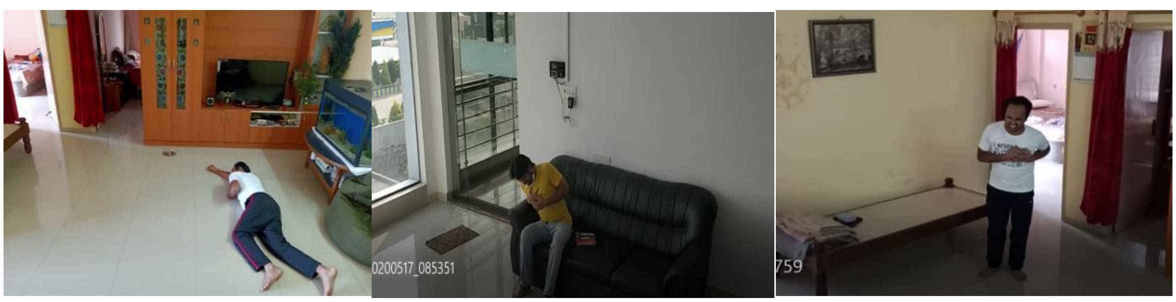
a) b) c)

d) e) f)

Figure 6: Levine’s pose, fall posture images from NTU RGB+D dataset.
a) Levine’s sign standing pose b) Partial fall pose
c) Levine’s sign sitting pose d) Partial fall pose
e) Partial fall pose f) Levine’s sign standing pose



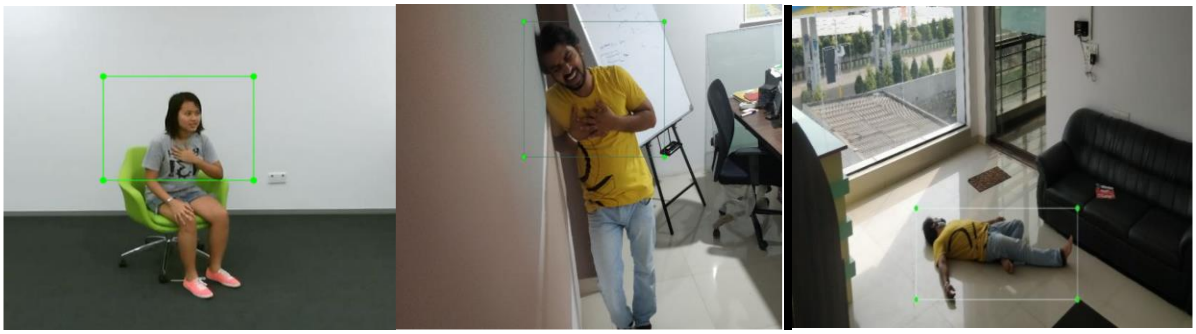
a) b) c)



d) e) f)

Figure 7: Levine’s sign, fall posture images from private RMS dataset
a) Levine’s sign standing pose b) Complete fall pose at office
c) Complete fall pose at corridor d) Chest pain standing pose
e) Levine’s sign sitting pose f) Levine’s sign standing pose

As a preprocessing technique the images obtained from video frames are scaled down in size to reduce the computational burden. Through Scaling process pixel width and height of the images were scaled down to 1067x800 pixels to avoid the quality of original image degradation. The images from both dataset combined were randomly split as train and test set in a ratio of 1:0.3. The annotation procedure is performed using Labellmg tool wherein Region of Interest (ROI) of images are selected as shown in Figure 8. Annotation technique is employed for manual labelling of every training image prior to the training process and the counterpart XML file format for target box location was generated. In this work, the images are classified into three main classes: i) chest pain posture ii) Partial fall iii) Complete fall.



a) b) c)

Figure 8: Region of interest selection for different postures.

3.4 Hardware Description

The present research work includes the concept of Edge AI where the signal acquisition, processing is performed locally on the embedded platform during real time. Training process is performed on a dedicated workstation since the deep learning model training with larger dataset demands high computational power. Later the trained model is deployed on the target hardware, Nvidia Jetson Nano for the inference process.

The state of art neural network architectures applied for performing specialized computations requires parallelized computing graphical processing unit (GPU). The parallel combination of Computational Processing Unit and Graphic Processing Unit is utilized for real word complex applications like object detection to achieve high throughput, accuracy, high bandwidth etc., To achieve high performance gaming and high end graphics applications rendering Nvidia Corporation developed with added techniques a new computed device architecture and cuda DNN library to enrich system performance. Each one of cuda cores or stream processors acts as sub units of GPUs which performs the tasks in parallel independently to accelerate the system performance. The salient features of NVIDIA Jetson Nano is light weight, low power consumption perfectly suited for graphical applications. In general algorithmic approaches implemented using deep learning techniques to train large data will increase the computational cost and memory bandwidth. To achieve this low cost and powerful accelerating hardware need to incorporate, one such hardware device is Jetson Nano NVIDIA. Jetson device is preconfigured with 2GB of reserved swap memory and 4GB total RAM memory. We have utilized the entire swap memory for executing the object detection code to avoid out of memory issue. Figure 9 (a) shows the Jetson Nano device and 9 (b) system interfacing.

a) **b)**

Figure 9: a) Jetson Nano b) Jetson NANO board configuration

3.5 Training

Model training is an automatic technique of parameter fitting in the deep learning network. For efficient training of the proposed network, the CNNs SSD Inception V2 and SSD MobileNet V2 used the pre-trained weights as the initialization from COCO dataset. The Pre-trained networks are being downloaded from official model Zoo - TensorFlow. Deep Neural Network architectures for a large datasets demands high performance computer system. The training of heavyweight DNNs on a powerful Graphics Processing Unit (GPU) for training improves the performance and also reduces the training time. In this present research work, we used computer system with a CPU: Intel(R) core(TM) i7-7700 CPU @3.60GHz, Graphics card: Intel R HD Graphics 630 and 12GB DDR4 RAM. A dedicated high performing GPU can be used for training the computationally expensive deep learning models. The Object detection dataset consists of three main posture classes: i) Chest pain ii) Partial Fall iii) Complete fall; Total data consisting of 3000 training images and 500 test images. Gradient descent optimization function is used to optimize the loss parameter while the learning rate is set as 0.004 and batch size as 24. The training took on average one hundred hours for SSD Inception V2 using the TensorFlow framework. The main termination criteria of training considered was minimum loss function and maximum mean average Precision. After training the deep learning model, it is deployed to Jetson Nano Embedded platform. Figure 9 (b) shows the Jetson Nano set up along with the camera for the real time performance evaluation of the model.

3.6 Performance Evaluation Metrics

COCO Evaluation object detection metrics was used to evaluate our CNN model. Here, bifurcating effectively the various problems of chest pain signs detection as well as fall detection is a major task in this work. So, SSD Inception/MobileNet models are incorporated to achieve detection performance accurately. In this performance evaluation, the key factors that are considered as Average precision, Average Recall, F1 Score, losses and Frames per second.

i) **Confusion Matrix(CM)** It is applied to outline the accomplishment of a classification model:

(a) True Positives (TP): Number of both true cases for classifier prediction and correctness of the class to point out as ground-truth bounding box.
(b) False Positives (FP) (Type I error): Indicates the count cases for True classifier prediction and false class in correction leads to improper object detection.
(c) False Negatives (FN) (Type II error): Indicates the count cases for False classifier prediction and True class in correction leads to improper object detection
(d) True Negatives (TN): Number of both false cases for classifier prediction and correctness of the class to point out as ground-truth bounding box.

It is observed from the literature survey that a true negative results in object detection provides only marginal information based on number of bounding boxes in the image/scene. Attention to be made to incorporate this ideology to enrich the performance of object detection analysis process. Figure 10 indicates the confusion matrix.

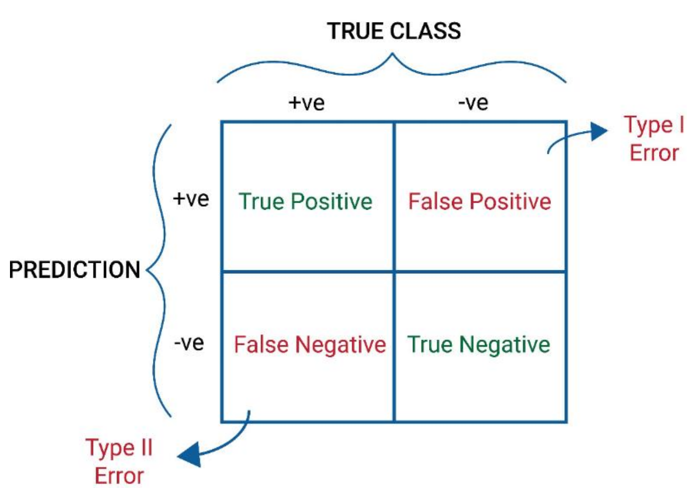


Figure 10: Confusion Matrix

ii) *Intersection over Union* (IoU): It is measurement of overlapping area using two bounding boxes: ground truth box $\boldsymbol{B_p}$ and predicted box $\boldsymbol{B_{gt}}$.

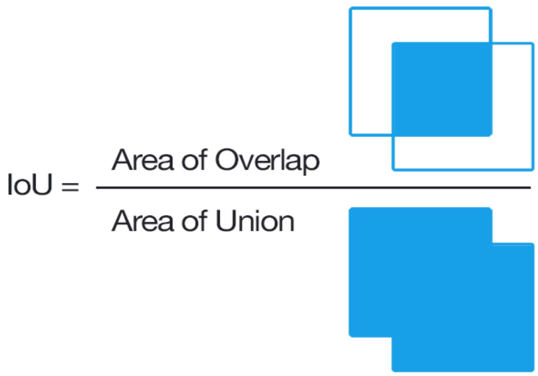


Figure 11: Intersection over Union

Finally, IoU can be used to measure the possibility of True Positive and/or True Negative cases in the detection process. This is defined as the ratio of an intersection area vs overlapped area of both bounding boxes (predicted and ground truth) as depicted in figure 11. The IOU is measured by using the equation 2 as,

$$\mathrm{IOU} = \frac{\mathrm{Area}(B_p \cap B_{gt})}{\mathrm{Area}(B_p \cup B_{gt})} \tag{2}$$

IOU threshold provides a metric to estimate the level of intersection between predicted bounding box and ground truth which in-turn helps is estimating True Positive, False Positive, True Negative and False Negative cases. For example, a value of 0.75 IoU threshold means that the intersection of both the bounding box is above 0.75%. While the case is considered true positive if the threshold of 0.75 is exceeded else marked as False positive. Depending upon the object detection application the threshold value is set for IoU in accurate decision making of true positive and False Positive. IoU with the defined threshold will be able to show the perfectness of overlapped bounding box areas.

iii) **Precision (P)**

Precision (P) is the measurement of positive predictions accurately. Precision is estimated by the ratio of true positives to the sum total of the positive predictions that provides the total positive predicted values. It is the similarity indexing factor of the machine learning network in evaluating the defined relevant objects. This metric is estimated using the equation 3 as,

$$\mathrm{P} = \frac{\mathrm{TP}}{\mathrm{TP+FP}} = \frac{\mathrm{TP}}{\text{Total number of ground truths}} \tag{3}$$

iv) **Recall (R)**

Recall (R) is the ratio of number of true positive cases detected against sum of true positive and false negative predictions. It indicates all relevant ground truth bounding boxes of a model. Recall predicts the true positive rate or sensitivity calculated by the ratio of true positive values (TP) and the sum total of true positives and false negatives. Recall is one of the performance evaluation metric of object detection model to determine all realizable ground-truth bounding boxes. Recall is measured using equation 4 as,

$$\mathrm{R} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} = \frac{\mathrm{TP}}{\text{Number of predictions}} \tag{4}$$

To measure the sufficient proposed overlapped bounding box overlaps reference to the ground truth and the evaluation metric intersection over union (IoU), is adopted to evaluate the accuracy in detecting the object.

v) **Average Precision (AP)**

The average precision (AP) is the mean value of precision reference to defined recalls (j).
For a given set of N images and stationary values of each IOU helps in evaluating the mean of AP in detection process. The metric AP is estimated using equation 5 as,

$$\mathrm{AP} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}}\frac{1}{\mathrm{M}}\sum_{\mathrm{j}=1}^{\mathrm{M}}\mathrm{Precision_i}(\mathrm{Recalls_j}) \quad (5)$$

Here, $Precision_i$ is a function of recalls$_j$. This Precision metric performs a major role in object detection model with prediction score value of each object to highlight the confidence level.

vi) **F1 Score**

The F1 score is expressed by a real integer that highlights the accomplishment rate is expressed by combination of two quantitative parameters: precision and recall rate. These two metrics plays a major role in estimating F1 score of the framework, and F1 score is calculated using equation 6 as follows

$$\mathrm{F1\ Score} = 2 * \frac{\mathrm{Precision} * \mathrm{Recall}}{\mathrm{Precision}+\mathrm{Recall}} \quad (6)$$

vi) **Losses**

The overall loss is splits into two main losses, i) localization loss, ii) Confidence loss. The localization loss gives an estimate of the mismatch between the final predicted bounding box and the ground truth box. SSD model mainly considers the predictions from Positive matches which are closer to ground truth boxes, there by ignoring the negative matches. The confidence loss is a loss value while performing the class prediction. It is a measure of the confidence of a network while estimating the object-ness of the computed bounding box. The total loss factor of SSD network can be evaluated from equation 7.

Let, x_{ij}^{p} = {1,0}; x_{ij}^{p} is a measure used for comparison the i^{th} default box to the j^{th} ground truth box of category P.

In this matching plan of action, $if\ \sum_i x_{ij}^{p} \geq 1$, the overall loss function $L(x,c,l,g)$ is expressed in terms of weighted sum of the localization (loc) and the confidence loss (conf),

$$L(x,c,l,g) = \frac{1}{N}[L_{conf}(\mathrm{x,c}) + \propto L_{loc}(x,l,g)] \quad (7)$$

Where,

N is the number of matched default boxes. If N = 0, we set the loss to 0.
The localization loss is a smooth L_1 loss between the predicted box (l) & the ground truth box (g) parameters.

We regress to offsets for the center (C_x, C_y) of the default bounding box (d) and for its width (w) and height (h).

$$L_{loc}(x,l,g) = \sum_{i\in Pos\ m\in\{C_x,C_y,w,h\}}^{N} \sum x_{ij}^{k}\ smooth_{L_1}(l_i^m - g_j^m) \quad (8)$$

Where,

$$g_j^{cx} = (g_j^{cx} - d_i^{cx})/d_i^{w} \qquad g_j^{cy} = (g_j^{cy} - d_i^{cy})/d_i^{h}$$

$$g_j^{w} = \log\left(\frac{g_j^{w}}{d_i^{w}}\right) \qquad g_j^{h} = \log\left(\frac{g_j^{h}}{d_i^{h}}\right)$$

The confidence loss is the softmax loss over multiple class confidences(c).

$$L_{conf}(x, c) = - \sum_{i\epsilon Pos}^{N} x_{ij}^{P} \log(C_i^{p}) - \sum_{i\epsilon N_{eg}}^{N} \log(C_i^{0}) \tag{9}$$

Where,

$$C_i^{P} = \frac{\exp(C_i^{p})}{\sum_p exp(C_i^{p})}$$

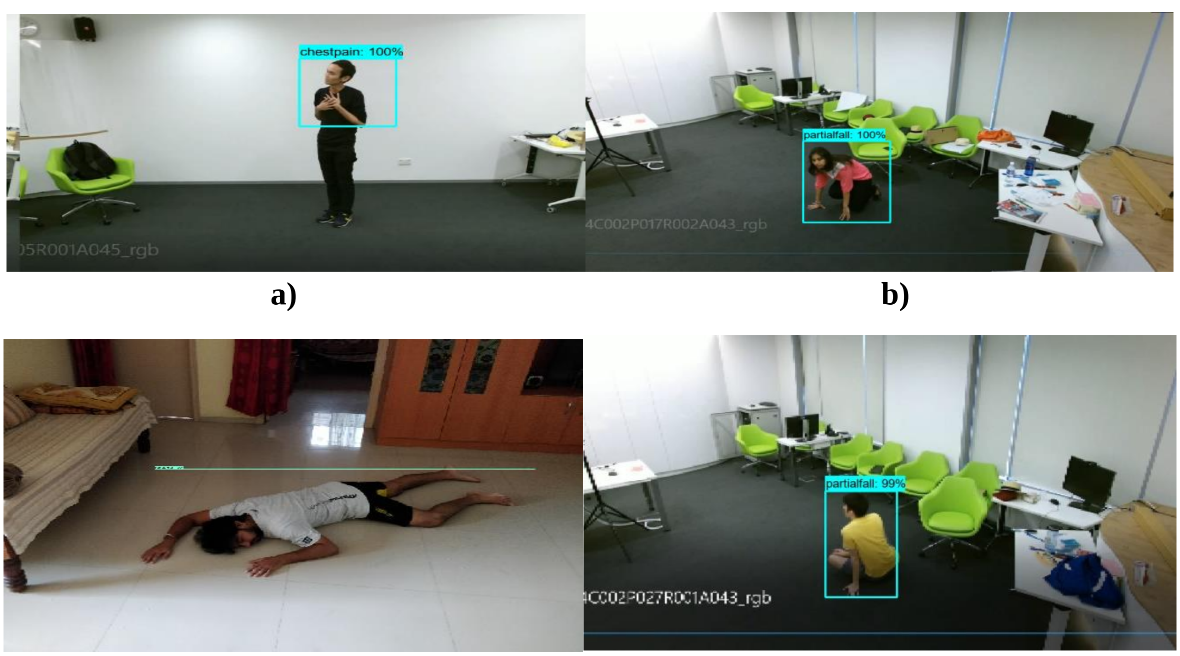
and the weight term ∝ is set to 1 by cross validation.

vii) **Frames Per Second(FPS)**

Frames Per Second is a unit which measures camera performance. Frame rate indicates amount of individual video frames that a camera captures, per second. FPS provides a performance measurement of motion videos on a display device.

4. Results and Discussion

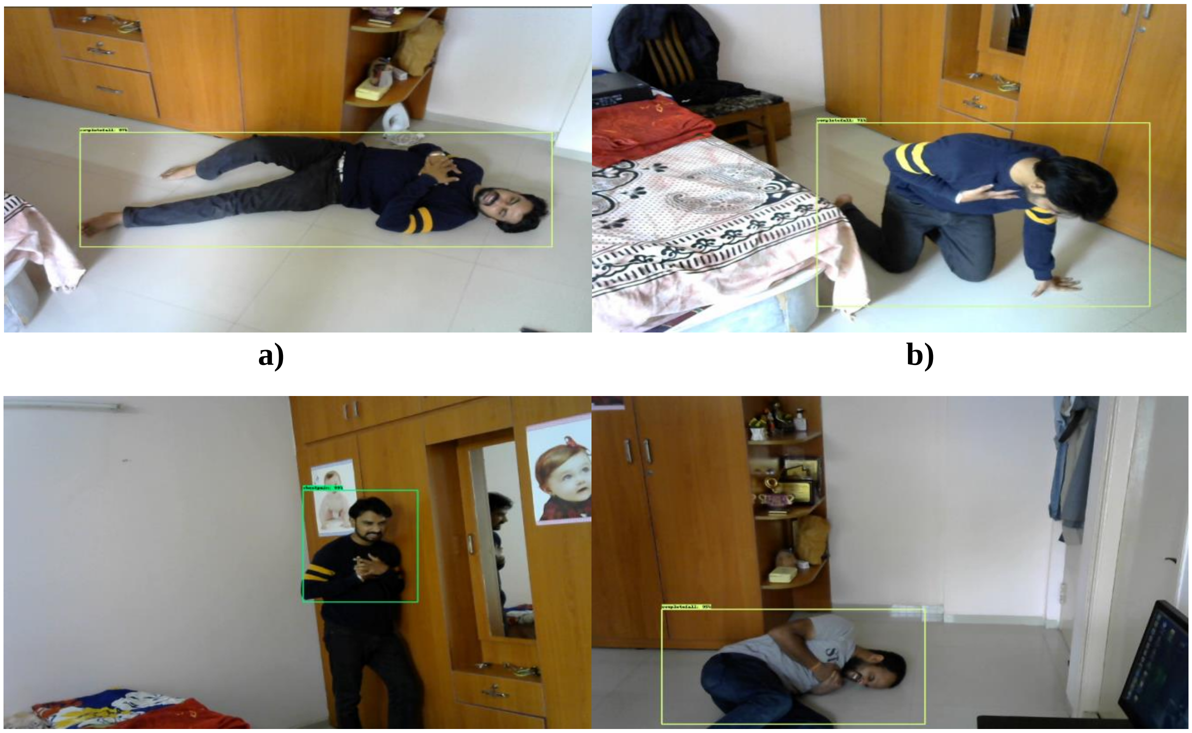
The deep learning CNN algorithmic model experiments were performed to examine the efficiency of the advocated neural network model for the application of vital signs of MI detection. The model utilizes Tensor-Flow library and Keras APIs and prototype model is built using python programming. The network was developed using deep learning framework is optimized to run on the parallelized CUDA architecture NVIDIA GPU for executing the kernels. The scores of the predicted box are displayed along with the chest pain Levin's posture and fall postures for the trained CNN SSD Inception V2 network as shown in Figure 12.



a) b) c) d)

Figure 12: Predicted simulated results of SSD Inception V2 for three classes of Vital signs Myocardial Infraction

A single camera modality connected to Jetson Nano is used for the real time inferencing. The trained CNN model successfully performs both classification and localization where in the Levine's sign and fall detections were identified. Figure 13 shows the real time detections of chest pain posture and fall along with their predicted score values.



a) b)

c) d)

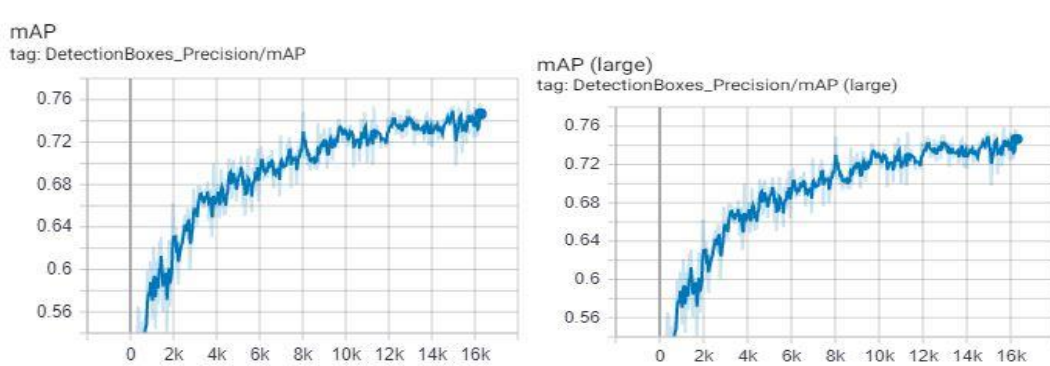
Figure 13: Real time detection of vital signs of Myocardial Infarction Levine's sign posture and fall condition postures with their score values.

a) Complete fall posture in living room conditions b) partial fall posture

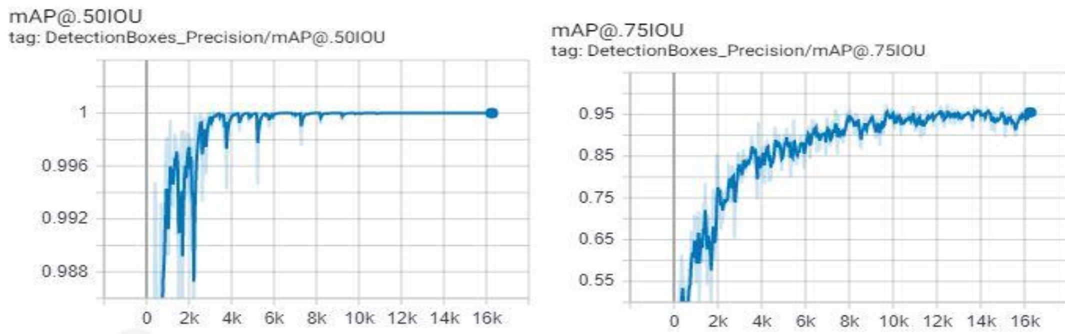
c) Levine's sign identification in indoor environment d) chest pain posture

i) Precision Evaluation

In computer vision object detection, the main performance metric measurement is mean Average precision. Considering COCO benchmark performance metric both Average precision and mean Average precision are evaluated as the same unique measure. mean Average Precision(mAP) values are plotted considering the IoU values ranging from 0.5 to 0.95 with an incremental step size values of 0.05 as shown in Figure 14 (a) and (b). The mAP graph plot for an IoU values of constant 0.5 and 0.75 are shown in Figure 14 (c) and (d) respectively. Table 4 indicated the results of mAP value, mAP large value and mAP at 0.5IOU and at 0.75 IOU.



a) b)



c) **d)**

Figure 14: mean Average Precision values of SSD InceptionNet V2 at various IOU

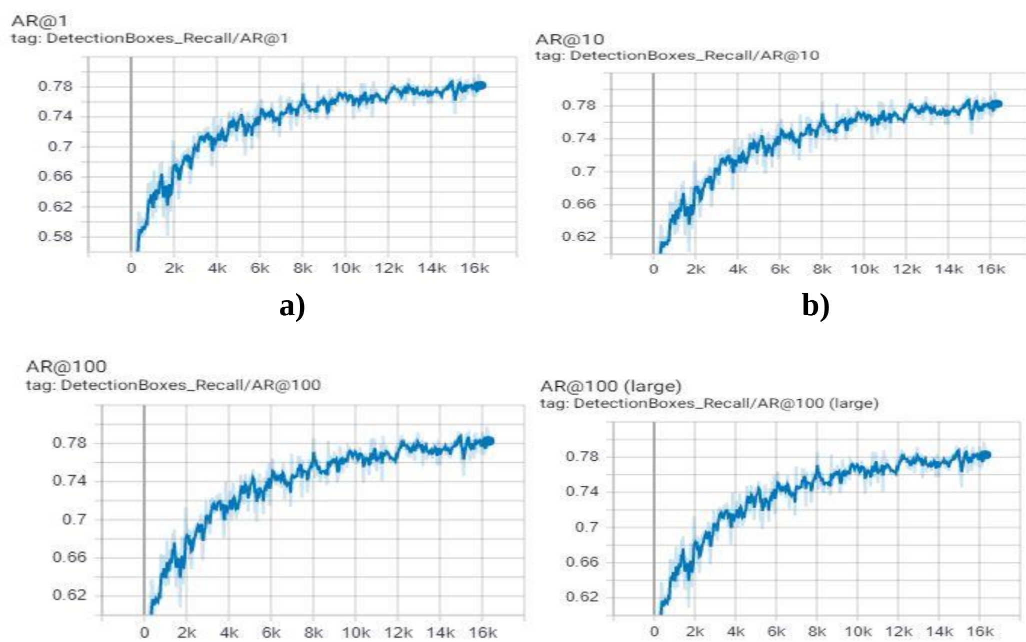
a) mAP values b) mAP (large) values c) mAP@0.5IOU d) mAP@0.75IOU

Table 4: Precision values of SSD Inception V2 at different IOU

No	Evaluation metric	Value
1	mean Average Precision (*mAP*)	76.4%
2	mean Average Precision (large)	76.4%
3	mean Average Precision @.50*IOU*	100%
4	mean Average Precision @.75*IOU*	96.5%

ii) Average Recall Evaluation

The COCO object detection criteria highlights about predefined areas of various sizes of objects in the image for detection process. Average recall evaluation designates total number of detection per image considering the object sizes: i) Size of object are less than 32^2 pixels is considered to be smaller. ii) Object size between 32^2 and 96^2 pixels are considered as medium, iii) Size of object greater than 96^2 pixels are treated as large. The standard areas mentioned is compared with the segmentation mask in an image for object detection process. Figure 15 show the graphical plots of Average recall values with different conditions of total detections per image as: 1, 10 and 100. The tabulation of results are shown in table 5.



a) **b)** **c)** **d)**

Figure 15: Average recall values considering number of detections per image

a) AR@1 values b) AR@10 values c) AR@100 values d) AR@100(large) values

Table 5: Average Recall values of SSD Inception V2

Sl. No	Evaluation metric	Value
1	Average Recall@1	80.0%
2	Average Recall @10	80.0%
3	Average Recall @100	80.0%
4	Average Recall @100 (large)	80.0%

iii) F1 Score Evaluation

F1 score is a measure of weighted average or harmonic mean between precision and recall values. F1 score/F1 measure mainly considers false positives and false negatives values and the range is between 0 and 1. Highest value of F1 score indicates low false positives and false negatives suggesting less false alarms in the model. The main aim of an object detection model is to achieve high precision, recall values in-tern obtaining high F1 score Table 6 highlights the mAP, Recall and F1 score values.

Table 6: mean Average Precision, Recall and F1 Score values of SSD MobileNet V2 and SSD Inception V2

Backbone DCNN	mean Average Precision%	Recall%	F1 Score%
SSD Inception V2 COCO	76.4	80.0	78.1
SSD Mobilenet V2 COCO	68.7	72.8	70.6

iv) Loss Function Evaluation

The overall loss function SSD network is evaluated as: classification, regularization and localization loss. The main objective of training our deep learning object detection model is to minimize the error function in-tern by reducing the total loss of the network. Figure 16 depicts the decreasing values of different losses values of the network and tabulated in Table 7.

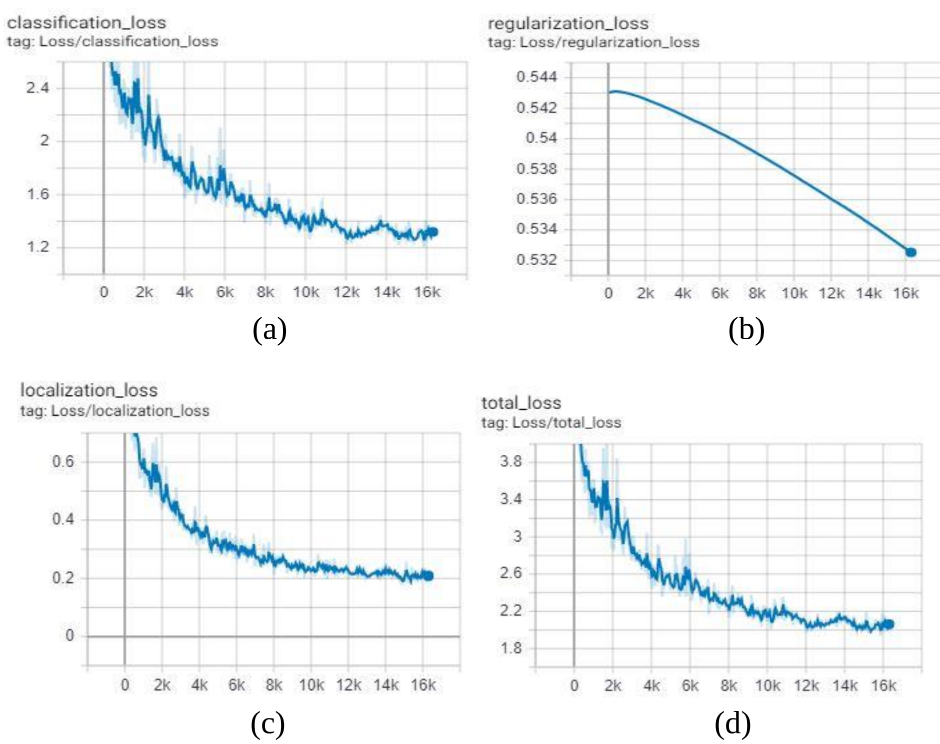


Figure 16: Loss curves of SSD Inception V2 model

Table 7: Loss factors measured by SSD Inception V2

Type of loss	Loss Type	Loss value
1	Classification loss	1.2836425
2	Regularization loss	0.5324396
3	Localization loss	0.19522505
4	Final Total loss	2.0113058

v) Training time Comparison

Training of two image datasets NTU RGB+D and RMS dataset is performed in the workstation with two CNN models. Table 8 shows training time taken for sixteen thousand steps. SSD MobileNet V2 COCO model takes more time in training compared to SSD Inception V2 COCO.

Table 8: Measurement of training time, number of steps of different SSD models

Convolution Neural Networks	Training Time (Hrs)	Number of Steps(K)
SSD Inception V2 COCO	93	16
SSD MobileNet V2 COCO	110	16

vi) Embedded Implementation

NVIDIA's Jetson Nano development kit platform incorporates neural network libraries and frameworks to efficiently implement the computer vision models practically. The training of the SSD CNN model is performed on the workstation and the model has been deployed on Jetson hardware platform, the performance was being tested in terms of frame rate. Table 9 shows the frames per second for two different CNN architectures that are implemented.

Table 9: Measurement of frames per Second

Edge Device	CNN Model	Power consumption – Watts	Frame per second
Jetson NANO	SSD MobileNet V2	10	2.135
		5	1.042
	SSD Inception V2	10	2.013
		5	0.996

4.1 Results Comparison with other work

From the extensive research survey that was carried out by the authors, a comparison is made with the object detection algorithmic based fall detection approaches. Table 10 gives the comparison with Xiang Wang [48] and Kun-Lin Lu [49]. Table 11 shows a performance comparison in terms of frames per second considering different hardware utilized to solve fall detection problem.

Table 10: Performance comparison with related works.

Paper	Model	Dataset	mean Average Precision%
Xiang Wang [48]	SSD321 ResNet101	VOC2007+2012	77.1
	SSD500 ResNet101	VOC2007+2012	80.6
Kun-Lin Lu [49]	SSD_MobileNet_V1_coco	Private	40.3
The Proposed work	SSD Inception V2 coco	Private RMS database and NTU RGB	76.4

Table 11: Time cost indicated in terms of Frames Per Second

Paper	Camera	Hardware	Frames per second
Rougier et al. [50]	4	Core 2 Duo	5
Yun et al. [51]	1	Core i7	0.076
Feng et al. [52]	1	Core i5	11
The proposed work	1	Core i7	7.583
The proposed work	1	Jetson Nano	2.135

4.2 Discussion

Computer vision domain has achieved a stupendous success lately and has attracted researchers to solve challenging applications of object detection. In this paper the authors made an attempt to evaluate the state-of-the art DNN object detection algorithms, SSD MobileNet and SSD InceptionNetV2 for the vital signs of heart attacks. This proposed work contemplates the image dataset from RGB videos of NTU RGB+D and proposed synthetic RMS database captured from high resolution cameras. The three prime possible vital sign postures of chest pain and fall were being analyzed and the ConvNet model was developed and the performance analysis is carried out. The three possible conditions of heart attack postures simulated helps in understanding the severity of the pain. This acts as a pain estimate to call for an emergency situation and help in diagnosing the patient at the earliest. The posture based data at the place of heart attack can act as a primary report for diagnosis. The work can be considered as an impressive example of GPU and CPU cooperation for implementing the deep learning architecture that enables highly accurate detection with lesser computational cost in more economical way. Through the experiments on Jetson Nano, real time performance is evaluated by considering the shortcomings of object detection algorithm on the embedded platforms. Our proposed MI vital signs detection system can run Inception V2 SSD and MobileNet V2 SSD CNN model on embedded GPU platform at frames per second that can be considered for practical implementation to emergency fall situations. We investigated different models such as SSD ResNet50 and SSD ResNet101 for training using the same database and deploying to Jetson Nano device. Although both models showed significant improvement in mAP and Recall values, the deep learning architectural is a complex and deploying on Jetson Nano edge device consumes lot of memory and shows extreme low frames per second on real time inference. The performance can be further enhanced through better optimization techniques using precision and inferencing using TensorRT for edge platform. During the course of experimentation Jetson Nano device encounters a rise in temperature for long durations of work with the large datasets. This overheating of the device could be avoided by installing a suitable ventilation system prescribed by NVIDIA.

4.3 Limitations

There are few shortcomings found in the present research work discussed as follows: Firstly number of RGB images used for training is 3000 from both private RMS and public NTU+RGB datasets combined. More number of images can be considered for training to enrich the performance of the DL-CNN model. Secondly the training images can be incorporated with data augmentation techniques to enrich the training of the CNN model which in turn could increase the performance metrics. Thirdly our model fails to recognize falls under extreme low light conditions. Finally high performance powerful GPU can be incorporated to reduce the number of training hours for the present object detection model.

5 Conclusion and Future work

Artificial intelligence based pain management strategies and automated fall identification through a specialist system, is an advancing area of research in smart health informatics. The AI procedure advocated in this present work can have useful implications for the medical diagnostic domain and opens up new possibilities for automatic pain therapeutics practices considering medical practitioners and other health care researchers. In this study, we propose a supervised learning object detection method from 3D RGB for enhancing the performance of vital signs MI fall detection. A state-of-the-art lightweight CNN structure InceptionNet V2 SSD and MobileNet V2 SSD is put forward for training the Levine's sign posture and fall posture RGB images from video frames for classification. In this proposed DNN CNN object detection model five performance parameters were estimated for optimum performance in Levine's chest pain posture, partial fall and complete fall. The performance evaluation highlights that InceptionNet V2 SSD can attain mean Average precision of (76.4%) and Recall (80%). The experimental results shows that our network can be used as a practical setting for real-time vital signs myocardial Infarction detection with GPU embedded implementation. The results highlights that the adopted deep learning model performs better than other existing object detection lightweight classification models. In the future work, the spatiotemporal video analysis will be considered for further enhancing the object detection model performance and to develop an intelligent video surveillance alarm system as a smart health care for detecting the emergency situation encountered with the heart attack falls for early assistance.

Acknowledgement

The authors would like to express sincere thanks to the organization Digital Shark Technology, Bangalore for providing hardware resources during the implementation of this work.

Declaration of conflict of interest

The authors declare that there is no conflict of interest.

References

1. Cristina Balla Rita Pavasini Roberto Ferrari, "Treatment of Angina: Where Are We?", Cardiology 2018, Vol. 140, pp 52–67, doi: 10.1159/000487936
2. Kurt J. Greenlund, Nora L. Keenan, Wayne H. Giles, MD, "Public recognition of major signs and symptoms of heart attack: seventeen states and the US Virgin Islands", Vol. 147, pp 1010-1016, doi:10.1016/j.ahj.2003.12.036
3. Karen L Smith, Peter A Cameron, Alastair Meyer and John J McNeil, "Knowledge of heart attack symptoms in a community survey of Victoria", Emergency Medicine Vol. 14, pp 255–260, 2002, doi: 10.1046/j.1442-2026.2002.00340.x
4. Johan herlitz, Ake hjalmarson, Finn waagstein, "Treatment of pain in acute myocardial infarction," British Heart Journal,Vol 61, pp 9-13, 1989, doi: 10.1136/hrt.61.1.9
5. James h. Behrmann., Harold r. Hipp, and Howard e. Heyer, "Pain Patterns in Acute Myocardial Infarction" American Journal of Medicine, Vol. 9, Issue 2, August 1950, Pages 156-163
6. Mair, Johannes; Puschendorf, Bernd; Smidt, Jörn; Lechleitner, Peter; Dienstl, Franz(1995)."A Decision Tree for the Early Diagnosis of Acute Myocardial Infarction in Non-traumatic Chest Pain Patients at Hospital Admission"Chest, 108(6), 1502–1509. doi:10.1378/chest.108.6.1502
7. Dr. Alan Leviton, Further Comments on the Levine Sign, July 29, 1965, N Engl J Med 1965;273:282 doi: 10.1056/NEJM196507292730523
8. W. M. Edmondstone (1995) Cardiac chest pain: Does body language help the diagnosis? 311(7021): 1660–1661, doi: 10.1136/bmj.311.7021.1660
9. Gregory M Marcus, Joshua Cohen, Paul D Varosy, The Utility of Gestures in Patients with Chest Discomfort, The American Journal of Medicine, Volume 120, Issue 1, January 2007, Pages 83-89, doi: 10.1016/j.amjmed.2006.05.045
10. World Health Organization, World Health Statistics Overview 2019, www.who.int/gho/publications/world_health_statistics/2019/en/

11. Delahoz, Y.S.; Labrador, M.A. Survey on Fall Detection and Fall Prevention Using Wearable and External Sensors. *Sensors* 2014, *14*, 19806-19842.
12. Guideline for the Prevention of Falls in Older Persons, Journal of the American Geriatrics Society 49(5):664 – 672, Dec 2001, doi: 10.1046/j.1532-5415.2001.49115.x
13. A systematic review on the influence of fear of falling on quality of life in older people: is there a role for falls, Clin Interv Aging, 2019 Apr 24;14:701-719. doi: 10.2147/CIA.S197857
14. Georg H. Eifert "Cardiophobia: An Anxiety Disorder in Its Own Right?" Behaviour Change , Volume 8 , Issue 3 , September 1991 , pp. 100 – 116 doi: https://doi.org/10.1017/S0813483900006690
15. Maw Pin Tan and Rose Anne Kenny, Cardiovascular Assessment of Falls in Older People, Clin Interv Aging. 2006 Mar; 1(1): 57–66, doi: 10.2147/ciia.2006.1.1.57
16. Trends in IoT based solutions for health care: Moving AI to the edge, Pattern Recognition Letters, Volume 135, July 2020, Pages 346-353, doi: 10.1016/j.patrec.2020.05.016
17. Ghulam Muhammed, Mohammed F Alhamid, "Edge computing with cloud for voice disorder assessment and treatment" IEEE Communications Magazine, Vol. 56, Issue 4, pp. 60-65, April 2018, doi: 10.1109/MCOM.2018.1700790.
18. J Pena Queralta, T N Gia, H Tenhunen, "Edge-AI in LoRa-based Health Monitoring: Fall detection system with for computing and LSTM recurrent Neural Networks" 2019 42nd International Conference on Telecommunications and Signal Processing (TSP), July 2019, doi: 10.1109/TSP.2019.8768883.
19. Xiangfeng Dai, Irena Spacic, Bradley Meyer "Machine Learning on Mobile: An on device inference app for skin cancer detection", 4th Int Conf. on Fog and Mobile Edge Computing (FMEC) June 2019, doi: 10.1109/FMEC2019 .8795362.
20. H. Mao, S. Yao, T. Tang, B. Li, J. Yao, Y. Wang, "Towards Real-Time Object Detection on Embedded Systems," in IEEE Transactions on Emerging Topics in Computing, Vol. 6, pp 417-431,2018, doi: 10.1109/TETC.2016.2593643.
21. Vittorio Mazzia, Aleem Khaliq, Francesco Salvetti, Marcello Chiaberge "Real-Time Apple Detection System using Embedded Systems with Hardware Accelerators: an Edge AI Application", Vol. 4, pp 1-13, 2016, doi: 10.1109/ACCESS.2017.
22. Luis Barba-Guaman, Jose Eugenio Naranjo, Anthony Ortiz "Deep Learning Framework for Vehicle and Pedestrian Detection in Rural Roads on an Embedded GPU" Vol 4, pp 1-17, 2020, doi:10.3390/electronics9040589
23. Victor Partel, Sri Charan Kakarla, Yiannis Ampatzidis "Development and evaluation of a low-cost and smart technology for precision weed management utilizing artificial intelligence" Computers and Electronics in Agriculture, Elsevier, Vol 157, pp 339–350,2019, doi:10.1016/j.compag.2018.12.048
24. Choi, Min-Kook & Park, Jaehyung & Jung, Heechul & Lee, Jin-Hee & Eo, Soo-Heang. "Fast and Accurate Convolutional Object Detectors for Real-time Embedded Platforms" Vol. 1, 2019
25. Ajeet Ram Pathak, Manjusha Pandey, Siddharth Rautaray, "Application of Deep Learning for Object Detection" International Conference on Computational Intelligence and Data Science (ICCIDS 2018), Procedia Computer Science, pp 1706–1717, Vol 132,2018
26. Weiming Chen, Zijie Jiang, Hailin Guo, Fall Detection Based on Key Points of Human-Skeleton Using Open Pose, Symmetry 2020, 12(5), 744; doi: 10.3390/sym12050744
27. Saleh, M. and Jeanne`s, R. L. B. (2019). Elderly fall detection using wearable sensors: A low cost highly accurate algorithm. IEEE Sensors Journal 19, 3156–3164, doi: 10.1109/JSEN.2019.2891128
28. Zitouni, M., Pan, Q., Brulin, D., Campo, E., et al. (2019). Design of a smart sole with advanced fall detection algorithm. Journal of Sensor Technology 9, 71, doi: 10.4236/jst.2019.94007
29. Wu, T., Gu, Y., Chen, Y., Xiao, Y., and Wang, J. (2019). A mobile cloud collaboration fall detection system based on ensemble learning, arXiv:1907.04788
30. Huang, Y, Chen, W, Chen, H., Wang, L, and Wu, K. (2019). G-fall: Device-free and training-free fall detection with geophones. In 2019 16th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON), 1–9, doi: 10.1109/SAHCN.2019.8824827
31. Tian, Y, Lee, G.-H, He, H, Hsu, C.-Y, and Katabi, D. (2018). Rf-based fall monitoring using

convolutional neural networks. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 2, 1–24

32. Wang, Y., Wu, K., and Ni, L. M. (2017b). Wifall: Device-free fall detection by wireless networks. IEEE Transactions on Mobile Computing 16, 581–594

33. Kerdjidj, O., Ramzan, N., Ghanem, K., Amira, A., and Chouireb, F. (2020). Fall detection and human activity classification using wearable sensors and compressed sensing. Journal of Ambient Intelligence and Humanized Computing 11, 349–361

34. Queralta, J. P., Gia, T., Tenhunen, H., and Westerlund, T. (2019). Edge-ai in lora-based health monitoring: fall detection system with fog computing and lstm recurrent neural networks. In 2019 42nd International Conference on Telecommunications and Signal Processing (TSP) (IEEE), 601–604

35. Han, Q., Zhao, H., Min, W., Cui, H., Zhou, X., Zuo, K., et al. (2020). A two-stream approach to fall detection with mobilevgg. IEEE Access 8, 17556–17566

36. Kong, Y., Huang, J., Huang, S., Wei, Z., and Wang, S. (2019). Learning spatiotemporal representations for human fall detection in surveillance video. Journal of Visual Communication and Image Representation 59, 215–230

37. Ko, M., Kim, S., Kim, M., and Kim, K. (2018). A novel approach for outdoor fall detection using multidimensional features from a single camera. Applied Sciences 8, 984

38. Shojaei-Hashemi, A., Nasiopoulos, P., Little, J. J., and Pourazad, M. T. (2018). Video-based human fall detection in smart homes using deep learning. In 2018 IEEE International Symposium on Circuits and Systems (ISCAS) (IEEE), 1–5

39. Min, W., Yao, L., Lin, Z., and Liu, L. (2018). Support vector machine approach to fall recognition based on simplified expression of human skeleton action and fast detection of start key frame using torso angle. IET Computer Vision 12, 1133–1140

40. Ozcan, K., Velipasalar, S., and Varshney, P. K. (2017). Autonomous fall detection with wearable cameras by using relative entropy distance measure. IEEE Transactions on Human-Machine Systems 47, 31–39

41. Gabriel Rojas Albarracín, Miguel Angel Chaves, Antonio Fernandez Caballero, "Heart Attack Detection in Colour Images Using Convolutional Neural Networks" Appl. Sci. Vol 9, pp 5065-5074, 2019, doi:10.3390/app9235065.

42. Deep Learning based systems developed for fall detection: Review, IEEE Access, Vol. 8, **pp.** 166117 – 166137, **doi:** 10.1109/ACCESS.2020.3021943

43. Shenshen Gu, Xinyi Chen, Wei Zeng, and Xin Wang "A Deep Learning Tennis Ball Collection Robot and the Implementation on NVIDIA Jetson TX1 Board" IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), Auckland, New Zealand, pp-170-175, 2018

44. Tijtgat, Nils, Wiebe Van Ranst, Bruno Volckaert, Toon Goedeme, "Embedded Real-Time Object Detection for a UAV Warning System" 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), pp 2110-2118, 2017

45. Howard, Andrew Zhu, Menglong Chen, Bo Kalenichenko, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications" pp 1-7, 2017

46. Wei Liu, Dragomir Anguelov, Dumitru Erhan, "SSD: Single Shot MultiBox Detector" Computer Vision and Pattern Recognition, 2016, doi: 10.1007/978-3-319-46448-0_2

47. Lin, Zitnick, Doll, "Microsoft COCO: Common Objects in Context", Computer Vision, ECCV, pp 740-755, 2015, arXiv:1405.0312.

48. Xiang Wang, Kebin Jia (2020) Human Fall Detection Algorithm Based on YOLOv3. IEEE 5th International Conference on Image, Vision and Computing,

49. Kun-Lin Lu, Edward T.-H. Chu (2018) An Image-Based Fall Detection System for the Elderly. Applied Sciences, 8(10):1995-2026

50. C. Rougier, J. Meunier, A. St-Arnaud, J. Rousseau, Robust video surveillance for fall detection based on human shape deformation, IEEE Transactions on Circuits and Systems for Video Technology. 21 (5) (2011) 611– 622.

51. Y. Yun, I. Gu, Human fall detection in videos via boosting and fusing statistical features of appearance,

shape and motion dynamics on Riemannian manifolds with applications to assisted living, Computer Vision and Image Understanding. 148 (2016) 111–122.
52. W. Feng, R. Liu, M. Zhu, Fall detection for elderly person care in a vision based home surveillance environment using a monocular camera, Signal, Image and Video Processing. 8 (6) (2014) 1129–1138.