VOLUME 16, N° 3 2022
Journal of Automation, Mobile Robotics and Intelligent Systems

Edge Artificial Intelligence-Based Facial Pain Recognition During Myocardial Infarction

Submitted: 23rd April 2021; accepted: 11th October 2021

Mohan H M, Shivaraj Kumara H C, Mallikarjun S H, Dr. Prasad A Y

DOI: 10.14313/JAMRIS/3-2022/23

Abstract:
Medical history highlights that myocardial infarction is one of the leading factors of death in human beings. Angina pectoris is a prominent vital sign of myocardial infarction. Medical reports suggest that experiencing chest pain during heart attacks causes changes in facial muscles, resulting in variations in patterns of facial expression. This work intends to develop an automatic facial expression detection to identify the severity of chest pain as a vital sign of MI, using an algorithmic approach that is implemented with a state-of-the-art convolutional neural network (CNN). The advanced object detection lightweight CNN models are as follows: Single Shot Detector Mobile Net V2, and Single Shot Detector Inception V2, which were utilized for designing the vital signs MI model from the 500 Red Blue Green Color images private dataset. The authors developed cardiac emergency health monitoring care using an Edge Artificial Intelligence ("Edge AI") using NVIDIA's Jetson Nano embedded GPU platform. The proposed model is mainly focused on the factors of low cost and less power consumption for onboard real-time detection of vital signs of myocardial infarction. The evaluated metrics achieve a mean Average Precision of 85.18%, Average Recall of 88.32%, and 6.85 frames per second for the generated detections.

Keywords: *Vital Signs, Myocardial Infarction, Facial Pain Expression, Computer Vision, Medical Assistance, Convolution Neural Network*

1. Introduction

The primary cause of human death across the globe is heart disease, specifically ischemia, and angina pectoris (chest pain) is its most common symptom [1]. Pain is a distressing experience, with actual or potential tissue damage associated with sensory, emotional, cognitive, and social components [2]. Pain is a publicly displayed visible event (usually demanding attention) and the facial expressions allow the observer to appropriately respond. Characteristics may be discovered on the observational scale of pain (sharp, intense or unusual) that may help to identify a warning signal as a potential danger threat. The patterns in facial pain expression induce social responses such as care, empathy, and nursing [3, 4]. Without pain, the total lifespan of human beings will be reduced drastically [5]. Pain is a multi-dimensional representation that incorporates behavioral, physiological, sociocultural, cognitive, and affective characteristics [6].

Clinicians perform their initial investigation through the patient's self-report, taking into consideration such as location, severity, sensory quality, temporal features, and aspects that escalate or diminish pain. A conscious verbal patient's self-report gathering can be adapted by different modes like verbal communication, gestures, nodding the head for a question, or writing. A non-verbal patient's self-report information collection can include searching for causes of pain, keen observation of patient behavior, the report from the patient caretaker, or through an analgesic trial. In a manual process, there are chances of missing certain information in a short interval of time due to high patient density in hospitals, miscommunication, late detection, and human reading errors leading to a faulty diagnosis. It is very necessary for the patient's diagnosis process to develop an accurate and automatic pain detection model that can eliminate all these human errors during the pain monitoring period [7].

Chest pain is an evident clinical attribute of myocardial ischemia during the suspected acute phase of myocardial infarction [8]. Both cardiac functional failure and pain fall under myocardial ischemia, caused by the imperative pumping of the heart against hypertensive pressure. It is almost certain that severe angina pectoris can occur in the presence of myocardial hypertrophy or preexistence of coronary artery disease [9]. One study tried to establish a relationship between MI pain duration and the mortality rate, and reports suggested that the highest mortality rate existed among the patients who succumbed to the longest duration of pain [10]. An investigation by researchers found a common pattern in facial expression of patients with chest pain diagnosed with necrosis or cardiac ischemia [11].

Pain is an individual experience and proves to be a complex phenomenon for automatic precise measurement and effective medical diagnosis using facial expressions. In recent decades, researchers have shown keen interest in exploring the challenges and problems of facial expression recognition (FER) in the growing medical allocation area of research such as computer vision and artificial intelligence domains [12]. The existing clinical practices in diagnosing chest pain treatment are time-consuming and involve costly procedures in the treatment process that need to be adapted. A focus has been made by researchers to develop more effective methods for evaluating chest pain caused by myocardial infarction. This idea helped us begin developing an edge-based AI model that interprets the expressions of the face to evaluate the intensity of angina pectoris. Therefore, the patients with high complications with MI who need immediate attention can be called in for an emergency procedure and be admitted to the hospital at the earliest possibility.

The research community has been constantly exploring integral solutions for remote monitoring of patients by generating reports to the clinicians for more than a decade. The primary motivation is to address healthcare issues at all levels: pediatric care, disease monitoring, elderly supervision, emergency patients handling, fitness, and private health management.

In recent years, efficient combinations of cloud computing and effective Internet of Things architectures, along with algorithmic approaches of artificial intelligence (AI), have been exploited to develop a product model for real-time smart health care applications. Data captured by embedded sensors, wearable devices, smartphones, and Internet of Things devices may help to explore the habits and patterns of a person and be effectively utilized in the healthcare domain to solve existing problems through state-of-the-art AI-based approaches [13]. The concept of edge intelligence ("Edge AI") is to "provide AI for every person, anytime, anywhere at all concerns." The Edge IoT devices were developed with built-in AI models to acquire the sensor data and decode its behavior to make accurate decisions and near-precise predictions. IoT devices with Cloud-based architecture disadvantages exhibit qualities such as non-safety, low latency, and soft real-time abilities, which are critical for IoT healthcare applications. However, considering the critical conditions of the patients under time-bound emergency conditions, those criticalities need high robustness, low latency, high bandwidth, and a large degree of reliable systems to avoid fatal consequences. Traditional cloud computing techniques based on IoT devices pose bigger challenges in health monitoring applications, such as bandwidth issues and reliability, latency, and privacy problems. In order to overcome these challenges, the concept of Edge AI has been introduced [14].

Presently, there are mainly three edge computing platforms, namely: i) on-device computation, in which AI computations are done locally on the end device; ii) edge-server architecture, in which the edge server does the computation task after gathering data from the edge devices called nodes; and iii) edge-cloud-based joint computation. Recently, considerable work has been carried out in the healthcare segment on edge platforms. Ghulam Muhammad et al. developed a voice disorder detection and classification system in smart health frameworks [15]. The acquisition of voice signals was carried out through IoT smart sensors. Processing and computing were done through edge computing and a cloud platform. J. Pena Queralta et al. proposed an LSTM Recurrent Neural Network technique for a fall detection system utilizing the state-of-the-art Low Power Wide Area Network (LPWAN). The authors utilized the state-of-the-art low power wide area network technology to overcome the network limitations in Edge AI [16]. Xiangfeng, Dai, et al. presented a mobile health platform for skin cancer detection, developing an inferencing platform on the device by combining deep learning and mobile health technology into a single technique for cancer detection and classification [17].

In this work, we have implemented an automatic facial pain recognition system using an embedded edge GPU platform, Jetson Nano; this was also to evaluate the state-of-the-art CNN lightweight architectures, SSD Inception V2 and Mobile Net SSD V2, by considering the following performance metrics: precision, recall, and frames per second. In this research work, the aim is to classify the severity of pain states of MI expressed as the vital sign exhibited in facial expressions as: normal, mild, and severe levels.

Our major contributions in this research paper are: (i) to create a chest pain facial expression dataset following the benchmark metrics of the Facial Action Coding System (FACS) along with suitable annotation; (ii) to choose a suitable real-time detection model, considering DCNN models SSD- MobileNetV2 and SSD Inception Net V2 for embedded platforms; and (iii) model optimization and performance tests being performed using NVIDIA Jetson NANO, and evaluating the inference speed during the real-time detection model on an embedded platform.

This research work is organized as follows. Section 2 explores the recent research works carried out in automatic pain expression extraction from facial expressions. Section 3 explains the overview of our proposed work and our Facial Action Coding System, and describes our dataset and the performance metrics used for evaluation. Section 4 elaborates and discusses the obtained simulation results. Lastly, Section 5 describes the overall conclusion.

2. Related Work

Scientific analysis of automatic facial expression analysis systems has been around for three decades. Early attempts failed to work for spontaneous facial expression detection and perform under a real-time environment due to a lack of powerful algorithms, capturing of quality datasets, and efficient hardware to process large datasets [18]. The recent revolution in computer vision techniques has made it possible to extract and analyze various health indicators from facial expressions, such as mental state, as well as physiological parameters like respiratory rate, blood pressure, ECG signals, etc. Automatic facial detection has high relevance and has attracted considerable interest in many applications like medical diagnosis, biometrics, forensics, defense, and surveillance. Automatic pain recognition requires a minimum of one sensory input channel, called modality, for extracting relevant information from the patient, and the same data is utilized for further processing in an embedded device or computer. It could be a behavioral feature or physiological feature of the person during observation. Behavioral features are based on body movements (head movements or restlessness), facial expressions, paralingual vocalizations (moaning or crying), or speech. Physiological modalities might be an electro-dermal activity, cardiovascular activity signal (ECG), or brain signal (EEG). A recognition system can be a unimodal or multimodal system [19]. Table 1 provides information related to 4 different metrics that have been incorporated by different authors using machine learning models.

One of the main hurdles in automatic pain detection (APD) research has been the widely accepted scientific dataset until 2009. Later, when a publicly available UNBC-McMaster Shoulder Pain Expression Dataset was released, there was increased interest in this field as well, which noteworthy publications have featured. However, compared to AFER research, APD related works are few, and still in their infancy [19]. Some of the benchmark datasets designed particularly for pain-related research, which have encouraged research in automatic pain detection and classification techniques, are specified in Table 2.

The UNBC-McMaster database comprises 200 videos captured from 25 patients suffering from shoulder pain. The video frames were labeled based on the golden standard for facial expressions: Facial Action Coding System (FACS), which was originally invented by Prkachin and Solomon [26]. Various approaches have been adopted by researchers considering the context of Automatic Pain Detection (APD), ranging from the traditional handcrafted feature extraction techniques of supervised learning to the recent AI algorithmic approaches [25]. A few supervising learning techniques like linear, logistic regression, and decision trees are least often used as learning methods. Other popular supervised machine learning techniques, like Support Vector Regressor (SVR) and Multiple Kernel SVM, are widely used. Semi-supervised learning methods have never been used until now [19].

Apart from AI-based approaches, several traditional approaches have been proposed by researchers. These methods are developed as emergency-based models by combining several facial descriptors such as shape appearance, facial texture, geometry, etc. Yang et al. presented a novel approach based on appearance-based facial descriptors in automatic pain assessment by analyzing the role of spatio-temporal information. Two pieces of descriptor information, namely, spatial texture features and spatio-temporal features, are extracted from video frames and video sequences respectively. The spatial descriptors extracted are mainly those that consist of binarized statistical image features (BSIF), binary patterns (LBP), and local phase quantization (LPQ) analyzed

Tab. 1. Various pain recognition modalities and approaches

Paper	Modality For Facial Expression	Clinical Context	Age	Stimuli	Model	Samples
Adibuzzaman [20]	Smartphone	Breast cancer	35-48	------	SVM-KNN	454+513
Ashouri [21]	Inertial sensor	Lower back pain	20-50	Trunk motion	SVM	52
Haque [22]	RGB, depth, thermal	-	22-42	Electrical	CNN+LSTM	2k
Rivas [23]	Hand movement, finger pressure	Stroke patients	Adult	Rehabilitation exercises	Semi-naïve Bayesian classifier	6K
Yang [24]	Physiological data	ICU	Adult	-------	Boltzmann machine	1K

Tab. 2. Publicly available pain recognition databases

Sl No	Database	Subjects	Stimuli
1	UNBC-McMaster shoulder pain [27]	Shoulder pain patients: 25 adults	200 range of motion tests with affected and unaffected limbs.
2	EmoPain [28]	22 adults with chronic lower back pain aged 50, 28 healthy adults aged ~37	Physical exercises (therapy scenarios)
3	BioVid Heat Pain [29]	90 healthy adults aged between 18-29	14k emotion elicitation heat pain; 41 posed expression cold pressor emotion elicitation
4	IIIT-S ICSD [30]	33 infants aged between 3-24 months	Immunization pain causes; non painful cry causes
5	BP4D Spontaneous [31]	41 healthy adults aged between 18-29	41 cold pressor tasks; emotion elicitation
6	Sense Emotion [32]	45 healthy adults aged around 26	8k heat pain (3 intensities x 30 repetitions x 2 stimulus sites x 45 participants)

from the videos utilizing Three Orthogonal Planes [33]. Juho Kannala et al. advocated an approach for constructing local image descriptors in order to encode textual information which would be suitable for histogram-based representation of image regions. This method generates binary code for each pixel, inspired by local binary pattern and local phase quantization, and provides advancement in overall performance compared to LBP and LPQ techniques [34].

Almost all of the research approaches have focused on facial pain evaluation where an input signal is either images or video samples. Initially, the preprocessing techniques are implemented using normalization or localization techniques for each input frame considered. In the later stage, characteristic feature extraction is carried out based on facial Action Units (AU), or extracting various fiducial points choosing specific facial landmarks. Finally, different algorithmic approaches are utilized to optimize a specific inference model. Lately, especially this decade, the most widely used unsupervised approaches have included Deep Neural Networks such as Recurrent Neural Networks (RNN), Conventional Neural Networks (CNN), and Long Short-Term Memory (LSTM), which have delivered high accurate results. Ghazal Bargshady et al. implemented an improvised deep neural network structure with four threshold levels for facial pain intensity detection. The VGG pre-trained model was adopted for feature extraction technique from the UNBC-McMaster Shoulder Pain Archive Database, and the principal component analysis dimensional reduction method was applied to improve efficiency. A hybrid approach of the deep learning CNN-BiLSTM model was incorporated for pain classification to achieve an accuracy of 90% and AUC of 98.4% [35]. Jing Zhou et al. proposed a novel technique for real-time automatic frame-level pain intensity estimation with an RNN technique by using regression framework. This work demonstrates the sliding window technique to acquire fixed-length input samples for RNN. The regressor approach of RNN provides a continuous score for the pain classification problem instead of discrete labels [36]. Marco Bellantonio et al. highlighted three major factors during the implementation of automatic pain detection: spatial information, temporal axis information, and variation in face resolution during the pain expression variations in video frames. A fusion of deep learning networks (CNN and RNN) was used to extract the features of pain patterns for the UNBC-McMaster Shoulder Pain database, using a super-resolution algorithm to generate the video frames of facial expression through a downsampling process with different resolution setups [37]. Paul et al. explored pain detection techniques to improve the medical diagnosis process with high accuracy and less computing time using the UNBC master shoulder pain database. The mechanism adapted to ensure improved metrics performance like accuracy, AUC, subject exclusive, and nonexclusive settings with a trained deep CNN model for estimating the percentage of pain level using RNN. These results showed a much better metrics performance during the evaluation process for the database as compared to the analysis report of the CK+ facial motion recognition database. The result of the aligned crop LSTM approach shows much better accuracy with an emotion classifier built on top of CNN. The results also highlight the correct classification with and without pain frames, with different facial gestures under various pain conditions for each subject, achieving the classification error [38]. Mohammad Tavakolian et al. adopted the facial pain frames as a compact binary code for intensity level classification by dividing the facial videos in terms of overlapping sequences. Feature extraction of frames was carried out using a CNN algorithm and aggregated as low-level structural information and high-level patterns [39]. Patrick Thiam et al. explored the spatial and temporal features of pain facial expression using attention networks and a mechanism of feeding the sequence of Motion Optical Flow Images (OFIs) and History Images (MHIs) to an attention networks-hybrid CNN and Bidirectional Long Short-Term Memory Recurrent Neural Network (BiLSTM RCNN) for the classification task. Performance analysis was carried out on the BioVid Heat pain database, sensing emotion database points, and achieving an improved performance compared to state-of-the-art methods [40]. Xiaojing Xu et al. established a three-stage DNN approach to evaluate the visual analog scale (VAS) for a video-level measure of pain intensity. The three-stage model includes i) a VGG Face neural network model for predicting the frame-level PSPI, ii) a fully connected neural network for estimating the sequence level pain score, and iii) a linear combination of multidimensional pain estimation of VAS [41]. Frerk Saxen et al. adopted lightweight CNN architectures for automatic face attribute detection systems for deploying the models on smartphones. NasNet-Mobile and Mobile-NetV2 models were used for classifying the custom facial dataset to achieve better accuracy with speed and ease to implement on mobile devices compared to other state-of-the-art methods [42]. From the extensive literature survey, the majority of the current research works on deep learning have adopted the Keras Tensorflow library for facial condition diagnosis. There are alternative powerful libraries in the Python and the C++ programming language, such as Microsoft CNTK, Theano, Caffe, Torch, and Sci-kit Learn, that can be adopted for facial pain analysis.

During recent years, various types of DCNNs have been utilized in object detection architectures, dramatically increasing performance with object detection algorithms. CNN-based object detectors have solved complex real-world problems, e.g., medical imaging, autonomous navigation, video surveillance, and machine vision [43]. Jiaxing Li designed a facial recognition system using the Faster R-CNN object detection algorithm [44]. Facial image feature extraction was carried out using a CNN layer, which was passed to the Region Proposal Networks for generating region proposals, and the classification layer consisting of SoftMax and regression layer. The efficient Faster RCNN network with the Chinese Linguistic Data Consortium (CLDC) dataset's video data for facial expression classification performed to achieve a better mean Average Precision (mAP) of 0.82 [44]. In spite of the significant progress in APD through facial expressions, more efforts should be made to collect an accurate pain database and to improve modelling for its effectiveness in real time clinical practices. From the literature survey, it was found that less attention is paid to chest pain-related facial expressions being used for pain detections. In this research work, we utilize Deep Convolution Neural Network object detection networks SSD Mobile Net V2 and SSD Inception V2 to extract more effective real-time performance on an embedded platform with limited computing resources. CNN lightweight models with high capacity have been adopted in feature selection and feature extraction and also effective transfer learning, thereby implementing automatic pain recognition model using facial expression images.

Recently, some researchers have worked on MI and cardiovascular diseases as medical emergency conditions to automatically detect the early symptoms in humans and prevent mortality. A deep learning-based artificial intelligence algorithm (DLA) has been adopted to detect MI using six-lead electrocardiography. A novel idea for a variational autoencoder was developed using the TensorFlow library for enhancing the performance of DLA. The results highlight that MI can be detected with high accuracy even with a 6-lead ECG device [45]. Mandair et al. utilized logistic regression and DNN algorithmic techniques for predicting MI from the known risk factors. ML packages such as sci-kit-learn and Keras were effectively utilized and implemented on a Google Cloud platform. Compared to the DNN algorithm, the traditional method of logistic regression offered better benefits in evaluating the disease factor from harmonized EHR data [46]. A novel work was advocated by Kwon et al. for estimating the risk strategy for the mortality of patients with acute MI. The authors identified the potential limitations of traditional methods and employed the deep learning-based approach using a multilayer perceptron built through the Tensorflow library. The prediction performance of the deep learning model designed for AMI patient outcomes was excellent [47]. A unique approach of the wearable ECG MI classifier was developed using CNN and recurrent neural networks with only a single lead recording. A stacking decoding method was adopted for the classification scheme of "MI," "healthy," "other," and "noisy" ECG signals to achieve superior performance [48]. Jyoti Metan et al. uniquely adapted an automatic detection technique based on a sandpiper-optimized CNN for detecting cardiovascular disease using cardiovascular magnetic resonance imaging [49].

3. Methodology

3.1. Overview of the Proposed Model

This work implements two high-performing deep learning CNN architectures merged into a single architectural model for an efficient implementation into a computationally intensive embedded platform.

3.1.1. Single Shot Detector (SSD)

The Single Shot Detector (SSD) model is devised to perform localization and classification tasks simultaneously. The SSD architectural framework consists of two stages: a backbone structure and SSD head. The first stage is the backbone structure with a pre-trained CNN network, which acts as an image feature extractor. The backbone structure is pre-trained on a large-scale benchmark dataset like COCO. ImageNet provides a solution to train a rich set of various features. In this work, Mobile Net V2 [50] and Inception V2 models were pre-trained from the COCO dataset for feature extraction or object prediction. The second stage extracts the semantic information from the image without losing the spatial information for classification. Here, the SSD-multi-box approach's core objective is to convert the bounding boxes into a set of default boxes with different aspect ratios and scales [51].

The SSD predicts the objects of different classes even though overlapped bounding boxes exist. During the prediction of objects in an image, the model creates scores in presence of each default box and produces adjustments to the box for better object shape matching. To achieve better detection, the model merges the predictions from multiple feature maps with different resolutions of various sizes of the objects. Figure 1 shows the generated default

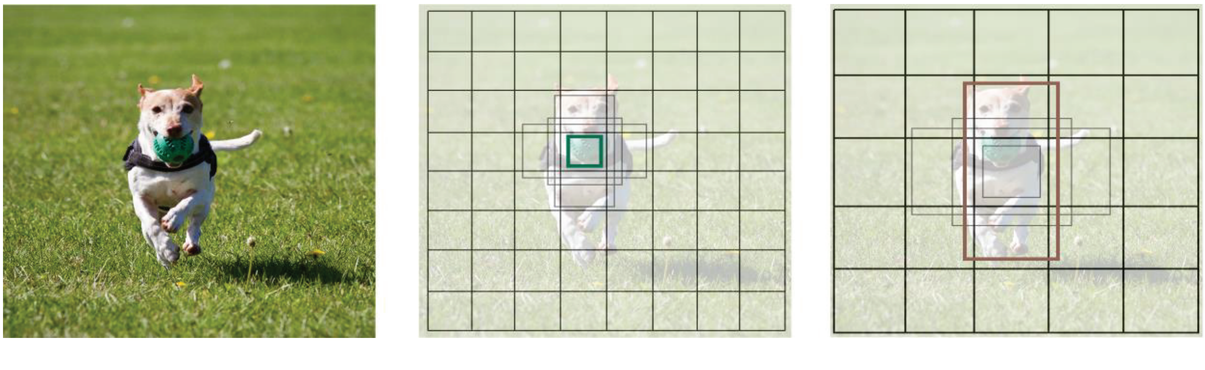


Fig. 1. Generated default boxes from the SSD Model

boxes for various aspect ratios vs. cell sizes. The adaptation of SSD–Mobile-Net V2 architecture, as shown in Figure 2, consists of a base CNN network as Mobile-Net V2 for image feature extraction, an SSD module for bounding box regressions, and a final classification step for accurate facial pain detection.

3.2. The Proposed Method

A block diagram of the advocated architectural framework in the present research is shown in Figure 3. The methodology incorporates three modules to improvise the efficacy of the proposed algorithm. Three stages are i) input stage, ii) training stage, and iii) detection or output stage. In the first stage, RGB original images from the chest pain dataset are transferred to the preprocessing stage wherein the cropping and resizing technique is applied. In the subsequent step, the region of interest of the facial expression is marked and prepared for the next feature extraction and model training phase. During the second stage, a pre-trained CNN network SSD Mobile-Net V2 and SSD Inception-Net V2 are selected for feature extraction and training the custom data. The training process is carried out in the workstation as more powerful hardware is required for training the deep neural network models. Later, the trained model is transferred to the Jetson Nano embedded GPU board, and real-time detection is carried out in the final detection stage for obtaining three distinctive classes of the vital signs of MI as chest pain facial expressions.

3.2.1. Facial Action Coding System (FACS)

The Facial Action Coding System (FACS), designed by Ekman and Friesen in 1976, is the most widely accepted set of standard criteria for facial expression research. FACS was proposed to provide a set of fine-grained, unified criteria for 6 basic emotions: surprise, joy, fear, sadness, disgust, and anger. The Action Units (AUs) defined can be used for all possible human facial anatomical expressions, and gave the researchers a new analytic powerful tool [18]. The investigation report by researchers has revealed that even a pain-related AU can be formulated. The pain evaluation metric PSPI is derived by Prkachin and Solomon [26], and using Equation 1, the metric PSPI is computed from different pain-related AU facial expression intensities.

$$\begin{aligned} \boldsymbol{PSPI} &= \boldsymbol{AU}4 + \max\left(\boldsymbol{AU}6;\boldsymbol{AU}7\right) \\ &\quad + \max\left(\boldsymbol{AU}9;\boldsymbol{AU}10\right) + \boldsymbol{AU}43 \end{aligned} \tag{1}$$

Intensity values of AUs are measured (0-5 from the weakest trace to maximum intensity). With the closing of eyes, AU43 is evaluated for the score values (either 0 or 1).

Here, the researchers adopt either frequency occurrences or pain baseline criteria to differentiate the patients with pain or without pain expression,

i) Frequency of occurrence criterion: The critical frequency level is being marked for a particular AU and, if exceeding the normal range, may be around 5-10%.

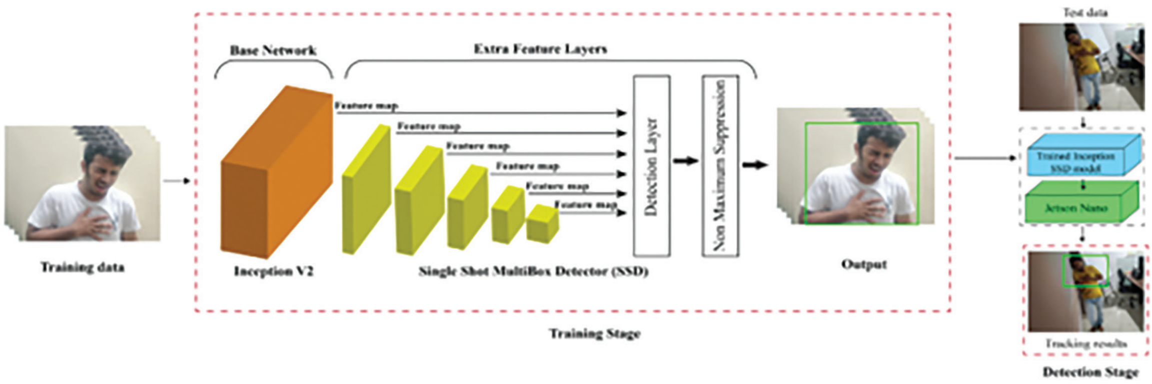


Fig. 2. Proposed SSD Mobile-Net V2 Architectural Model designed for facial expression chest pain detection

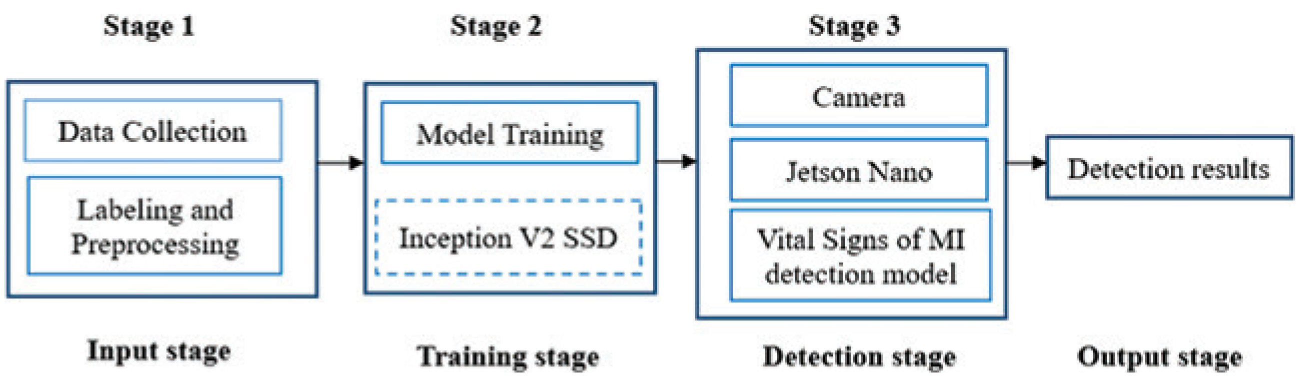


Fig. 3. Method for Training and Detecting Real-Time Vital Signs of Myocardial Infarction

ii) Pain baseline criterion: There is a painful baseline condition being set. The AU is defined as pain-related when it occurs more frequently in pain patients compared to non-pain patients [52].

From this private chest pain facial expression PM dataset, 11 AUs with a more relevant and believable connection to chest pain facial expression [14] are expressed in Table 3. All pain-related studies deviate in selecting the possible AUs for their specific medical application. Only a few AUs are listed here. Figure 4 shows the simulated chest pain facial expression taken from the PM private dataset, considering a few AUs chosen from Table 3 [11, 52].

publicly available datasets. Pain-induced or simulated facial expressions as custom-made datasets are hard to collect. An optimal dataset has to include high-quality annotations, be multimodal, and also have other relevant states to access specificity corresponding to pain against the false alarm rate trade-off [54]. The UNBC-McMaster database is a challenging dataset, where in some cases it is difficult to predict whether a person is in pain or not, even for medical professionals [55]. Facial expression analysis models that are designed for young adults would not also generalize to older age groups [19]. Thus, a dataset of participants aged more than 65 years old has been included for evaluating performance for the age group

Fig. 4. Facial expression related to pain from PM dataset

Tab. 3. Pain Related Action Units

FACS Action Units	Description	Muscular Basis
AU4	Eyebrow lowering	Depressor glabellae, Depressor supercilii, Corrugator supercilli
AU6	Cheek raising	Orbicularis oculi; pars obitalis
AU7	Eyelid tightening	Orbicularis oculi; pars palebralis
AU9	Nose wrinkling	levator labii superioris alaeque nasi
AU10	Upper lip raising	Levator labii superioris; caput infraorbitals
AU20	Lip stretching	Risorius
AU26	Jaw dropping	Masetter; temporal and internal pterygoid relaxed
AU27	Mouth stretching	Pterygoids, digastric
AU43	Eyes closing	Relaxation of levator palpebrae superioris
AU51	Head turning left	Sternocleidomastoid
AU55	Head tilting left	Sternocleidomastoid

3.2.2. Dataset

One of the major challenges researchers face in automatic pain recognition is the availability of suitable

of 16-80 years. Table 4 indicates the classification scheme with different pain score levels for the private PM database.

Tab. 4. Different pain intensity levels in the proposed work database

VAS Score/ PSPI Score	Pain Level	Number of Images
0	No Pain/ Normal	160
0-3	Mild Pain	165
3-5	Severe Pain	175

While acquiring our custom-made chest pain dataset PM, the Action Units mentioned in Table 3 and following points were considered: i) FACS coding pattern for pain fulfilling the critical frequency level, ii) male and female subjects being considered in equal proportion, iii) the participants' age group being from 16–80 years, and v) the type of pain. The images were captured using the OnePlus 5 smartphone camera with a resolution of 16 MP of original frame size 4608x3456. The images are scaled down to lower dimensionality (1067x800 pixels) to minimize computational complexity, in turn enhancing the processing speed. The pain facial expressions simulated are ensured to look like real-time scenarios of a heart attack for the observer. The present work dataset consists of three classes: i) normal pose, ii) mild pain pose, and iii) severe pain pose, as shown in Figure 5.

3.2.3. Hardware Description

With advancements in complex architectures of deep learning networks for performing object detection and classification tasks, high speed parallel computing architectures play a major role. The choice of an edge device and AI algorithm for a specific application are coupled with each other. A careful analysis has to be made while choosing hardware-based architecture models on certain factors, such as cost, energy consumption, accuracy, and throughput. To enrich the optimal performance of computing deep learning models, Nvidia Corporation has developed GPU-enabled parallel processing Cuda core architecture-based embedded boards in recent years. A low-cost, powerful Nvidia Jetson Nano embedded system platform is adopted with the cutting-edge technology of Edge AI in this research work to achieve high accuracy and throughput. Utilizing Jetson Nano's full potential involves an optimization of effective algorithms, as well as hardware, to achieve impressive real-time performance. Figure 6 shows the Jetson Nano board and system interfacing.

3.2.4. Training

The lightweight Deep Convolution Network models in this proposed work are: SSD InceptionNet V2 and SSD MobileNet V2, which were downloaded from the TensorFlow model Zoo, and are pre-trained networks. The pre-trained weights were initialized by training the model using the COCO dataset. Our PM dataset is organized into three main facial expression classes: i) normal, ii) mild pain, and iii) severe pain. The dataset consists of 350 training images and 150 test images. Training a Deep Neural Network model requires high-performance systems with GPUs for effective ad-

Fig. 5. Chest pain facial expression dataset consisting of i) normal, ii) mild, iii) severe expressions

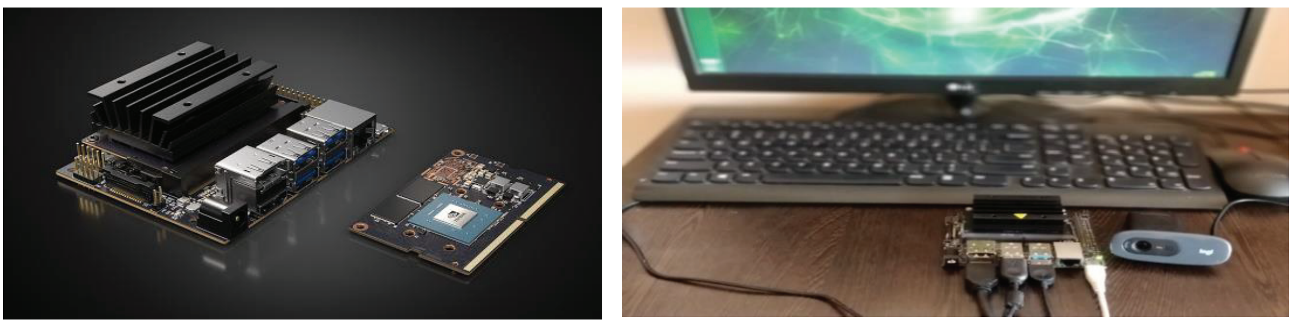


Fig. 6. a) Jetson Nano b) Jetson Nano board configuration

vanced training with a high computation speed. With the help of the class labels mentioned, the goal is to train a DCNN model which predicts chest pain facial expression detection directly from video sequences. Hyper-parameter tuning of the neural network models is based on the evaluation of training/validation learning curves. The number of epochs, training time, and stopping criteria for training the model were decided based on the careful examination of learning curves.

3.2.5. Performance Evaluation Metrics

COCO evaluation object detection metrics have been used in this work to validate the effectiveness of CNN models [56]. The classification problem of chest pain facial expressions as normal, mild, and severe conditions have been evaluated, and SSD Inception/MobileNet model performance has been tested using COCO performance metrics. The mean Average Precision (mAP), Average Recall (AR) and F1 score are used in this evaluation process. The frames per second parameter is also used as a key element in the evaluation process to implement the real-time embedded applications. The bounding box location and class confidence are defined as the predicted outputs.

Intersection of Union (IOU) indicates the scaling factor at which the predicted bounding object box matches with the ground truth box. It brings the relation between the common intersection area over the summation of their areas.

$$IoU = \frac{A \cap B}{A \cup B} \quad (2)$$

For evaluation, the preferred performance metrics in object detection algorithms are: precision (P), Average Precision (AP), mean Average Precision (mAP), Average Recall (AR), and F1 Score. The criteria for performance optimization are designed to find out mispredictions, wrong localization, and any duplications involved during object detection.

Considering the test for the object detection problem, for the i-th image and j-th prediction, the algorithm is expected to find a predicted bounding box b_{ij}. Denoting the confidence value as c_{ij} and threshold of confidence as t, $s_{ij} = 1$, if $c_{ij} = t$; otherwise, $s_{ij} = 0$. The value $z_{ij} = 1$, if the confidence value exceeds the threshold t when the detection prediction j on image i matches a ground truth box; otherwise, $z_{ij} = 0$. Four metrics are defined as follows.

Recall (R_{ot}) is indicated by a proportionality constant of perfect predictions with respect to total number objects in the images for any object classification. Recall R_{ot} given by equation 3.

$$R_{ot} = \frac{\sum_{i=1..N_i} \sum_{j=1\,N_{ij}} z_{ij}}{N} \quad (3)$$

where ot - object threshold value, Ni - number of images, Nij - total number of detections on image i, and N - maximum number of objects in s given class considering total images.

Precision (P_{ot}) is a scaling factor expressed in terms of exact predictions of object over total predictions. Precision P_{ot} is given by equation 4 as

$$P_{ot} = \frac{\sum_{i=1..N_i} \sum_{j=1\,N_{ij}} z_{ij}}{\sum_{i=1..N_i} \sum_{j=1\,N_{ij}} s_{ij}} \quad (4)$$

Here, s_{ij} is set at 1 when the algorithm determines that detection j in image i is an object in the given class, and z_{ij} indicates if it is a correct object under the same class.

Average Precision (AP) for multiple object detection, the precision factor is inversely proportional to recall threshold value. Thus, the metric Average Precision is generally adopted to evaluate by using Equation (5). It is an integral function of precision with respect to the recall over a boundary [0-1] (where r stands for recall). Average Precision (AP_ is given by Equations 5 and 6.

$$AP \quad \int precision(r)dr \quad (5)$$

Under practical considerations, Average Precision is calculated on different recall levels. Let us assume the difference between two close recall levels to be dr. Then Average Precision can be defined as a`verage of precision results over different recall levels.

$$AP = \frac{\sum_{recall=\frac{0,1}{dr}1} precision}{d_r + 1} \quad (6)$$

Mean Average Precision (mAP) the total number of classes of objects in measured images is To. mAP is defined as the mean of APs over total number of classes To. mAP is adopted as the main metric for object detection applications. mAP is defined as the average of AP taking into consideration all classes. mAP is given by equation 7 as

$$mAP = \frac{\sum_{n=1,2..N_0} AP_n}{T_0} \quad (7)$$

The mAP metric evaluates the algorithm's performance over all recall levels and all classes.

Losses (L) the total loss consists of two main losses: localization loss and confidence loss. The localization loss gives an estimate of the mismatch between the final predicted bounding box and the ground truth box. The SSD model mainly adopts the predictions from positive matches, which are closer to ground truth boxes, and the negative matches are ignored. The confidence loss is a value while performing the class prediction. It is a measure of the confidence of a network while estimating the objectness score of the computed bounding box .

Let $x_{ij}^{p} = \{1, 0\}$ be an Indicator for matching the i^{th} default box to the j^{th} ground truth box of category P. In this matching strategy, If $\sum_i x_{ij}^{p} \geq 1$. The overall loss function is a weighted sum of the localization (loc) and the confidence loss (conf):

$$L(x,c,l,g) = \frac{1}{N}\left[L_{conf}(x,c) + \propto L_{loc}(x,l,g)\right] \quad (8)$$

where N - number of matched default boxes.

If N = 0, The loss value is set to 0.

The localization loss is a smooth L_1 loss between the predicted box (l) & the ground truth box (g) parameters. Considering offsets for the center (C_x, C_y) of the default bounding box (d), width (w), and height (h).

$$L_{loc}(x,l,g) = \sum_{i\in Pos}^{N} \sum_{m\in\{C_x,C_y,w,h\}} x_{ij}^{k}\, smooth_{L_1}\left(l_i^m - g_j^m\right) \quad (9)$$

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

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

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

$$L_{conf}(x,c) = -\sum_{i\epsilon Pos}^{N} x_{ij}^{P}\log\left(C_i^p\right) - \sum_{i\epsilon N_{eg}}^{N}\log\left(C_i^0\right) \quad (10)$$

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

where the weight term $\propto$ is set to 1 by cross validation.

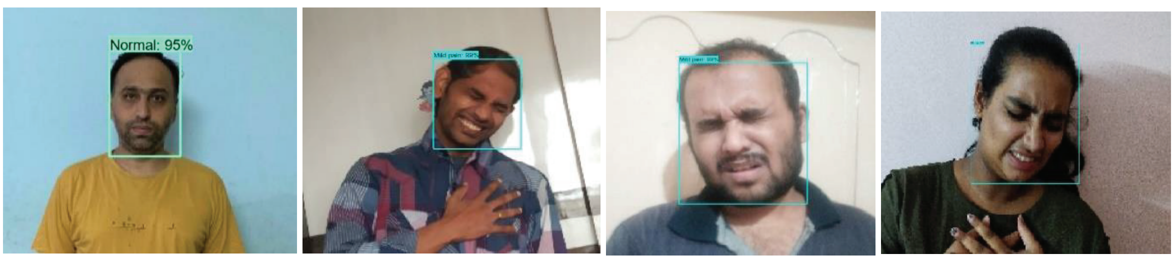
4. Results and Discussion

The self-exploratory modelling experiments were performed to analyze the efficacy of the DCNN proposed models and implemented using Intel(R) core (TM) i7-7700 CPU @3.60GHz and 12GB DDR4 RAM. The current work implements the algorithmic model from Tensorflow object detection API, and the prototype model was developed using the Python programming platform. A powerful deep learning library, Tensor-Flow, and Keras were used for easy and faster prototyping. Figure 7 shows the experimental results of the trained CNN SSD Inception V2 model. The ground truth boxes and the predicted box scores are displayed in the image section of the Tensor board visualization.

Figure 8 depicts three facial expressions captured from the camera. The Tensor board visualization toolkit is designed as a scalar dashboard that is utilized for visualizing the performance metrics. Here, the TensorFlow object detection model is user-friendly; APIs are used for visualizing the evaluated metrics like mean Average Precision, Average Recall, loss function, test images ground truth, and predicted values. Figure 9 shows the mAP graph plotted using the Tensor Board window.

4.1. Precision, Average Recall and F1 Score Evaluation

The parameter metrics are visualized graphically in regular intervals of checkpoints on the Tensor board and the same results are updated. The IOU parameter helps in evaluating the detection to be correct or incorrect with a given threshold. The IOU ratio is equal to 0.5, which indicates the overlapping area of the ground truth box with the bounding. Precision indicates whether the object detection model is identifying relevant objects in a given class, and gives correct positive predictions in terms of percentage. Average Precision elucidates maximum detections per image, considering the predefined standard areas defined in Common Objects in Context metrics, like i) Smaller sized objects <= pixels, ii) Medium-sized objects > to <=, and iii) Larger sized objects > pixels. According to standard COCO metrics, AP and mAP represent the same identity. Figure 9a shows a set of mAP values plotted in the range of IOU values from 0.5 to 0.95, with step size of 0.05. The main classes considered are: normal,



a) b) c) d)

Fig. 7. Predicted simulated results of SSD Inception V2 for 3 classes of chest pain facial expression detection

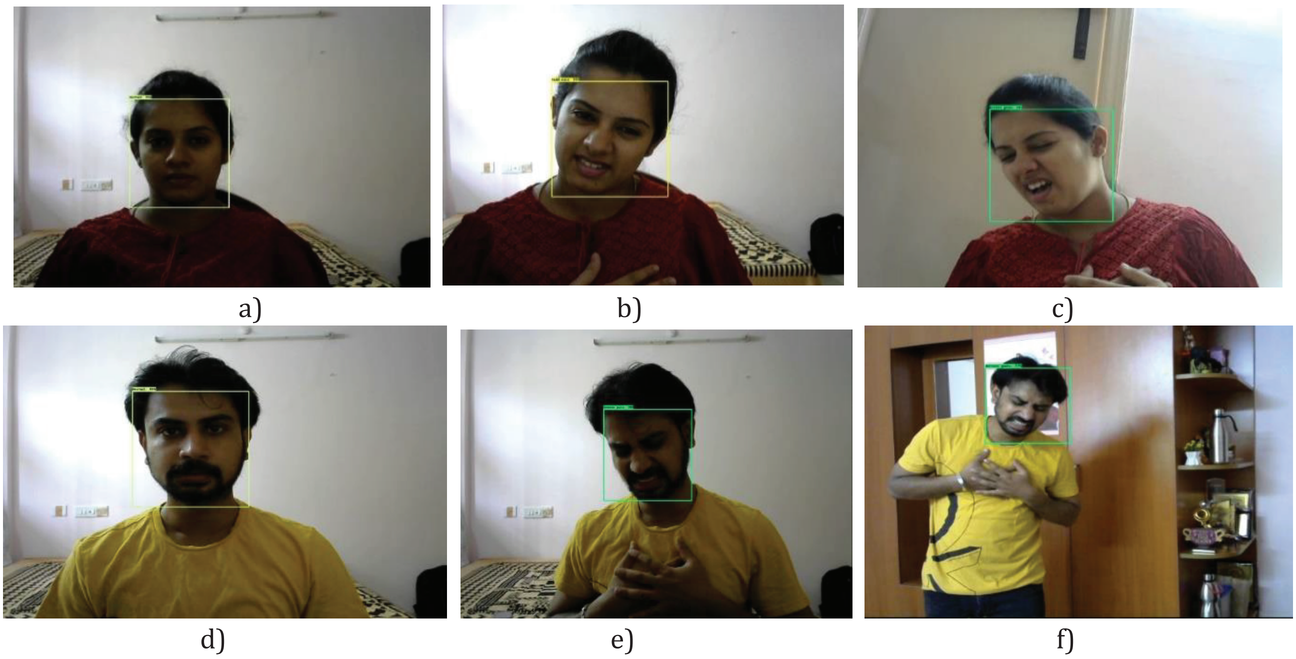


Fig. 8. Real-time detection of facial pain expression

mild pain, and severe pain cases. Average Recall shows the ability of this proposed model to determine the correct positive predictions against all proposed ground truths in a given class and is expressed as a percentage value. Figure 9 (b) shows a set of Average Recall values plotted over categories and IoUs.

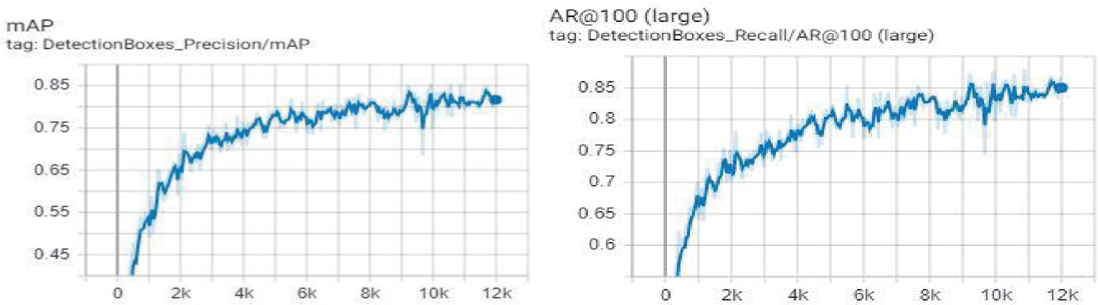
From the results obtained for Precision and Recall metrics, they were used to classify the performance of model detections into 3 types: i) Maximum number of detections – High Precision and High Recall, ii) Maximum detected objects are incorrect indicating maximum false positives – High Recall and Low Precision, and iii) All predicted boxes are correct and maximum faulty ground truth objects indicating false negatives – Low recall and High Precision. The model's aim is to achieve high recall and High Precision Condition as a high-performing system. Table 5 gives the values of mean Average Precision, Average Recall, and F1 Score in terms of percentage.

Tab. 5. Measurement of mean Average Precision, Recall and F1 Score values of SSD MobileNet V2 and SSD Inception V2

Convolutional Neural Networks	Mean Average Precision	Average Recall	F1 Sore
SSD Inception V2 COCO	85.18	88.32	86.72
SSD Mobilenet V2 COCO	82.7	85.5	84.07

4.2. Loss Function Evaluation

The total loss of SSD Inception V2 and SSD InceptionNet V2 models occurring at different stages of the training process is framed into three main losses: classification, regularization, and localization loss. The four possible cases of loss functions were simulated using the Tensor-flow framework whose simulated results in four cases were represented in Figure 10a-c, and the overall loss function arealso represented in



(a) (b)

Fig. 9. Mean Average Precision and Average Recall values of SSD InceptionNet V2 COCO a) mean Average Precision b) Average Recall (large)

Figure 10d. Table 6 gives the different loss values obtained from simulation results for two different CNNs.

Tab. 6. Various loss factors measured by SSD Inception V2

Loss Type	Loss	Loss value
A	Classification loss	1.36578
B	Regularization loss	0.25272
C	Localization loss	0.124921
E	Final total loss	1.7434

4.3. Training Time Comparison

Training time infers total time taken in training the deep learning model. Training for 12,000 numbers of steps took approximately 49 hours for SSD MobileNet V2 COCO. The training time of SSD InceptionNet V2 COCO took lesser time when compared with SSD Mobilenet V2 COCO. Table 7 shows the evaluation results for the training period vs. the number of steps for different DCNN SSD COCO models.

Tab. 7. Measurement of training time v/s number of steps for different SSD models

Convolution Neural Networks	Training Time (Hrs)	Number of Steps (x1000)
SSD Inception V2 COCO	35	12
SSD MobileNet V2 COCO	49	12

4.4. Embedded Implementation

After training the model in a high-performance system, it was deployed to the Jetson Nano GPU board. The accuracy of the proposed chest pain face detection CNN model was estimated by the evaluation of a custom-made chest pain dataset and inference speed on Jetson Nano board, which is tested to verify real-time performance on an EDGE-AI embedded device. The results highlight that the CNN models tested ensure balanced performance in inference speed and also in terms of accuracy in embedded system platforms. Figure 4 shows the experimental setup for inference evaluation of the model. Table 8 shows Jetson Nano's performance considering inference speed measured in terms of frames per second.

Tab. 8. Measurement of Frames per Second for 2 CNN models

Device	CNN Model	Power Consumption (Watts)	Frame per Second (FPS)
Jetson NANO	SSD Mobile-Net V2	10	6.85
		5	3.32
	SSD Inception V2	10	6.26
		5	3.18

4.5. Results Comparison With Other Work

Even though authors of different papers may have used identical databases, comparing the results of different papers is not generally a good approach. The limitations leading to incomparability are due to some of the following differences: i) using subsets of custom-made data, ii) evaluating with different performance measures, iii) evaluation methodologies followed, and iv) prediction tasks. As per the extensive literature survey by the authors, none of the

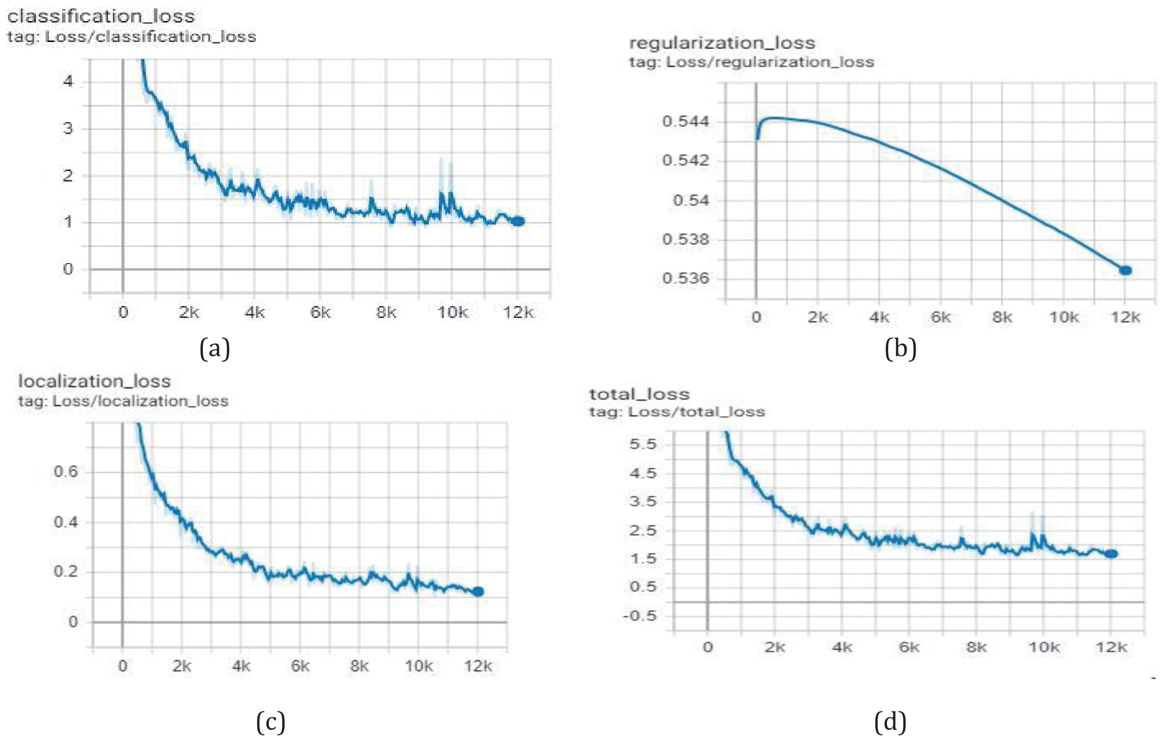


Fig. 10. Loss curves of SSD Inception V2 model

pain facial expression databases have been evaluated using object detection CNN models. However, facial expression recognition has been carried out using the popular object detection Faster RCNN algorithm. Table 9 shows the comparison of performance metrics of our proposed work with the results of Jiaxing Li [44].

Tab. 9. Measurement of mAP metric of this proposed work

Paper	Model	Dataset	mean Average Precision
Jiaxing Li [44]	Faster RCNN with VGG 16 backbone	Chinese Linguistic Data Consortium [CLDC]	81.6
Proposed work	SSD inception V2	Custom chest pain database	85.18

4.6. Discussion

Automatic pain detection is a much-anticipated remedy to the prevailing acute and chronic pain management in the expert medical domain. Computer vision-based analysis offers a promising viable solution for chest pain facial expression for efficient pain detection. In this work, the author has endeavored to evaluate the two state-of-the-art DCNN algorithms, SSD InceptionNetV2 and SSD Mobile Net V2, for chest pain facial expression being considered as a vital sign of heart attacks. It considers the image dataset from custom-made RGB chest pain facial expression images from a high-resolution camera. Three main vital sign postures of facial expression images have been contemplated manually and evaluated in interpreting the severity of the pain as a sign of heart attack detection. In order to evaluate the CNN algorithm, this experiment was carried out to find the best training model incorporating the best test-train configuration ratio for satisfying minimum loss criteria. The process of automatic feature extraction from the training images is an advantage compared to traditional facial expression feature extraction algorithms. The main limitation is a lack of publicly available standard databases for chest pain facial expression, and it was a challenging task to gather the custom-made dataset, annotate it, and build an accurate pain-based facial recognition system for the object detection algorithm.

5. Conclusion

Automating pain detection and estimating the pain intensity level based on facial images via suitable pain management strategies can emerge as a lifesaver in medical health informatics. The artificial intelligence approach adopted in this research plays a significant role for researchers and medical professionals working with pain management practices. The authors have developed a real-time chest pain-based facial expression pain detector that guarantees a pain estimator and detection solution under myocardial infarction emergency conditions to save lives. Two deep conventional neural networks were developed based on algorithmic implementation: SSD InceptionNetV2 and SSD Mobile Net, which have been deployed to evaluate the chest pain facial expression recognition task using CNN networks. The results have been shown to accomplish a state-of-the-art performance using the classification task in TensorFlow object detection-API. Training a CNN model end to end achieved better metrics, with a mean Average Precision of 85.18% and Recall 88.32%. An embedded GPU platform, Jetson Nano, estimates the real-time performance using an object detection algorithm, and the results achieved 6.85 frames per second in the pain detection technique. In the future, this can lead to a road map for researchers by incorporating knowledge-based ideas to develop an embedded system solution that can be designed based on our model as an emergency alarm indicator based on the severity of the pain score during potential life-threatening cardiac arrest situations.

AUTHORS

Dr. Mohan H M* – HMI, Digital Shark Technology Pvt. Ltd, Bangalore, Karnataka, India, e-mail: mohanhm@gmail.com.

Shivaraj Kumara H C – Tecplix Technologies Pvt. Ltd, Bangalore, Karnataka, India, e-mail: hivaraj.k@tecplix.co.

Mallikarjun S H – Government Polytechnic, Kampli, Karnataka, India, e-mail: vmakdree@gmail.com.

Prasad A Y – Dept. of CSE, SJB Institute of Technology, Bangalore, VTU, India, e-mail: prasadadguary@gmail.com.

*Corresponding author

REFERENCES

[1] Cristina Balla Rita Pavasini Roberto Ferrari, "Treatment of Angina: Where Are We?", Cardiology journal, vol.140, 2018, pp.52–67 10.1159/000487936

[2] Amanda Williams, Kenneth D. Craig, "Updating the definition of pain" PAIN, vol.157, no.11, 2016, pp. 2420-2423, 10.1097/j.pain.

[3] George R. Hansen, Jon Streltzer, "Psychology of pain", Emerg Med Clin N, vol. 23, 2005, pp.339–348, 10.1016/j.emc.2004.12.005.

[4] Steven J. Linton, William S. Shaw," Impact of Psychological Factors in the Experience of Pain" Physical Therapy, vol.91, no.5, 2011, pp.700-711, 10.2522 /ptj. 20100330.

[5] James C Eisenach, "Textbook of Pain, 4th edition", vol.12, no.3, July 2000, 276-278.

[6] Xiaojing Xu, Jeannie S Huang, Virginia R De Sa, "Pain Evaluation in Video using Extended Multitask Learning from Multidimensional Measurements", Proceedings of Machine Learning for Health NeurIPS Workshop, PMLR 116: 2020, pp.141-154.

[7] Philipp Werner; Daniel Lopez-M, Walter St , "Automatic Recognition Methods Supporting Pain Assessment: A Survey", IEEE Trans on Affective Computing, vol.13, no.1, Oct 2019, pp. 530-552, 10.1109/ TAFFC.2019.2946774.

[8] Johan Herlitz, Ake Hjalmarson, Finn W, "Treatment of pain in acute myocardial infarction," British Heart Journal, vol.61, 1989, pp. 9-13, 10.1136/hrt.61.1.9.

[9] Richard Gorlin, "Pathophysiology of Cardiac Pain" Circulation, vol.32, July 1965

[10] James H Behrmann, Harold R Hipp, Howard E Heyer, "Pain Patterns in Acute Myocardial Infarction" American Journal of Medicine, vol. 9, no.2, Aug 1950, pp.156-163, 10.1016 0002-9343(50)90018-0.

[11] J A Dalton L Brown, J Carlson, R McNutt, S M Greer "An evaluation of facial expression displayed by patients with chest pain", vol. 28, no.3, May-Jun 1999, pp.168-74, 10.1016/s0147-9563(99)70056-7

[12] Patrick Thiam, Hans A Kesler, "Two-Stream Attention Network for Pain Recognition from Video Sequences", Sensors, vol. 20, no.3, pp.839 2020, 10.3390/s20030839

[13] Luca Greco, Gennalo Percannella, Pierluigi Ritrovato, "Trends in iot based solutions for health care moving ai to the edge", Pattern Recognition Letters, vol.135, July 2020, pp. 346-353, 10.1016/j.patrec.2020.05.016.

[14] Massimo Merenda, Carlo Porcaro, Demetrio Lero, "Edge machine learning for ai-enabled iot devices a review", Sensors, Vol. 20, no.9, 2020, pp.25-33, 10.3390/s20092533.

[15] Ghulam Muhammed, Mohammed F Alhamid, "Edge computing with cloud for voice disorder assessment and treatment" IEEE Communications Magazine, vol. 56, no. 4, April 2018, pp. 60-65 10.1109/MCOM.2018.1700790.

[16] J. P. Queralta, T. N. Gia, H. Tenhunen and T. Westerlund, "Edge-AI in LoRa-based Health Monitoring: Fall Detection System with Fog Computing and LSTM Recurrent Neural Networks," 2019 42nd International Conference on Telecommunications and Signal Processing, Budapest, Hungary, 2019, pp. 601-604, doi: 10.1109/TSP.2019.8768883.

[17] X. Dai, I. Spasić, B. Meyer, S. Chapman and F. Andres, "Machine Learning on Mobile: An On-device Inference App for Skin Cancer Detection," 2019 Fourth International Conference on Fog and Mobile Edge Computing (FMEC), Rome, Italy, 2019, pp. 301-305, doi: 10.1109/FMEC.2019.8795362.

[18] Dianbo Liu, Dan Cheng, Timothy T Houle, Lucy Chen, Wei Zhang, Hao Deng, "Machine learning methods for automatic pain assessment using facial expression information", vol.97, no. 49, Medicine 2018, 10.1097/MD.0000000000013421.

[19] Teena Hassan, Dominik Seuß, Johannes Wollenberg, Katharina Weitz, Miriam Kunz, Stefan Lautenbacher, Jens-Uwe Garbas, Ute Schmid, "Automatic Detection of Pain from Facial Expressions: A Survey" IEEE Transactions on pattern analysis and machine intelligence, 2019, 10.1109/TPAMI.2019.2958341

[20] M Adibuzzaman; Colin Ostberg; S Ahamed, Richard P; "Assessment of Pain Using Facial Pictures Taken with a Smartphone", IEEE 39th Annual computer Software and Applications conference, vol.2, July 2015, pp. 726-731. 10.1109/OMPSAC.2015.150.

[21] Sajad Ashouri, Mohsen Abedi, Masoud Abdollahi, "A novel approach to spinal 3-D kinematic assessment using inertial sensors: Towards effective quantitative evaluation of low back pain in clinical settings", Comput Biol Med, vol.89, Aug 2017, pp.144-149, 10.1016/j.compbiomed.2017.08.002.

[22] M. A. Haque et al., "Deep Multimodal Pain Recognition: A Database and Comparison of Spatio-Temporal Visual Modalities," 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), Xi'an, China, 2018, pp. 250-257, doi: 10.1109/FG.2018.00044.

[23] J. J. Rivas et al., "Automatic recognition of pain, anxiety, engagement and tiredness for virtual rehabilitation from stroke: A marginalization approach," 2017 Seventh International Conference on Affective Computing and Intelligent

Interaction Workshops and Demos (ACIIW), San Antonio, TX, USA, 2017, pp. 159-164, doi: 10.1109/ACIIW.2017.8272607.

[24] Lei Yang,Shuang Wang, Xiaoqian Jiang, "PATTERN: Pain Assessment for patients who can't tell using Restricted Boltzmann machine," BMC Medical Informatics and Decision Making, vol.73, July 2016, pp.190-208, 10.1186/s12911-016-0317-0.

[25] Zhanli Chen, Rashid Ansari, Diana J Wilkie, "Automated Pain Detection from Facial Expressions using FACS: A Review", Nov 2018, pp. 1-19.

[26] Kenneth M Prkachin 1, Patricia E Solomon, "The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain", Pain, vol.139, Oct 2008, pp. 267-74, 10.1016/j.pain.2008.04.010.

[27] Patrick Lucey, Jeffrey F. Cohn, Kenneth M. Prkachin "Painful data: The UNBC-McMaster shoulder pain expression archive database", 2011 IEEE International Conference on Automatic Face & Gesture Recognition, March 2011, pp. 1-9, 10.1109/FG.2011.5771462.

[28] Min S. H. Aung, Sebastian Kaltwang, Bernardino Romera-Paredes, "The Automatic Detection of Chronic Pain-Related Expression: Requirements, Challenges and the Multimodal Emo Pain Dataset", 2016 IEEE Transactions on Affective Computing, vol.7, no. 4, Oct 2016, pp.435-451, 10.1109/ TAFFC.2015.246283.

[29] S. Walter, S Gruss, H Ehleiter, "The biovid heat pain database data for the advancement and systematic validation of an automated pain recognition system," 2013 IEEE International Conference on Cybernetics (CYBCO), Lausanne, Switzerland, 2013, pp. 128-131, doi: 10.1109/CYBConf.2013.6617456.

[30] Mittal V K., "Discriminating the Infant Cry Sounds Due to Pain vs. Discomfort Towards Assisted Clinical Diagnosis", SLPAT 2016- on Speech and Language Processing for Assistive Technologies, 2016, pp.37-42, 10.21437/SLPAT. 2016-7.

[31] Xing Zhang, Lijun Y, "BP4D-Spontaneous: a high-resolution spontaneous 3D dynamic facial expression database", Image and Vision Computing, vol.32, no. 10, Oct 2014, pp.692-706, 10.1016/j.imavis.2014.06.002.

[32] Maria Velana, Sascha G, G Layher, "The Sense Emotion Database: A Multimodal Database for the Development and Systematic Validation of an Automatic Pain- and Emotion-Recognition System", Multimodal Pattern Recognition of Social Signals in Human-Computer-Interaction, vol. 10183, June 2017, pp. 127-139.

[33] Ruijing Yang, Shujun Tong, Miguel Bordallo "On Pain Assessment from Facial Videos Using Spatio-Temporal Local Descriptors", IEEE 6th International Conference on Image Processing Theory, Tools and Applications, IEEE 2016, pp.1-6. 10.1109/IPTA.2016.7820930.

[34] Juho Kannala, Esa Rahtu, "BSIF: Binarized Statistical Image Features", Proceedings of 21st International Conference on Pattern Recognition, ICPR2012, Nov 2012, pp.1363 – 1366.

[35] Ghazal Bargshady, Xujuan Zhou, Ravinesh "Enhanced deep learning algorithm development to detect pain intensity from facial expression images", Expert systems with applications, vol.149, 1 July 2020,

[36] Jing Zhou; Xiaopeng Hong "Recurrent Convolutional Neural Network Regression for Continuous Pain Intensity Estimation in Video" 2016 IEEE Conf on Computer Vision and Pattern Recognition, 10.1109/CVPRW.2016.191.

[37] Marco Bellantonio, Mohammad A. Haque, Pau Rodriguez, "Spatio-temporal Pain Recognition in CNN-Based Super-Resolved Facial Images" FFER 2016: Video Analytics. Face and Facial Expression Recognition and Audience Measurement, March 2017, pp.151-162.

[38] Pau Rodriguez, Guillem C, Jordi Gonalez "Deep Pain: Exploiting long short Term Memory Networks for facial expression classification" IEEE Transactions on Cybernetics, Feb 2017, pp.1-11, 10.1109/tcyb.2017.2662199.

[39] M. Tavakolian and A. Hadid, "Deep Binary Representation of Facial Expressions: A Novel Framework for Automatic Pain Intensity Recognition," 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 2018, pp. 1952-1956, doi: 10.1109/ICIP.2018.8451681.

[40] Patrick Thiam, Hans A. Kestler, Friedhelm Schwenker "Two-Stream Attention Network for Pain Recognition from Video Sequences" Sensors, vol. 20, no.3, 2020, 10.3390/s20030839.

[41] Xiaojing Xu, Jeannie S. Huang, Virginia R. de Sa, "Pain Evaluation in Video using Extended Multitask Learning from Multidimensional Measurements", Proceedings of Machine Learning Research, vol.116, 2020, pp.141–154.

[42] F. Saxen, P. Werner, S. Handrich, E. Othman, L. Dinges and A. Al-Hamadi, "Face Attribute Detection with MobileNetV2 and NasNet-

Mobile," 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA), Dubrovnik, Croatia, 2019, pp. 176-180, doi: 10.1109/ISPA.2019.8868585.

[43] Min-Kook Choi, Jaehyung Park, Heechul Jung "Fast and Accurate Convolutional Object Detectors for Real-time Embedded Platforms" Computer Vision and Pattern Recognition, 2019, arXiv:1909.10798.

[44] Jiaxing Li, Dexiang Zhang, Jinging Zhang "Facial Expression Recognition with Faster R-CNN", Procedia Computer Science, vol.107, 2017, pp.135-140, 10.1016/j.procs. 2017.03.069.

[45] Cho, Y. et al. (2020) 'Artificial intelligence algorithm for detecting myocardial infarction using six-lead electrocardiography', Scientific Reports 2020 10:1. Nature Publishing Group, vol.10, no.1, pp.1–10. 10.1038/s41598-020-77599-6.

[46] Mandair, D. 'Prediction of incident myocardial infarction using machine learning applied to harmonized electronic health record data', BMC Medical Informatics and Decision Making. BioMed Central, vol.20, no.1, 2020, pp.1–10. 10.1186/S12911-020-01268-X.

[47] Kwon, J. 'Deep-learning-based risk stratification for mortality of patients with acute myocardial infarction', PLOS ONE. Public Library of Science, vol.14, no.10, 10.1371/JOURNAL. PONE.0224502

[48] Lui, H. W. and Chow, K. L. (2018) 'Multiclass classification of myocardial infarction with convolutional and recurrent neural networks for portable ECG devices', Informatics in Medicine Unlocked. Elsevier, vol.13, pp.26–33. 10.1016/ J.IMU.2018.08.002

[49] Jyoti Metan, A.Y. Prasad, K.S. Ananda Kumar et al. "Cardiovascular MRI image analysis by using the bio inspired (sand piper optimized) fully deep convolutional network (Bio-FDCN) architecture for an automated detection of cardiac disorders" Biomedical Signal Processing and Control, vol.70, 2021, 10.1016/j.bspc.2021.103002

[50] Howard, Andrew Zhu, Menglong Chen, Bo Kalenichenko, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications", 2017, pp.1-9. arXiv preprint ar Xiv:1704.04861.

[51] Wei Liu, Dragomir Anguelov, Dumitru Erhan, "SSD: Single Shot MultiBox Detector" Computer Vision and Pattern Recognition, 2016, 10.1007/978-3-319-46448-02.

[52] M. Kunz, D. Meixner, S. Lautenbacher, "Facial muscle movements encoding pain–a systematic review," Pain, vol.160, no. 3, March 2019, pp.535–549.

[53] Wilkie, Diana J. "Facial Expressions of Pain in Lung Cancer" Analgesia, vol.1, no.2, 1995, pp.91-99, 10.3727/ 107156995819564301.

[54] P. Werner, D. Lopez-Martinez, S. Walter, A. Al-Hamadi, S. Gruss and R. W. Picard, "Automatic Recognition Methods Supporting Pain Assessment: A Survey," in IEEE Transactions on Affective Computing, vol. 13, no. 1, pp. 530-552, 1 Jan.-March 2022, doi: 10.1109/TAFFC.2019.2946774.

[55] Pau Rodriguez; Guillem Cucurull; Jordi Gonzàlez P. "Deep Pain: Exploiting Long Short-Term Memory Networks for Facial Expression Classification," in IEEE Transactions on Cybernetics, vol. 52, no. 5, pp. 3314-3324, May 2022, doi: 10.1109/TCYB.2017.2662199.

[56] Lin, Zitnick, Doll, "Microsoft COCO: Common Objects in Context", Computer Vision, ECCV, vol. 8693, 2014, pp. 740-755.