Mathematical Biosciences and Engineering

MBE, 18(6): 8727–8757.
DOI: 10.3934/mbe.2021431
Received: 05 July 2021
Accepted: 09 September 2021
Published: 13 October 2021

http://www.aimspress.com/journal/MBE

Research article

Applied convolutional neural network framework for tagging healthcare systems in crowd protest environment

Gaurav Tripathi[1], **Kuldeep Singh**[2] **and Dinesh Kumar Vishwakarma**[3,*]

[1] Department of ECE, Delhi Technological University, Delhi 110042, India
[2] Department of Electronics & Communication Engineering, MNIT, Jaipur
[3] Department of IT, Delhi Technological University, Delhi 110042, India

* **Correspondence:** Email: gauravtripathy@gmail.com.

Abstract: Healthcare systems constitute a significant portion of smart cities infrastructure. The aim of smart healthcare is two folds. The internal healthcare system has a sole focus on monitoring vital parameters of patients. The external systems provide proactive health care measures by the surveillance mechanism. This system utilizes the surveillance mechanism giving impetus to healthcare tagging requirements on the general public. The work exclusively deals with the mass gatherings and crowded places scenarios. Crowd gatherings and public places management is a vital challenge in any smart city environment. Protests and dissent are commonly observed crowd behavior. This behavior has the inherent capacity to transform into violent behavior. The paper explores a novel and deep learning-based method to provide an Internet of Things (IoT) environment-based decision support system for tagging healthcare systems for the people who are injured in crowd protests and violence. The proposed system is intelligent enough to classify protests into normal, medium and severe protest categories. The level of the protests is directly tagged to the nearest healthcare systems and generates the need for specialist healthcare professionals. The proposed system is an optimized solution for the people who are either participating in protests or stranded in such a protest environment. The proposed solution allows complete tagging of specialist healthcare professionals for all types of emergency response in specialized crowd gatherings. Experimental results are encouraging and have shown the proposed system has a fairly promising accuracy of more than eight one percent in classifying protest attributes and more than ninety percent accuracy for differentiating protests and violent actions. The numerical results are motivating enough for and it can be extended beyond proof of the concept into real time external surveillance and healthcare tagging.

Keywords: healthcare; internet of things; deep learning; convolutional neural network; protest;

violence detection

1. Introduction

Internet of things (IoT) is an interconnection of networks where multiple devices are connected through Internet [1]. The devices can range from camera, computers, sensors etc through which information is ingested and processed in the ubiquitous IoT framework. The information can be accessed anywhere, anytime. Wireless Sensor Networks (WSN) [2] propels IoT system framework which has revolutionized the manner of data transmission. Within a WSN the job of the sensors is to sense the respective environment and send it to the appropriate devices such as local host computer or any central cloud server. Modern IoT systems have various smart systems such as smart dustbin [3], smart irrigation system [4], smart healthcare system [5], and smart traffic control [6] to name a few. The focus of this proposed research is on the smart healthcare systems. The focus in the smart healthcare system is usually on the reactive state of healthcare systems. This reactive state deals after an event have happened and the patient has been admitted in the hospital where his vital life signs are monitored, stored and analysed. Smart healthcare is one of the serious challenges that needs to be addressed especially for elderly patients. Smart wearable devices for elderly patients [7] can be included in intelligent framework for smart cities. The devices record the sensor data and transmit it to a secured cloud-based platform. Smart healthcare has multiple attributes such as remote patient monitoring, wearable devices based smart computing and preparing response for emergency services.

Intelligent transport systems in smart cities works assists emergency situations [6].Intelligent transport improves the chances of saving lives in case of accidents, natural disasters and man-made situations such as violent protests. Smart cities must have provisions for dealing emergency situations**.** Proactive emergency response preparation is the key to face crowd protests and violence**.** Smart healthcare systems tag heterogeneous fields and different departments to ensure safety of patients, especially elderly patient as well as health care institutions. Smart heterogeneous applications like smart homes [8] have proven that the convergence of AI and IoT technologies in the best optimized ways during video streaming through wireless micro medical devices (WMMDs) in smart healthcare homes. Constant research in improving surveillance framework for managing crowd protests must be developed. This framework will assist in the deduction of severity of protests and can alert the healthcare systems for the availability of healthcare professionals and specialist.

1.1. Psychology of protest

The psychology of protest has been studied and the researchers have approached the idea of protests and demonstrations using social psychological approach. Van Stekelenburg et al. [9] presented motives for protests and demonstration by humans. The paper presented five unique domains which accelerates the initiation and processing of crowd protest behavior. These five things are as follows:

a) Grievances: Public anger about situations drives the initiations of protests. Joining of any popular political personality can drive agitations and protests [10].

b) Efficacy: It is an individual’s person belief that can change their present conditions, policies through protest and demonstrations.

c) Identity: Identification with group, belief or any organization can give the impetus to join the protests.

d) Emotions: Anger is the main emotion that gives the energy to humans to go for protests.

e) Social Embeddedness: Sharing grievances of society as a whole with other fellow members can initiate the drive for protests.

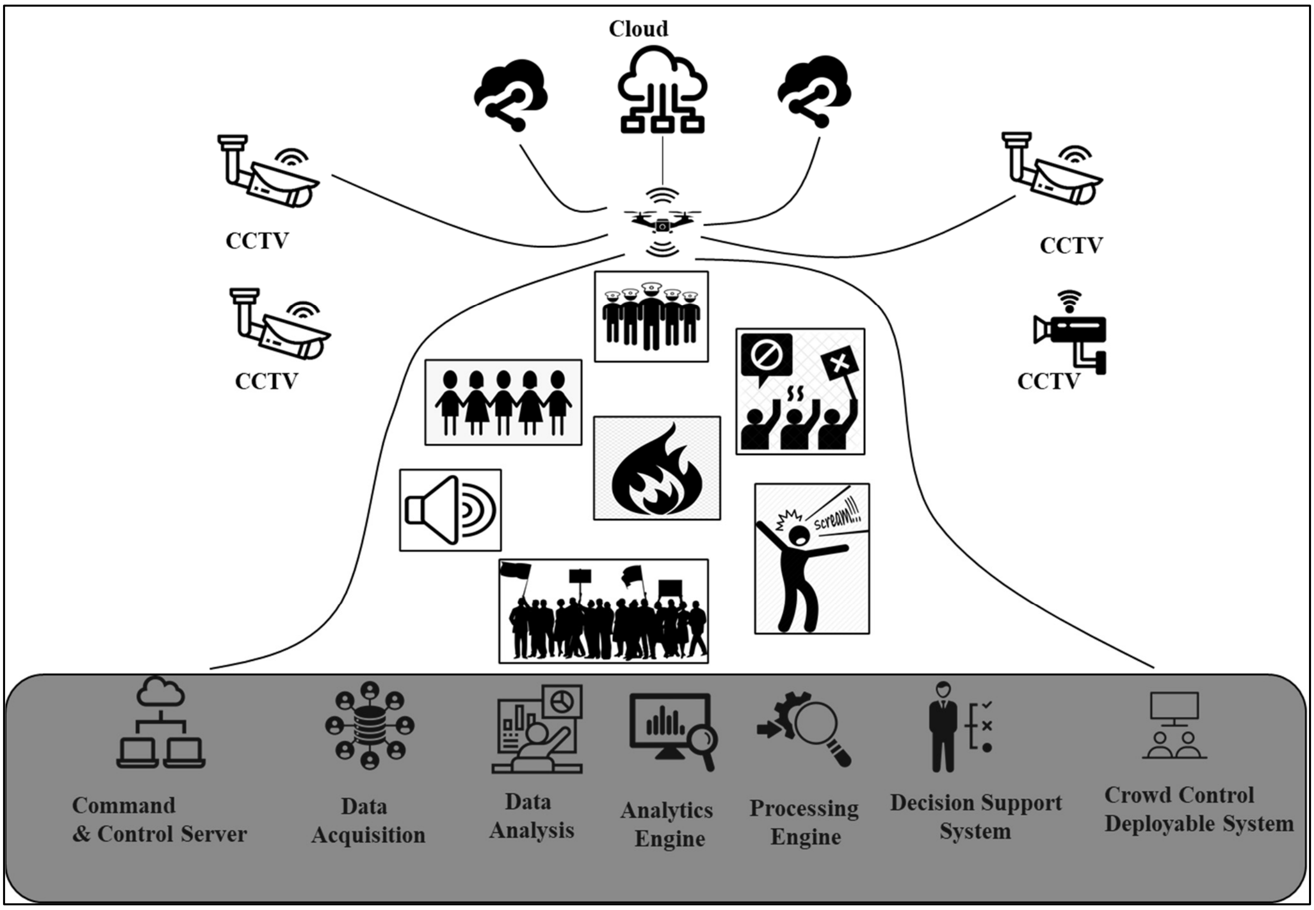

Figure 1. Overview for management of crowd protests using AI and IoT.

As per Figure 1, an overview of the envisioned IoT environment-based framework is presented. The framework is designed for a crowd protest environment. It can be clearly observed that crowd protests activities must be monitored through CCTV and drone cameras in a more sophisticated environment. This is a centrally distributed approach. The feeds are transferred to the central cloud device where the crowd protests feeds are deciphered and classified into multiple stages based on the severity of the protests. Deep learning is used for training the crowd protests models and then further classification is done on the live feeds of the protesting crowd. Based on the severity of crowd protests appropriate tagging is initiated for available healthcare systems and professionals. IoT devices such as CCTV sensors and cloud computing servers would be needed to integrate the seamless transmission of information across the city. The aim of the IoT service is to make sure that primary healthcare infrastructure is secured and upgraded and people dissenting, protesting and indulging in violent activities are attended to in any emergency situation. Thus, an effective response is computed and the specialist healthcare infrastructures and professionals are tagged in such unfortunate event.

1.2. Psychology of violence

The psychology of violence believes that there is a difference that exists in the domain of biology, psychology, sociology, culture, and mass communication which leads to the idea of violence. A general theory breaking down violence is by describing their respective attributes. External as well as internal environment of individual is responsible for motivation to carry out violence. DeKeseredy et al. [11] came up with strong theory on psychology of violence in twelve partitions which include exchange theory, subcultural theory, resource theory, patriarchal theory, ecological theory, social learning theory, evolutionary theory, sociobiological theory, pathological conflict theory, psychopathological theory, general systems theory, and inequality theory. The same book cites reasons such as feelings of alienation, shame, humiliation by individual or group of people, mortification, rejection by individual or society as a whole, abandonment, denial, mental health ailments leading to depression, anger due to societal injustice, hostility within communities, projection, and displacement of individual or a group of people can cause extreme protests and violence.

1.3. Subsystems of healthcare

Healthcare systems consists of internal as well as external healthcare subsystems. Internal framework systems focus on the individual patient level interactions admitted to the hospital. External subsystem analyzes the external environment to decipher the need of healthcare systems, professionals, specialist doctors, emergency services in a real time situation. These subsystems facilitate the tagging of healthcare systems to the needy people. Figure 2 highlights the subsystems of healthcare framework.

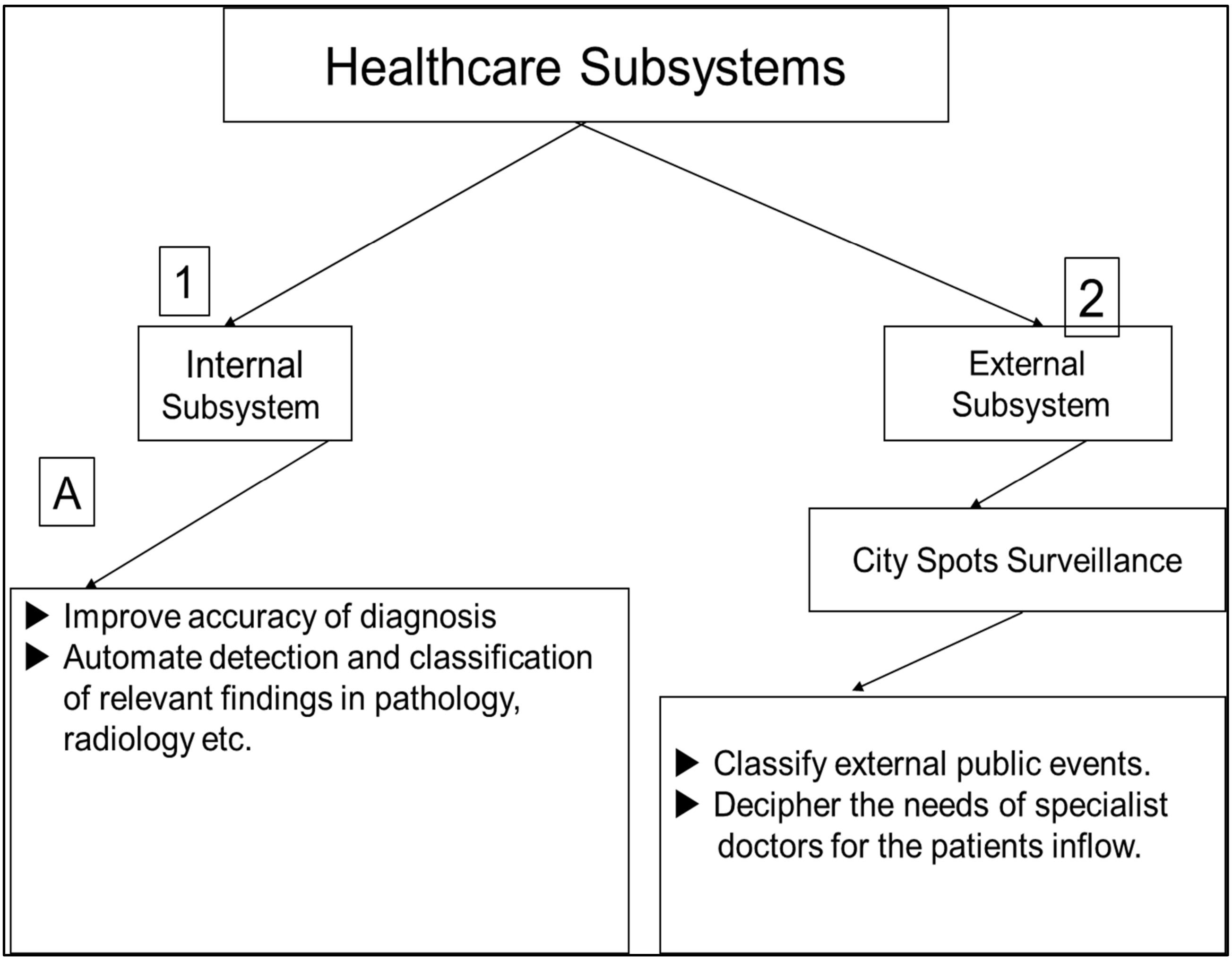

Figure 2. Healthcare system: internal and external subsystems.

The focus of this research is to study external crowd subsystems where crowded places and their events are observed in a camera based IoT environment. The present research focuses on external subsystem where concentration is on protests sites and violent crowds' sites. Scenarios concerning these two events are included in the current study for tagging healthcare systems.

Figure 3. Protestors sample images.

Crowd protests consists of demonstration, remonstration or remonstrance and it is an expression of disagreement or objection towards an idea or action, typically a political one [12,13]. Protests have been related to cause stir in the mentality of the nation [14,15]. Social movements [16] can trigger protests in countries [17]. Protests can spiral in to violence if the demands of protesters are not met. Aggravated people can resort to violent methods to show their disapproval. Violence classification assumes prime significance in securing human life and public places and properties [18]. Figure 3 gives us the fair idea about how protests is organized using group of people, banners, children, banners, shouting and so on. Protesting crowds have high energy [19] which can translate itself in to violence laced situation. Interesting cues regarding images and sounds have been proposed in last decade with aim to accelerate the research [20,21]. This research paved the path for more research in computer vision domain.

Convolutional neural networks (CNNs) are a new framework for deep learning solutions. The paper shall refer Convolutional neural networks (CNNs) as ConvNets. ConvNets have been applied for different computer vision applications [22–25]. A sample ConvNets is shown in the figure below.

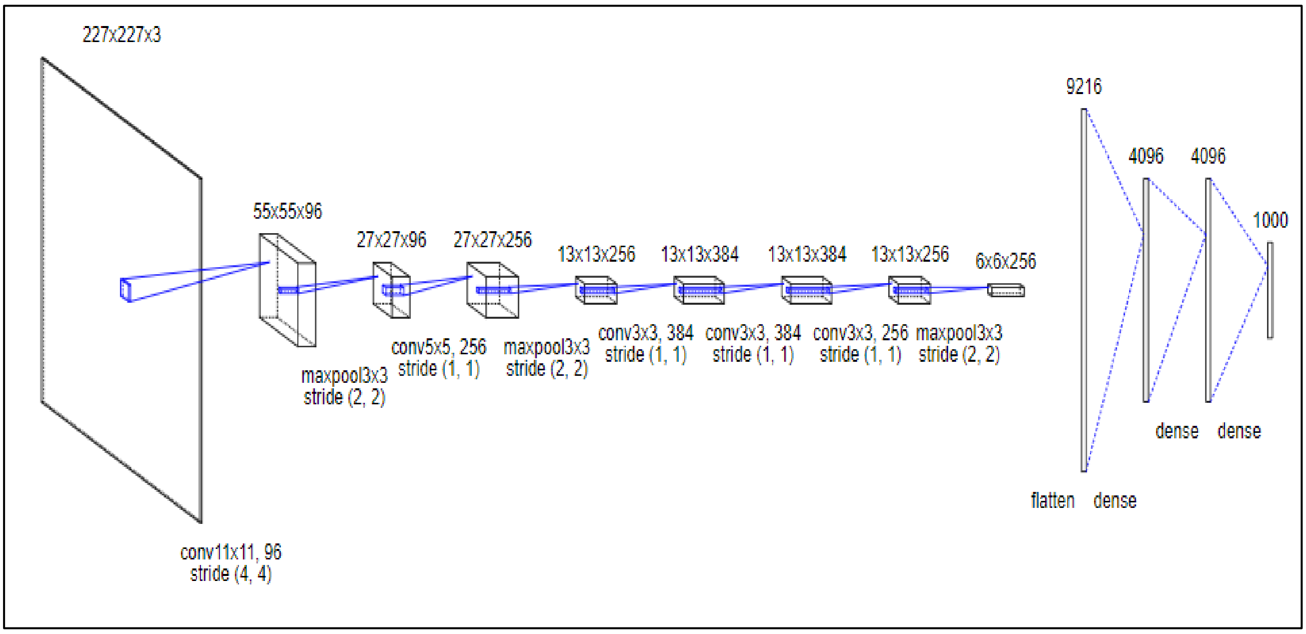

Figure 4. A Standard ConvNet.

The ConvNets have the ability to retain significant image features which it uses to learn the respective classification. The paper presents a novel application of using pre-trained ConvNets to classify stages protests and violence. Popular pretrained models are VGG16 [26], Inception ResNet V2 [27]. Each of these architectures have performed exceptionally well in ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [28]. In the event of moderate protests to extreme protests, it is the responsibility of police and law enforcement agencies to control the protests. The response from the police team for countering protests and violence is use of tear gas, pepper spray, batons, shields, and rubber bullets. The consequences of using these measures presents grave dangers to the protestors and rioters. In the midst of the protests and counter protest processes injuries are observed to the crowds. Most common injuries in case of moderate to extreme protests are ocular trauma and subsequent vision loss [29]. Thus, the paper introduces the framework of crowd behavior situations concerning protests and violence, use of deep learning to classify the situations and popular injuries that happens to the participating crowd in such situation. The contribution of the paper is as follows:

- Utilization of an existing protests datasets [30].
- Curation and utilization of violence dataset from YouTube.
- Explain the rule base to categorize the protests.
- Fuzzification of rules for degree of protests seen in the images/videos.
- Based on the degree of protests tagging of hospitals and healthcare professionals is generated.
- The idea can be spread to various other crowd domains where individual or mass of people gets injured. The whole scenario is envisaged for deployment in an IoT based smart city framework.

Our paper follows the following structure of the paper. Section 2 discusses the brief related work of the latest protest and violence detection methods. Section 3 discusses the proposed methodology describing the pre-processing philosophy of the input and current proposed ConvNets frameworks. Section 4 describes the experimental setups along with the datasets, Section 5 describes the protests detection performance, Section 6 comes up with the discussion of the paper and Section 7 deals with the conclusion of the paper.

2. Related work

A novel multi-task convolutional neural network was developed for recognizing protesters. This model deciphered the protestors activities using their visual attributes and estimate the perceived violence in the image [30]. Our literature survey specifically focusses on deep learning-based methods for classifying protests or protests like events. Computer vision can be used for deciphering political events related questions using large-scale visual content. A study has been performed for classification of facial features for deciphering their biases towards political movements and leaders using twitter users' photographs [31]. Another study has been performed on the images of political leaders that has been shared on social and multimedia platforms [32] for their perceived emotions and personality traits. Image classification domain depends upon visual features in the input media content. These features can help classify the content in to creative content, interesting content and adult provocative themes [33]. Social event detection has been a constant source of research using computer vision and a deep learning approach has been applied using Graph Convolutional Network [34]. Computer vision has also been used for detecting extreme stages of protests such as arson and stone pelting [35]. A ConvNets and recurrent neural network approach known as collective action from social media (CASM) image data and recurrent neural networks with long short-term memory on text data was used in a two-stage classifier to identify offline collective action from social media data [36]. A novel research study was dedicated for extracting protests news by using Long Short-Term Memory and combined with a deep learning model developed on the inter country datasets [37].Our study also includes related work on violence detection methods which can be a direct result from the protest event. Novel method of Bidirectional Convolutional LSTM has been given by applied by Hanson et.al [38]. A multistream ConvNets architecture using 3D CNN and ConvNets was proposed in [23] and conducted on the CCTV-Fights dataset. A flow mechanism was deployed to detect violent event in crowds using deep learning [39] and is entirely based on the local level conditions of Bangladesh. Song et al. [40] used application of 3D CNN for violence detection approach. 3D CNN was especially used for convolutional operations on extracting the spatio-temporal features of the frames. Xu et al. proposed P3D-LSTM violent video classification [41]. Li et al. [42] synthesized a novel method by using 3D CNN utilizing multiplayer violence detection procedure. Another excellent method using 3D CNN method was proposed by Tran et al. using 3D CNN. The 3D ConvNets were designed to learn the procedure of learning the spatiotemporal features of the input videos and then classifying the videos using machine learning procedures [43]. Zhou et al. applies 3D-CNN for getting deep image features for violence classification [44]. Mandal et al. experimented with discriminative features using fine-tuned deep convolutional neural residual network for classifying crowd behavior [45]. Ammar et al. worked extracted video frames and applied CNN and LSTM layers for classification [46]. Meng et al. experimented with trajectory-based motion features and deep CNN for violence detection [47].

Table 1. Deep learning techniques for Protest and violence detection.

Authors	Methodologies	References
Perez et al.	3D ConvNets	[23]
Won et al.	Visual features for perceived violence.	[30]
Wang et al.	Facial features classification	[31]
Chen et al.	Political leaders' images linkage to protest sentiments.	[32]
Ganguly et al.	Provocative theme features extraction	[33]
Peng et al.	Graph Convolutional Network	[34]
Tripathi et al.	ConvNets used for Arson and stone pelting	[35]
Zhang et al.	ConvNets and recurrent neural network	[36]
Thenmozhi et al.	LSTM based feature extraction	[37]
Hanson et al.	Bidirectional Convolutional LSTM	[38]
Sumon et al.	Flow mechanism for violence detection	[39]
Song et al.	3D ConvNets	[40]
Xu et al.	P3D-LSTM	[41]
Li et al.	Multiplayer violence detection using 3D ConvNets	[42]
Tran et al.	3D ConvNets	[43]
Zhou et al.	3D- ConvNets	[44]
Mandal et al.	Fine-tuned ConvNets for crowd behavior	[45]
Ammar et al.	LSTM based ConvNets	[46]
Meng et al.	Normal ConvNets	[47]

Solving protest environment using ConvNets is a challenging task. A simpler and yet a novel approach is to use pretrained ConvNets subjected to fine tuning which can classify the protests and can then the situation can be classified in to their respective fuzzy categories. The authors believe that the crowd formation itself presents an opportunity to tag healthcare system although then the threat level is very less and tagging is not necessary in the normal crowd situation. The approach presented in the current paper is novel, simple and easy to tag the healthcare systems.

3. Proposed methodology

We approach the problem of deciphering two domains and the convergence of one domain in to one another. Crowd protests can transform themselves in to crowd violence and in some case crowd violence reverts back to crowd protests. As long as the protests are under control, peaceful, local law and order authorities have no problems with the demonstrators. This research article tries to decipher the various stages of convergence of protests in to violence using pretrained ConvNets. The idea of this research is to assess the protesting site for any need of the specialist healthcare workers/doctors. Escalation of protest stage to violent stage can have serious ramifications. To prove a point pretrained ConvNets are utilized in the current research. We utilize the cluster of pretrained ConvNets architecture and combined them for deducing the protests-violence classification using VGG16 [26] and Inception ResNet V2 [27]. It is a deeper network and pre-trained on more than a million images on the famous ImageNet database [48]. These ConvNets have been chosen as they have shown better performance in tackling the problem of vanishing gradient [48].

Table 2. Subtle differences in existing methodology and proposed methodology.

Authors	Methodologies	Proposed Methodology
Perez et al. [23]	3D ConvNets	3D ConvNets are known to acquire higher computational resources. Pretrained models can be easily trained and used effectively. It can also be used for mobile based devices.
Won et al. [30]	Visual features for perceived violence.	Features are extracted by pretrained model itself.
Wang et al. [31]	Facial features classification	The present approach focusses on the whole image and not just the facial features classification.
Chen et al. [32]	Political leaders' images linkage to protest sentiments.	General public images and their actions are classified for clearer protests stage deciphering.
Ganguly et al. [33]	Provocative theme features extraction	Highly dependent on the 12 features that is being classified. Has the ability to add more attributes and analyze protests more comprehensively.
Peng et al. [34]	Graph Convolutional Network	Pretrained Convnets
Tripathi et al. [35]	ConvNets used for Arson and stone pelting	Covers more protests attributes rather than just two scenarios.
Zhang et al. [36]	ConvNets and recurrent neural network	Pretrained Convnets
Henmozhi et al. [37]	LSTM based feature extraction	Pretrained Convnets
Hanson et al. [38]	Bidirectional Convolutional LSTM	Pretrained Convnets can be combined with Bidirectional Convolutional LSTM in future protests scenarios
Sumon et al. [39]	Flow mechanism for violence detection	Pretrained Convnets
Song et al. [40]	3D ConvNets	Pretrained Convnets
Xu et al. [41]	P3D-LSTM	Pretrained Convnets
Li et al. [42]	Multiplayer violence detection using 3D ConvNets	Pretrained Convnets for crowd behavior analysis.
Tran et al. [43]	3D ConvNets	Pretrained Convnets for crowd behavior analysis.
Zhou et al. [44]	3D- ConvNets	Pretrained Convnets for crowd behavior analysis.
Mandal et al [45]	Fine-tuned ConvNets for crowd behavior	Efficient pretrained models are used for protests as well as violence classification
Ammar et al. [46]	LSTM based ConvNets	Pretrained Convnets for crowd behavior analysis.
Meng et al. [47]	Normal ConvNets	Pretrained Convnets for crowd behavior analysis.

The previous section has highlighted the methods that is being applied for the protests and violence detection methods. Latest literature has few methods that directly deals with crowd protests. Also, violence detection using ConvNets are also fewer. The methods that are being touched here utilizes the pretrained Convnets. The main objective of choosing this method is to make sure that this proof of concept transforms itself in to real time program which can be used for Android and web-based devices equally. Table 2 presents a simpler difference approach between existing methodologies and proposed methodology.

Let us assume that X to be the of the protest dataset $|X|$ The error $\in$ is calculated as:

$$\in (W, X) = \frac{1}{|\mathrm{D}|} \sum_{i=1}^{|\mathrm{D}|} \mathcal{L}\left(\mathcal{F}(X_i, w), \widehat{c_i}\right) \tag{2}$$

where X_i is the i^{th} image of training dataset X, $\mathcal{F}(X_i, w)$ is the ConvNet function that predicts the class c_i of X_i given the respective weights w,and $\widehat{c_i}$ is the ground-truth class of the i^{th} image, and $\mathcal{L}(c_i, \widehat{c_i})$ is a penalty function for predicting c_i instead of $\widehat{c_i}$. We can set $\mathcal{L}$ to the logistic loss function [49]. In the proposed method we Adaptive Gradient Descent (AGD/AdaGrad) [50].The cost function to compute the optimization error is as follows:

$$\mathcal{L} = -\frac{1}{|X|} \sum_{i}^{|\mathrm{D}|} \ln\left(p(c_i|X_i)\right) \tag{3}$$

$p(c_i|X_i)$ denotes the probability by which X_i is correctly classified in the protest dataset. If W_h^t denotes the weights in h^{th} hidden convolutional layer at iteration t, and $\hat{\mathcal{L}}$ denotes the cost over a mini-batch of size M, then the updated weights in the next iteration are computed as below:

$$\eta^t = \eta^{\left\lceil \frac{tM}{|\mathrm{X}|} \right\rceil} \tag{4}$$

$$U_h^{t+1} = \rho U_h^t - \eta^t \alpha_h \frac{\partial \hat{\mathcal{L}}}{\partial W_h} \tag{5}$$

where α_h is the learning rate for the h^{th} layer, ρ is the momentum that signifies the contribution of the last weight update in current iteration and η is the scheduling rate that decreases α at the end of each epoch. The Adaptive Gradient Descent was computed for optimizations vide the following equations with respect to the mini batches formed similar to that used in SGD.

$$U_h^{t+1} = \rho U_h^t - \frac{\eta}{\sqrt{K_{h,h}^t + \beta}} . \alpha_h \frac{\partial \hat{\mathcal{L}}}{\partial W_h} \tag{6}$$

$$W_h^{t+1} = W_h^t + U_h^{t+1} \tag{7}$$

where $K_{h,h}^t \in R^{d\times d}$ represents a matrix such that the diagonal elements (h, h) equals the sum of gradients.

Following the above calculation, the weights in the fine-tuning architecture were updated for optimization as indicated by Eq (7). We can clearly infer from Eqs (3)–(7), the back-propagation activity was carried out iteratively to fine tune the pretrained model. Following two pretrained ConvNets are used in deciphering the multiclass protests and then used for computing severity of the protests.

3.1. VGG16 ConvNets

VGG16 is one of the popular pretrained ConvNets and had a good classification accuracy in ImageNet large scale visual recognition challenge (ILSVRC) [28]. This is simple to use and is popular in the transfer learning domain. In the process of fine tuning VGG16, the fully connected top layers are removed and new fully-connected layers are added for our classification of tasks. The model utilizes a set of weights pre-trained on ImageNet. VGG16 uses gradient descent combined with a loss function over a sufficiently large dataset propels the ConvNets to learn the patterns in the complex visual space data that is fed to the network. VGG16 consists of 16 convolutional layers. These layers help the architecture in feature extraction. VGG16 has three main constituents: convolution, pooling, and fully connected layers. VGG16 proceeds with two convolution layers. Then a pooling layer is added. The pattern is repeated again. Then three convolutions are repeated with pooling attached. Finally, three FC (Fully Connected) layers are added.

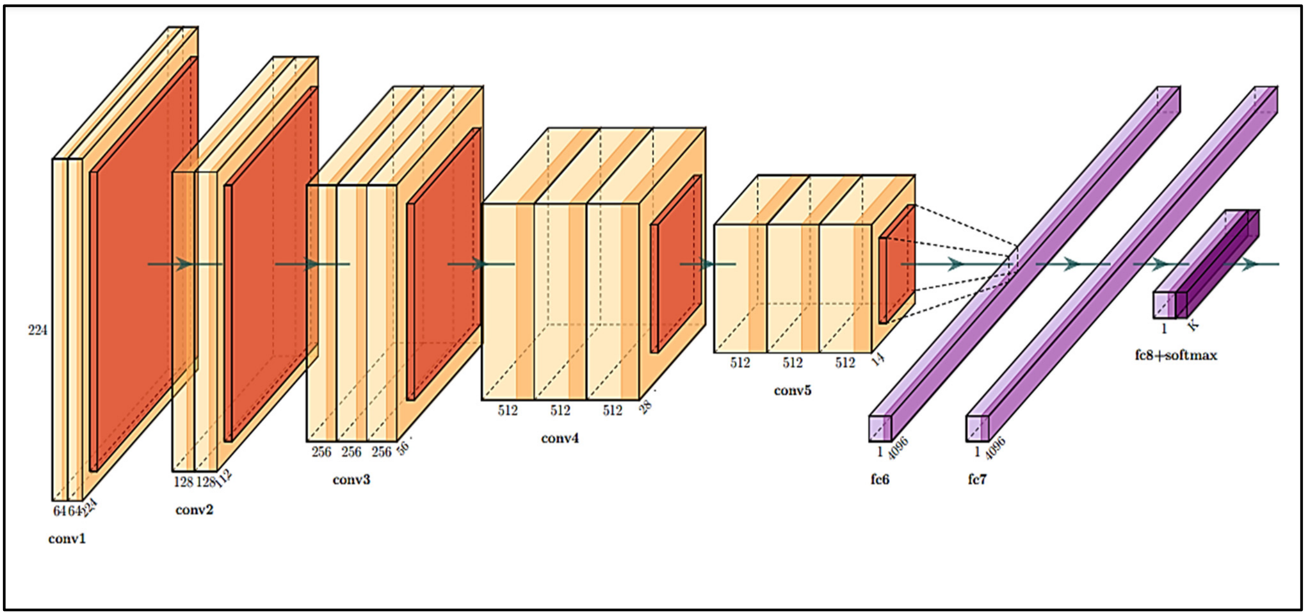

Figure 5. VGG16 Architecture.

Figure 5 presents the VGG16 architecture which is used in the current research to highlights the working of ConvNets and is generated using [51].The above architecture has single global max pooling layers, single fully connected layers with 512 units, single drop out layer and single softmax activation layers of violence enabled inputs. The layers in VGG16 are frozen and feature extraction is performed for the additional FC layer.

3.2. Inception ResNet ConvNets

With the introduction of residual connections with normal inception blocks has yielded good performance in 2015 ILSVRC challenge and the new architecture is known as Inception ResNet V2 [27]. These residual connections facilitate the shortcuts in the model to train deeper neural networks, Hence the performance has increased for inception blocks.

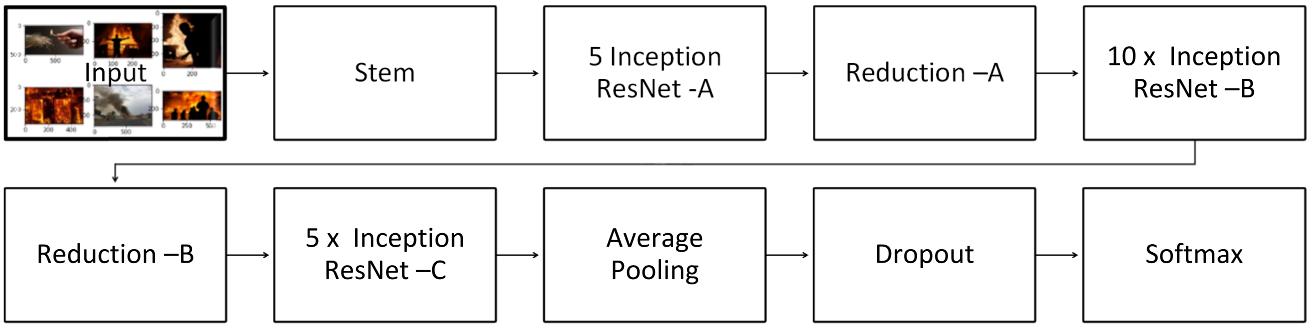

Figure 6. Inception-ResNet architecture.

Figure 6 shows the steps of the input images to the final classification of the protests. This model reduces the representational bottleneck by introducing factorization of large block in to two smaller blocks. Also, it expanded the filter banks horizontally rather than vertically. This ensured less loss of information. The above steps were incorporated to build three different types of inception modules known as Figure 6. Instead of training the model from scratch the paper relies upon the transfer learning concept and does fine tuning of the Inception-ResNet V2 on the protest dataset.

4. Experimental approach

We have evaluated our approach on standard imbalanced protests datasets [30] and have tried to classify various classes of protests dataset using transfer learning. To further prove the research analysis, we have curated a violence related dataset mostly from YouTube and other related videos. Protest dataset [30] have been used extensively for classification using three pretrained ConvNets. The number of images in this dataset is 40,000 images. We have added an additional category of loudspeakers for protest dataset and removed violence category from this dataset. Instead, we have curated our own images and added to the violence category. The approach is twofold. One is to analyse protests from the angle of Table 2 and find out the severity of the protest events. The second is to classify protests and violence. We are focussing on protests dataset and their combinations among them for the fuzzification of the severity of protest. We design a rule base for identifying the stages of protests and its severity. Table 3 presents the rule base in a tabular form:

Table 3. Rule base for severity of protests.

Situation Scenario	children	fire	flags	20 people	100 people	Night	Photo	Police	Protest	Shouting	Sign	Loud speakers	Fuzzy Protest
1.	–	√	–	–	–	√	–	√	√	–	–	–	Medium
2.	√	–	–	–	–	√	–	√	√	√	–	–	Severe
3.	–	√	–	–	√	√	–	–	–	√	–	√	Severe
4.	–	√	–	–	–	√	–	–	–	√	–	–	Severe
5.	–	√	√	–	√	–	√	–	√	√	–	√	Severe
6.	–	√	√	–	√	–	–	√	–	√	–	√	Severe
7.	–	√	–	–	–	–	–	√	–	–	–	√	Medium
8.	–	√	√	–	√	–	–	–	√	√	√	√	Severe
9.	–	√	–	–	√	–	–	–	–	√	–	√	Severe
10.	–	√	–	–	√	–	–	√	–	√	√	√	Severe
11.	–	√	–	–	√	–	–	√	–	–	√	√	Medium
12.	–	√	–	–	√	–	–	√	–	√	√	√	Severe
13.	–	–	√	–	√	–	–	–	√	–	√	√	Normal
14.	√	√	√	–	√	–	√	–	√	–	√	√	Normal
15.	√	√	√	–	–	–	–	–	–	–	–	–	Normal

The above table gives a fair idea about how to interpret the table. There are twelve attributes which decide the situation. Based on the presence of few or more attributes beyond a threshold, the protest situation is classified in to different normal, medium and extreme level protests. For example, if there is a presence of night, children, fire, shooting and police presence the protests situation is in severe category and appropriate solution is suggested. One thing we noted that combinations of fire, night, police, children, shouting and loudspeakers and group of more than 100 denoted severe protests in to combining in to violence. These combinations have inherently more chances of transformations of protests in to violence. As soon as these combinations are detected an alert notification is pushed to the all the available health care centres of the city. Based on the severity categories an alert for specialist healthcare personnel is generated. Based on the presence of these twelve attributes of any crowd protest event, healthcare systems are tagged and a ready alert is generated. A report by Amnesty international [52] discussed about the police actions on the protestors. The report discussed about the actions taken by police in tackling unarmed protestors and use of weapons to neutralize them. Serious injuries amounting to loss of lives have been reported. As stated in the literature, the standard police response to control protests is to use multitude of weapons, such as tear gas, pepper spray, rubber bullets, batons, riot shields, and TASERS [53]. Similarly violent protestors display their potential by using rocks, sticks, knives, guns, and Molotov cocktails [54]. Taser devices are used by law enforcement agencies to control violent protests and riots. These devices are known to reduce the risk of injury and death by minimizing the use of lethal force. Spectrum of injuries reported by usages of taser [55] is shown in the Table 4:

Table 4. Spectrum of injuries reported by usages of taser [55].

S.No	Type of injuries reported by usage of Tasers	Types of healthcare professional needed
1.	A basilar skull fracture, right subarachnoid haemorrhage, and left-sided epidural haemorrhage necessitating craniotomy	✓ Neurosurgeon Surgeon [56] ✓ Maxillofacial Surgeon [57] ✓ Plastic Surgeon [58]
2.	A concussion, facial laceration, comminuted nasal fracture, and orbital floor fracture	
3.	Penetration of the outer table and cortex of the cranium by a Taser probe with seizure-like activity reported by the officer when the Taser was activated	
4.	A forehead hematoma and laceration	

Hence locating these surgeons nearby to the protests and violence can help save many lives and prevent major injury losses. Below table is a research article that focusses on the types of injuries that is faced by the crowd protestors.

Table 5. Type of injuries possible during protests [59].

Protests Response	Ocular Injuries	Chest Injuries	Fractures	Skin Injuries	Bullet Injuries
Tear gas	√	–	–	√	–
Pepper spray	√	–	–	√	–
Rubber bullets	√	√	–	–	√
Batons	√	√	√	–	–
Riot shields	√	–	–	–	–
TASERS	√	√	√	√	√

Table 5 describes the types of serious injuries that happens in severe protests and violence. Table 6 describes the types of doctors needed for the kind of injuries reported in Table 4.

Table 6. Protest's injuries and their respective specialist.

Typical Injuries	Type of specialist needed
Ocular Injuries	Ophthalmic Surgeon [60]
Chest Injuries	Orthopaedic Surgeons [61]
Fractures	Orthopaedic Surgeons [61]
Skin Injuries	Plastic Surgeon [58]
Bullet Injuries	Orthopaedic Surgeons [61]

The above tables clearly describe the types of police response used, types of injuries inflicted and types of specialists needed in case of severe protests. Thus, a proper relationship between protests, response, injuries and specialist healthcare professionals has been established in this research. Now based on the severity of protests, general physicians as well as specialists' doctors would always be available in the hospitals during protests. For the implementation of proposed protests classification and their respective typical injuries detection in crowded scene algorithm, we used standard Keras application with Python on Ubuntu operating system and experiments were conducted on a standard machine (Intel Core i5,8 GB RAM) equipped with NVIDIA TITAN X GP.

5. Protest detection performance

We deployed the above-defined architecture. The results are divided in to two approaches:

- The first approach uses pretrained ConvNets for multiclass classification of protests datasets. We use VGG16 and Inception Resnet V2 for the protest's classification.
- A rule based is defined in the previous section. Based on this rule base an alert is generated which can be sent to the nearest healthcare centers demanding the readiness of specialist's doctors as described in the previous sections. Sending the alerts can be automatic or through a human in the loop [62]. This research believes that the human is necessary to broadcast the protest level to all the healthcare centers in the vicinity of the protest's sites. The whole research is dedicated to the protest's behavior and its respective severity.

Total dataset is revised by adding loudspeakers to the protests datasets and violence category is separated from the original dataset. Then the utility is invoked for displaying of dataset images. The image is as follows:

Figure 7. Protest images in the actual dataset.

Then the experiment proceeded with VGG16 and Inception Resent V2 pretrained ConvNets. The experiment utilizes k-fold cross Validation technique. Cross-validation is a unique procedure used for evaluation of deep learning models on the limited data sample. As specified in the literature [63] ,the choice of k is normally taken to be 5 or 10. Arbitrary increase in the value of k cannot guarantee better results as resampling subsets gets smaller. The experimental procedure is as follows:

1. Protest dataset is shuffled randomly.
2. Split the dataset into k = 5 groups
3. For each such group:
    - ✓ Take the group as a hold out or test data set
    - ✓ Take the remaining groups as a training data set
    - ✓ Fit a model on the training set and evaluate it on the test set
    - ✓ Retain the evaluation score and discard the model
4. Model evaluation scores and respective confusion matrix is described thereafter.

Evaluation metrics: The performance of the proposed research concept was tested and on the parameters of accuracy, specificity, sensitivity, precision, and f1-score. The parameters used in the following equations are true-positive N_{TP}, number of true negative N_{TN}, number of false positive N_{FP}, and number of false negatives N_{FN} to measure predicted classes which are based on the evaluation metric mentioned in Eqs (8–12):

$$\text{Accuracy}=\frac{\mathrm{N_{TP}}+\mathrm{N_{TN}}}{\mathrm{N_{TP}}+\mathrm{N_{TN}}+\mathrm{N_{FP}}+\mathrm{N_{FN}}} \tag{8}$$

$$\text{Specificity}=\frac{\mathrm{N_{TN}}}{\mathrm{N_{TN}}+\mathrm{N_{FP}}} \tag{9}$$

$$\text{Sensitivity}=\frac{\mathrm{N_{TP}}}{\mathrm{N_{TP}}+\mathrm{N_{FN}}} \tag{10}$$

$$\text{Precision}=\frac{\mathrm{N_{TP}}}{\mathrm{N_{TP}}+\mathrm{N_{FP}}} \tag{11}$$

$$\text{F1- Score}=2\times\frac{\mathrm{recall}\times precision}{precision+recall} \tag{12}$$

VGG16 confusion matrix and classification results are as follows. The respective confusion matrix is as follows:

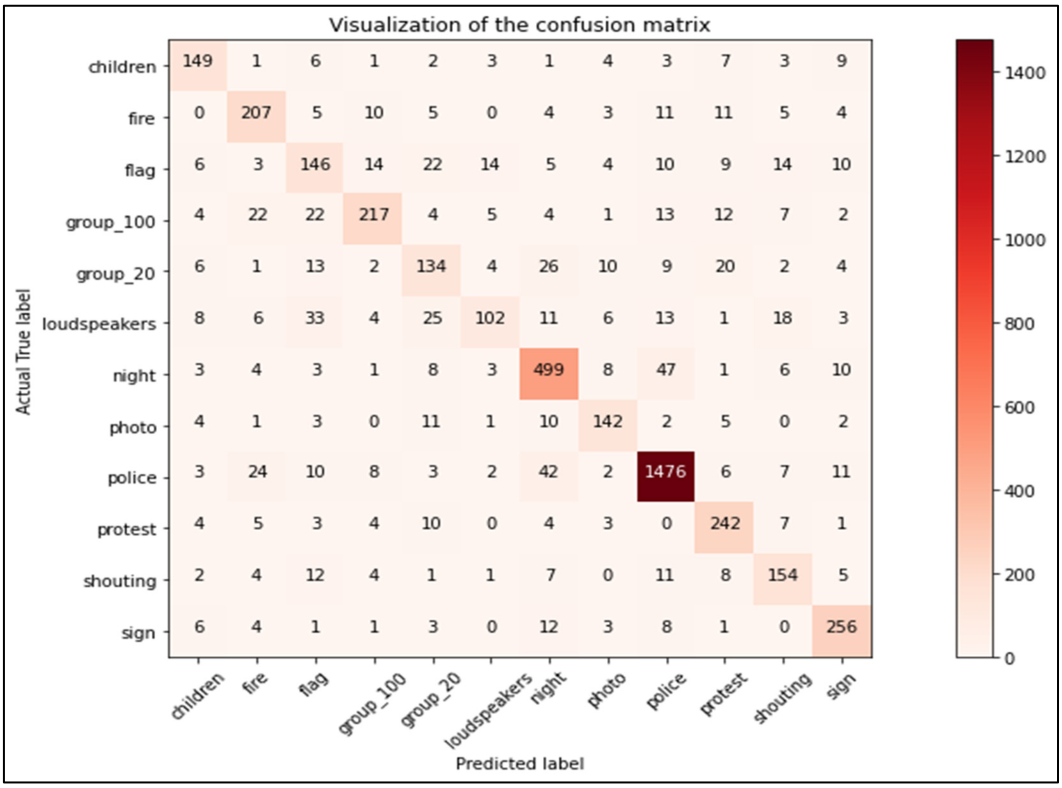

Figure 8. VGG16 confusion matrix for protest dataset.

The training accuracy is a follow:

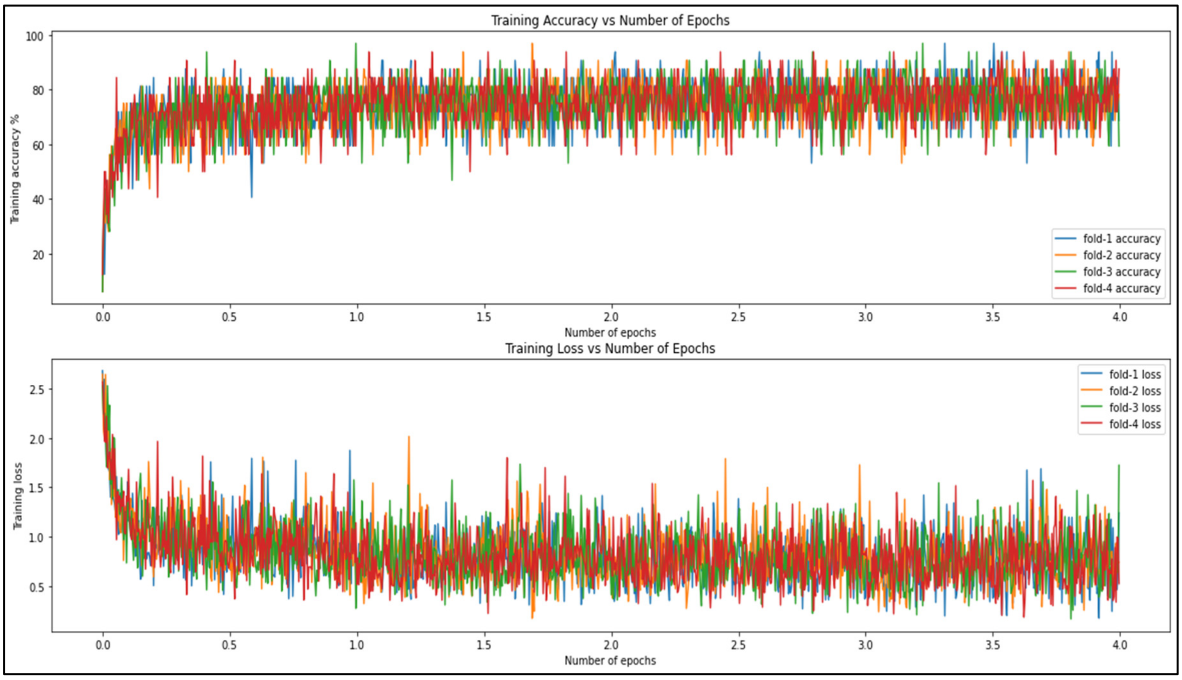

Figure 9. Fivefold accuracy training accuracy vs number of epochs for VGG16.

The results of Inception Resnet V2 in the form of confusion matrix and five-fold accuracy graphs is as follows:

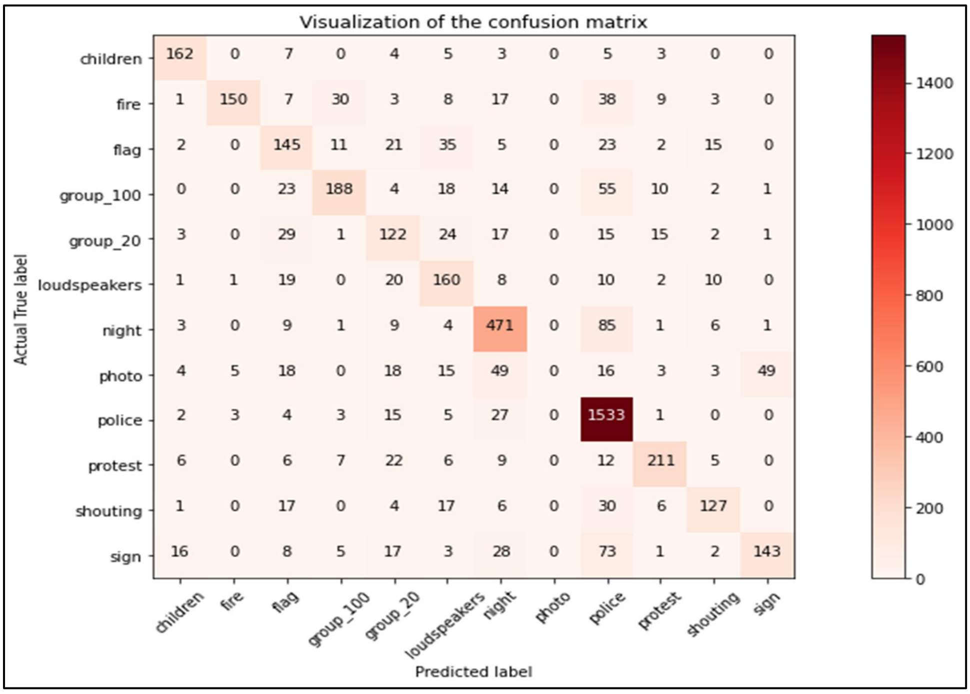

Figure 10. Confusion matrix for protests datasets using Inception Resnet V2.

The 5-fold accuracy graph for all the three approaches is as follows:

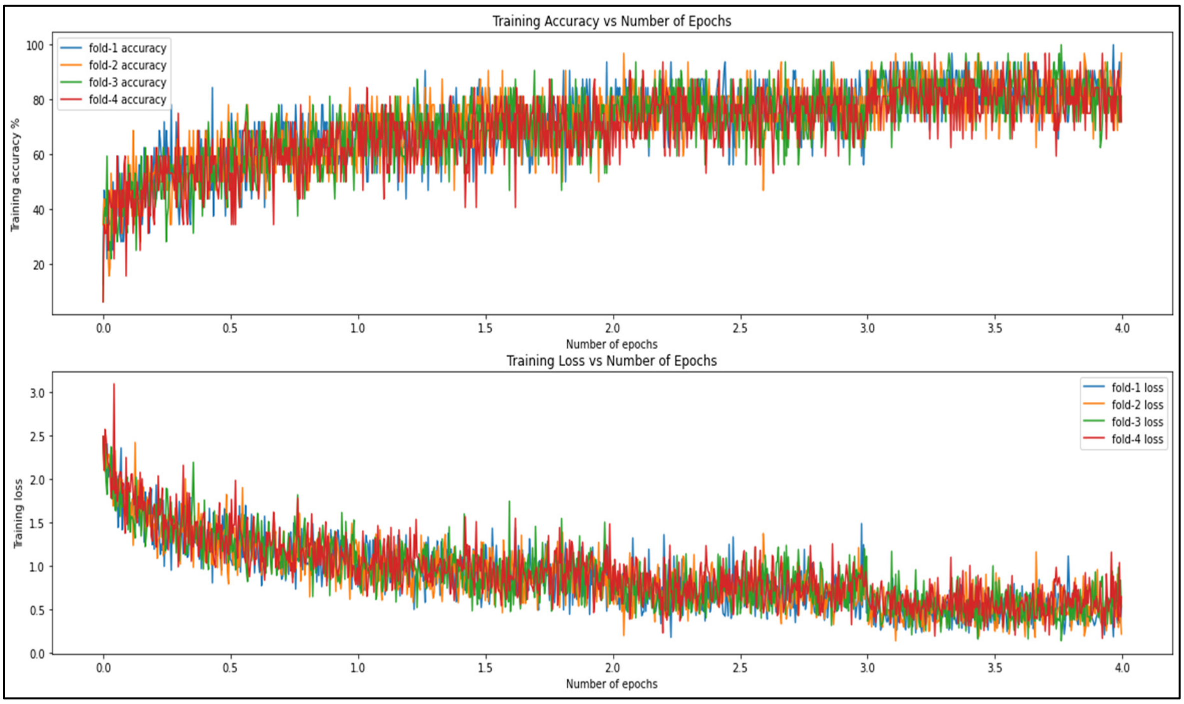

Figure 11. Five-fold training accuracy vs number of epochs graph.

Table 7. Accuracies comparison table.

	Accuracy	F-score	Recall	Precision
VGG16 [26]	80.259	0.738	0.739	0.746
Inception ResNet V2 [27]	82.259	0.758	0.789	0.776

From the above Table 7, it becomes clear that our proposed framework achieves substantial performance on the protest dataset proving that our association strategy of pretrained ConvNets for protest detection have given practical results. The results are promising and they can be further improved using more balanced dataset by improving protest dataset. The results in the multiclass classification are returned in the probability. The returned probabilities of the classes beyond a threshold are taken as the entity present in the input images. Multiclass classification helps in classification of image datasets and helps in classifying the images. Based on this multi class classification various rule set can be framed and a decision support system can be easily made for assisting humans. Further a simple binary classification for classification of protests and violence category is also complied in this research. Here a binary classification between protest and violence dataset is undertaken. The aim for such a classification is that the detection of violence in the input images/videos can make the alert system more perfect for alerting healthcare systems personnel.

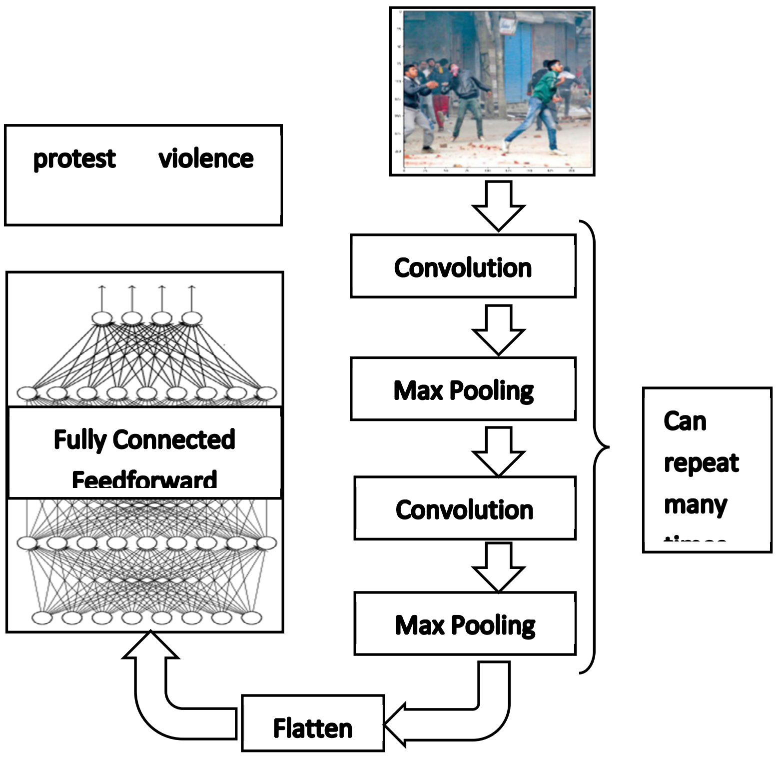

Figure 12. A simplistic diagram of ConvNets module for image classification.

The protest and violence dataset were subsampled to total of 9000 images each. A classification scheme using pretrained ConvNets was utilized. We use VGG16 for the classification purpose. Finetuning of VGG16 is undertaken. The top fully connected layers are removed newly fully connected layer is added to the existing pretrained ConvNets. This new layer consisted of global max pooling layers, one fully connected layer, one drops out layer with 0.49 rate and one softmax activation function for 2 types of categories. Feature extraction is performed for the newly added FC layer. This ensure that the weights in the system is not random and gradient would be under control. Post feature extraction the final ConvNets layer is unfreeze and fine tuning of the model starts for father iteration. The learning is rate is 0.001 and the Ada Grad optimizer is applied on the classification algorithm.

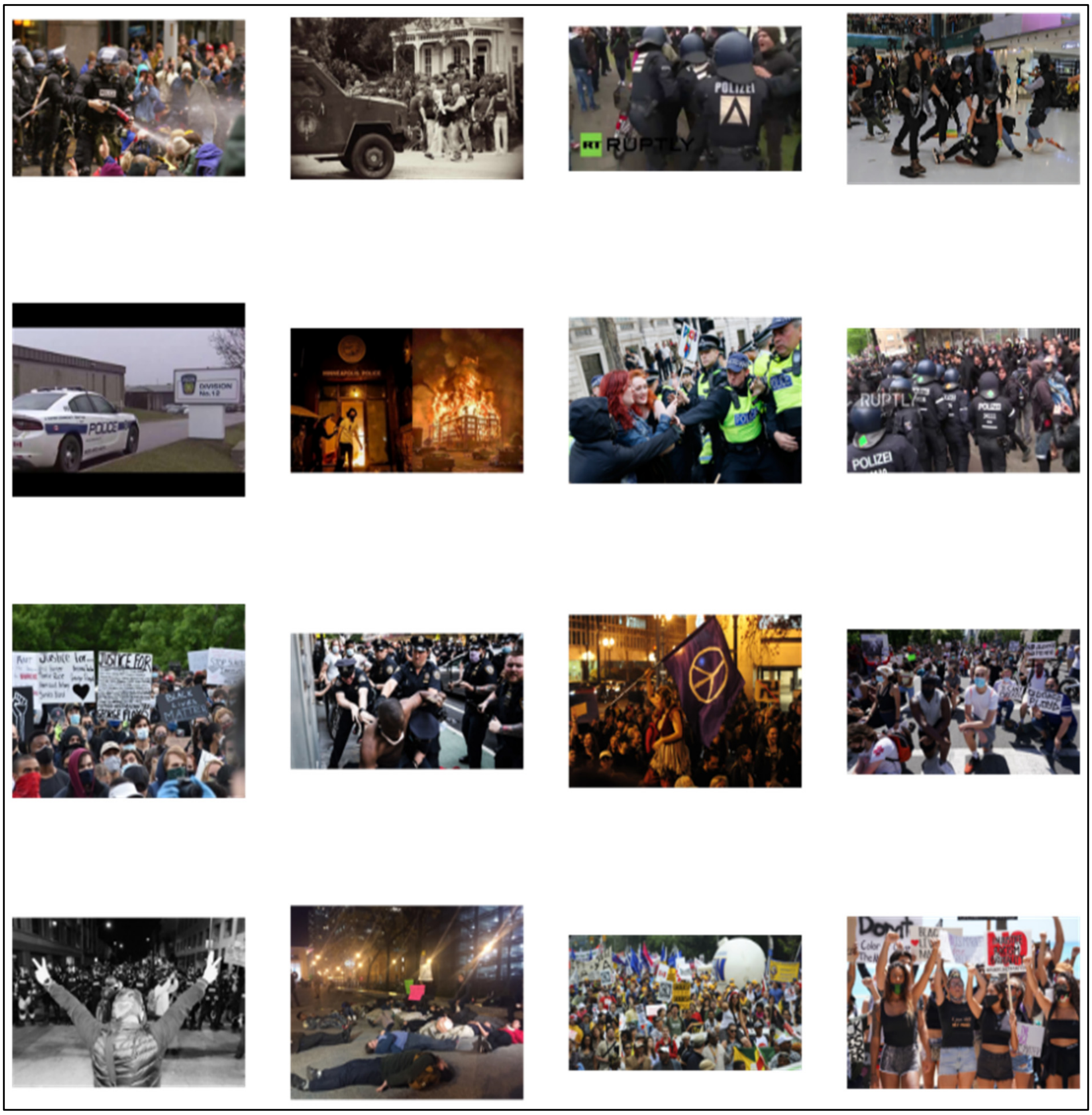

Figure 13. Sample images of protests and violence.

The results of the system are as follows:

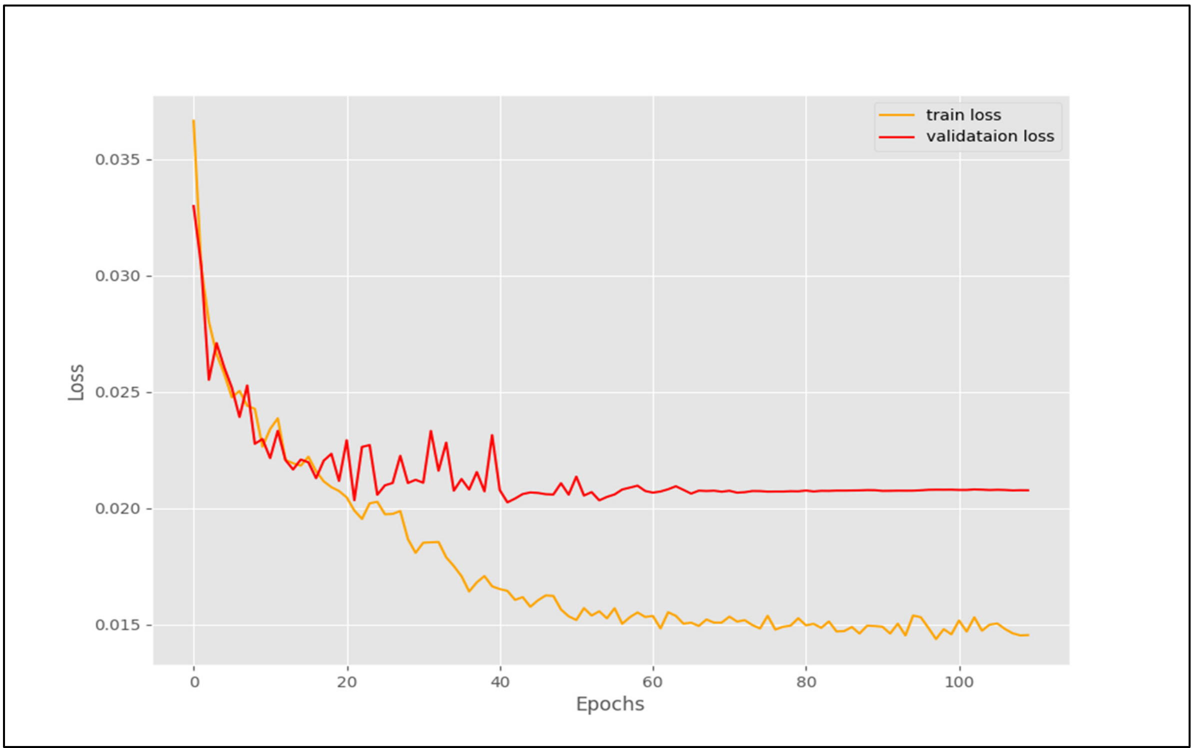

Figure 14. Training loss and validation for protest and violence classification.

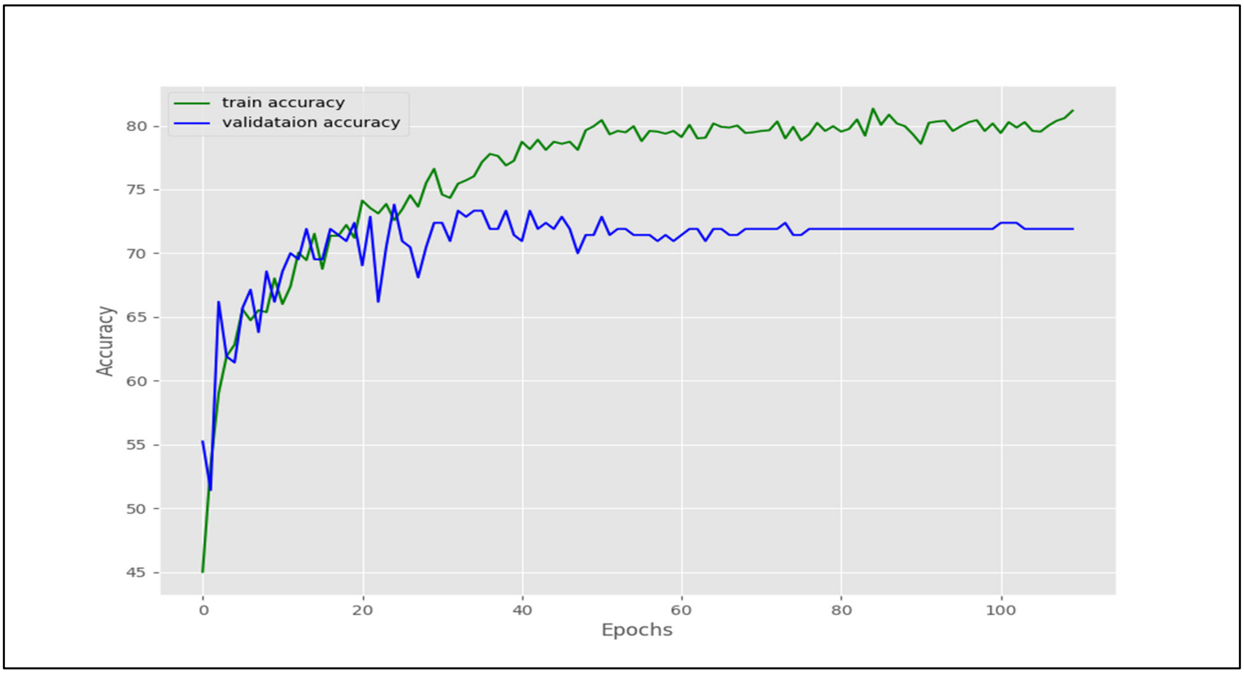

Figure 15. Training accuracy and validation accuracy for protest and violence classification.

Fine tuning the whole model using transfer learning gives better result and the model converge faster.

Table 8. Accuracy of the protests and violence dataset.

Protests and Violence dataset	Accuracy
VGG16 [26]	92.789%

Figure 15 and Table 8 shows the accuracy and loss diagrams of the VGG16 approach for classifying protest and violence. To the best of knowledge, this is a unique classification scheme of classifying image categories of protests and violence. Violent activities automatically qualify for tagging of specialist healthcare system and surgeons. Figures 14 and 15 describes the result and shows overfitting results. The sample is size is limited which may not generalize well on unseen data. As per Figure 12 a simplistic framework for classification is shown where convolution and pooling layers combinations extend in to the deep learning architecture. The approach to tackle overfitting is to optimize and reduce the deep learning capacity by removing convolutional layers as shown in Figure 12 and reducing the number of elements in the hidden layers. Application of regularization and usage of dropout layers can help in reducing the overfitting in the model performance. A recent research article in the medical field has proposed a novel deep learning model for more efficient and precise recognition of Tetralogy of Fallot (TOF) [64] has discussed a novel method of deploying stochastic pooling methods to counter overfitting models. A rank-based stochastic pooling (RSP) for breast cancer classification has been recently published [65] . A future version of this paper can be extended by combining two advanced neural networks namely Graph convolutional network (GCN) and ConvNets (CNN).

6. Discussion

Protest detection for crowd management is a challenging job. This research has presented a way to link crowd behaviour to deep learning and tagging healthcare system to the safety of people. There are various reasons which affects individuals and group of people which can trigger the cause of protests. Protest can manifest itself due to various reasons which includes anger, emotions, identity to social embeddedness. The idea percolated in this research is that protests are collective crowd behaviour. If not handled properly, protests can spiral in to violence. The paper explores post protests experiences and safety of the people who are in protesting, police who are controlling the protestors and normal crowd who are just bystanders, onlookers who are witnessing the protesting events. Thus, a lot of people gets involved, voluntarily or involuntarily. The paper explores this crowd behaviour using protest dataset. Using multiple attributes of the protest dataset this research classifies the severity of crowd protests using rule set belonging to the respective attributes of protest dataset. Converging of protests in to violence can cause stern retaliatory action from the law enforcement agencies. The significance of this research is to come out with a safety plan of people who may get injured in the protests. This research highlighted the common injuries that happens during severe protests. These injuries are sustained by the protestors as well as the police forces. Even the bystanders can get fatally injured in these protests. The highlight of the research is to make a protest storyboard. People protesting would definitely face police controllable actions. Protesting people are united either by anger or identity issues, they usually turn violent to make their point of view heard by the government system and authorities. Amidst all this chaos, if the healthcare systems, professionals, ambulances, medicine stores etc are ready for protest and violence situation, the objective of this research is achieved. At the end of the day no human lives must be lost due to violent protests. Crowd behaviour like protests and riots is a very hard problem to solve using any images datasets. Further the current research links the type of injuries that can happen during protests and severe protests. Each such injury is mapped to a specialist health care personnel to specific surgeons. Also, it is imperative to view the tagging of healthcare systems from two point of view. The first view is dependent solely on the multiclass protest classification. Based on their respective threshold probability score, the rule sets are fired and an alert is generated for tagging healthcare systems. The second view is to generate the classification of each frame of the real time of the input video in to protests and violent categories. This research believes that if the violent state is detected then it automatically is qualified for tagging specialist healthcare systems and doctors to be at alert for saving human lives. This research can be a game changer in the IoT enabled smart cities and devices. Future versions of this can be Android complied and an app can be made functional for directly pushing notifications to the healthcare professionals. All the medicines, injections, anti-biotics, oxygen cylinders needed for such scenarios could be readied beforehand and the human lives can be saved even in the midst of sever protests and violence.

6.1. Contexts, scenarios and solution

People believe in the power of protests. The time frame of acceptance of the protesters' demands may vary. In the initial stages of protests, the event looks like a failure but the power of protests becomes visible in the long term. A system is envisaged which interconnects the subsystems of healthcare tagging. This system is connected to the videos / live stream of the protesting sites. These frames from these feeds are extracted and pumped into the trained model which classifies the protest attributes. Each combination of protest attributes forms a rule which is triggered respectively. There are usually two types of healthcare centers. Primary healthcare centers are controlled by the government and the private hospitals. All the primary health care centers are geo mapped. As soon as the protests' severity begins, their respective classification and notifications regarding the state of the protests begin to generate. These notifications are pushed through the pipeline and concerned authorities are informed about the ongoing situations. Following prototype is envisaged to prove our points of tagging healthcare amidst protests and violence. There are eight camera sub screen which can receive the video feeds from the CCTV cameras via RTSP protocols. For testing purpose, the videos have been taken from Internet and a proof-of-concept procedure is developed. The images shown are the camera feeds of the protesting and violent crowds and the respective police presences. Protest severity is determined and push notifications to the ten primary health care centers are sent to be ready for receiving patients with external injuries. This research can be used in different protests scenarios. For example, in a developing country like India, two recent protests have created a widespread headline all over the world. These protests are farmers protests [66] which is still in continuing and protests against the Citizenship Amendment Act [67] have given the impetus to authors to analyze the protests scenarios and minimize injuries to people, police and innocent bystanders. Below there are samples of images that highlights the proof of system images where feeds can be received on line from a connected CCTV / camera. As it has already been mentioned in Figure 1 of introduction section, the concept is to make a full fledge system where a command-and-control center deals with all the feeds. It is accompanied by the preprocessing and decision support system. All these operations converge onto crowd controlling deployable system using the deep learning framework.

The video streams undergo deep learning based analysis and severity of protests is ascertained. The final results goes to the decision support system which helps the human operator to take a final call on the prevailing situation.

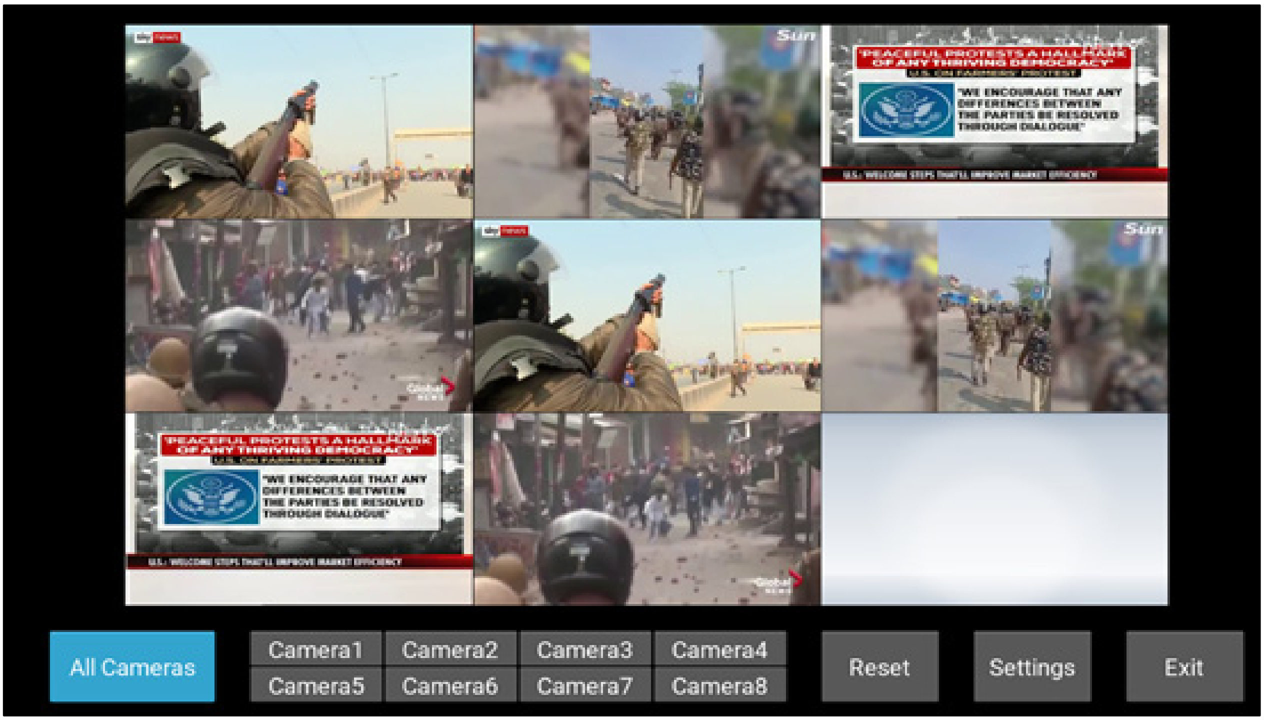

Figure 16. Main screen for the video feeds from the protest's sites.

The system shall take feeds from various CCTV feeds. For testing purpose, the system is shown to have eight screens. It can be extended to multiple screens. The system can see all feeds in one screen or if a user wants to explore the feeds, individual camera screen can be accessed.

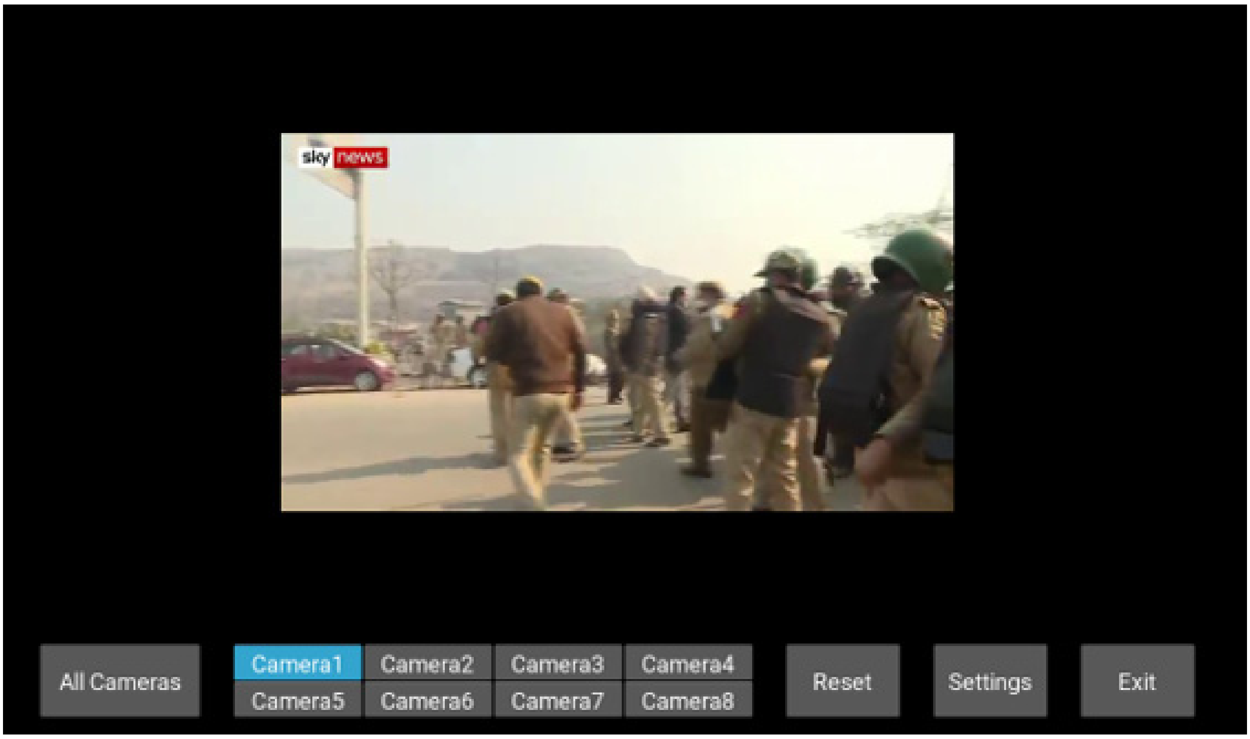

Figure 17. Camera 1 screen for the video feeds from the protest's sites**.**

Figure 18. Camera 2 screen for the video feeds from the protest's sites.

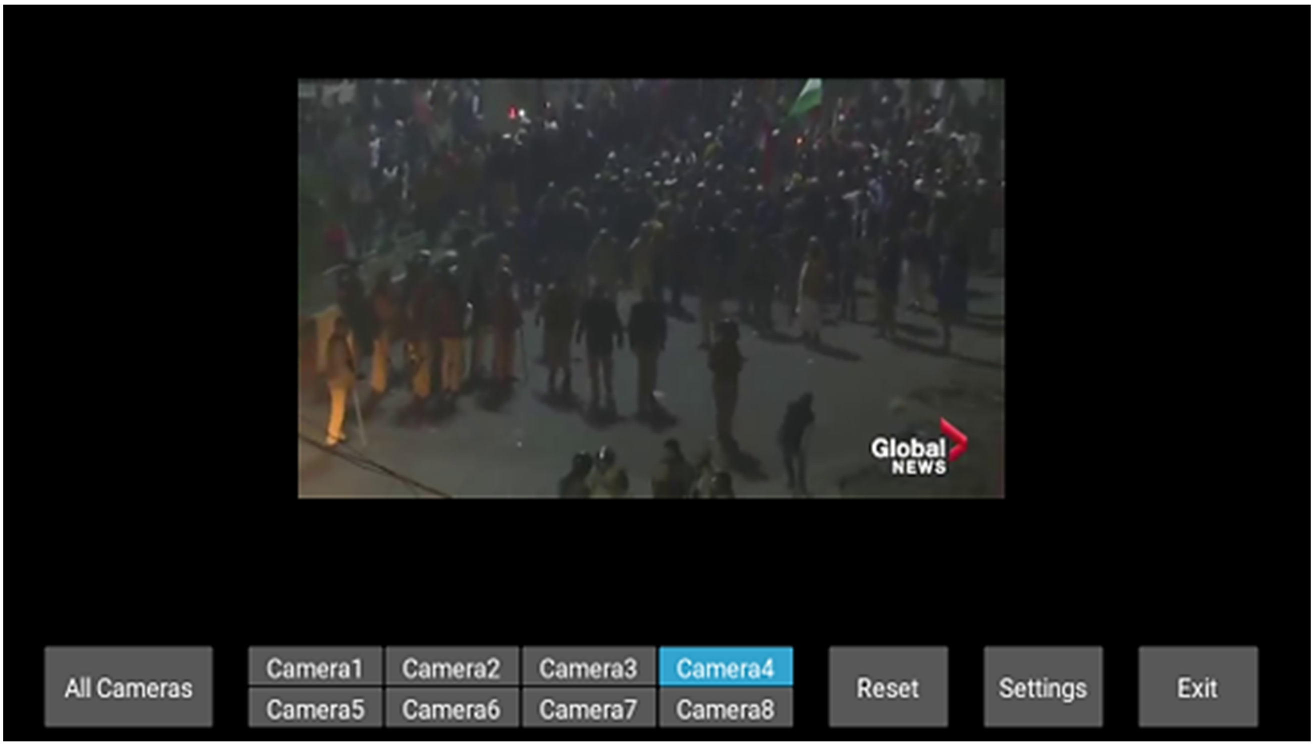

Figure 19. Camera 3 screen for the video feeds from the protest's sites.

The feeds that are received is ingested by the trained model. The model classifies it presence of various protests attributes. As soon as the protests attributes are received in the single frame, the rules are fired. Based on the rules, the ten nearest primary healthcare centers are identified and notifications to these centers are pushed via messaging application. The researchers assume that there are two levels of healthcare centers involved in this strategy. First the Primary Health Care centers of the Government, and top government hospitals and secondly all private hospitals which are well equipped to handle severe traumatic cases. The assumption is that in the actual implementation details all the health care centers would be marked on a map and a central authority would be responsible for sending the grievances messages to the nearest healthcare centers. Based on their responses, actual actions would be initiated to inform the driver of the ambulance which in turn can take the injured persons to the nearest healthcare centers.

7. Conclusions

This research is a unique attempt to link the protest storyboard to healthcare systems. This research uses deep learning for multiclass classification of the protest datasets. Based on the respective thresholds of the protest's attributes, a rule set has been presented which classifies the severity of protests in to normal, medium and severe category. Also, it shows an exclusive differentiation between protests and violence. In case of violence a direct tagging of healthcare infrastructures and doctors can be raised by the human in the loop. This is a novel study to alert the specialist healthcare doctors to be on standby in case the severity of protests breaches the thresholds. This framework can be integrated further with an IoT environment and Android application so that respective messaging push notification can be send to directly to all the concerned specialist healthcare personnel. The researchers have assumed that it is full fledge Internet based system with linkage to CCTV cameras, the video streaming, the geo-location of all the healthcare centres situated in the city on a map, and a human in the loop operator which can guide the messaging of injured person to the nearest health care centres. As soon as the healthcare institution accepts the injured person, the system moves on the next phase of taking care of the injured persons with specialist doctors, surgeons, equipment and so on. The main objective is to save human lives of the protestors as well as the police personnel. Automatic surveillance becomes a lot easier and any sort of protests in public places can be tactfully handled. A real time framework of assessing the protests site can help prepare for any eventuality. The current research can be a big boon for tagging healthcare systems and can ensure availability of specialist doctors as per the severity of doctors, medicines, ambulances and other healthcare entities.

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. D. Singh, G. Tripathi, A. J. Jara, A survey of Internet-of-Things: Future vision, architecture, challenges and services, in *2014 IEEE world forum on Internet of Things (WF-IoT)*, 2014.
2. C. Guy, Wireless sensor networks, in *Sixth international symposium on instrumentation and control technology: Signal analysis, measurement theory, photoelectronic technology, and artificial intelligence*, (2006), 635711.
3. N. S. Kumar, B. Vuayalakshmi, R. J. Prarthana, A. Shankar, IOT based smart garbage alert system using Arduino UNO, in *IEEE Region 10 International Conference TENCON*, 1028–1034.
4. M. S. Munir, I. S. Bajwa, A. Ashraf, W. Anwar, R. Rashid, Intelligent and Smart Irrigation System Using Edge Computing and IoT, *Complexity*, 2021.
5. R. Rajavel, S. K. Ravichandran, K. Harimoorthy, P. Nagappan, K. R. Gobichettipalayam, IoT-based smart healthcare video surveillance system using edge computing, *J. Ambient Intell. Humaniz Comput.*, **3** (2021), 1-13
6. F. Malik, M. A. Shah, H. A. Khattak, Intelligent transport system: An important aspect of emergency management in smart cities, in *2018 24th International Conference on Automation and Computing (ICAC)*, 2018.
7. A. H. Sodhro, M. S. Obaidat, A. Gurtov, N. Zahid, S. Pirbhulal, L. Wang, et al., Towards wearable sensing enabled healthcare framework for elderly patients, in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, 2020.
8. A. H. Sodhro, A. Gurtov, N. Zahid, S. Pirbhulal, L. Wang, M. Rahman, et al., Toward convergence of AI and IoT for energy-efficient communication in smart homes, *IEEE Int. Things J.*, **8** (2020), 9664–9671.
9. J. V. Stekelenburg, B. Klandermans, The social psychology of protest, *Curr. Sociol.*, (2013), 886–905
10. J. Joo, W. Li, F. F. Steen, S. C. Zhu, Visual persuasion: Inferring communicative intents of images, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014.
11. W. S. DeKeseredy, B. Perry, Advancing critical criminology: theory and application, *Lexington Books*, 2006.

12. https://www.merriam-webster.com/dictionary/protest, Available: https://www.merriam-webster.com/dictionary/protest. [Accessed 13.2.2021 2 2021.]
13. https://www.macmillandictionary.com/dictionary/british/protest_1, Available: https://www.macmillandictionary.com/dictionary/british/protest_1. [Accessed 13 02 2021.]
14. Z. C. Steinert-Threlkeld, Spontaneous collective action: Peripheral mobilization during the Arab Spring, *Am. Polit. Sci. Rev.*, (2017), 379–403
15. Z. C. Steinert-Threlkeld, D. Mocanu, A. Vespignani, J. Fowler, Online social networks and offline protest, *EPJ Data Sci.*, (2015), 1–9
16. S. González-Bailón, J. Borge-Holthoefer, Y. Moreno, Broadcasters and hidden influentials in online protest diffusion, *Am. Behav. Sci.*, (2013), 943–965.
17. R. Enikolopov, A. Makarin, M. Petrova, Social media and protest participation: Evidence from Russia, *Econometrica*, (2020), 1479–1514.
18. T. Senst, V. Eiselein, A. Kuhn, T. Sikora, Crowd Violence Detection Using Global Motion-Compensated Lagrangian Features and Scale-Sensitive Video-Level Representation, in *IEEE Transactions on Information Forensics and Security*, (2017), 2945–2956.
19. J. Lofland, Protest: studies of collective behaviour and social movements, *Routledge*, 2017.
20. H. Grabner, F. Nater, M. Druey, G. L. Van, Visual interestingness in image sequences, in *Proceedings of the 21st ACM international conference on Multimedia*, 2013.
21. M. Redi, N. O'Hare, R. Schifanella, M. Trevisiol, A. Jaimes, 6 seconds of sound and vision: Creativity in micro-videos, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014.
22. A. Krizhevsky, I. Sutskever, E. H. Geoffrey, Imagenet classification with deep convolutional neural networks, *Adv. Neur. Inf. Process. Syst.*, 2012.
23. M. Perez, A. C. Kot, A. Rocha, Detection of Real-world Fights in Surveillance Videos, *Process. IEEE Int. Conf. Acoust. Speech Signal Process.*, 2019.
24. K. Simonyan, A. Zisserman, Two-stream convolutional networks for action recognition in videos, *Adv. Neur. Inf. Process. Syst.*, 2014.
25. R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, *Comput. Sci.*, 2014.
26. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, preprint, arXiv:1409.1556.
27. S. Christian, I. Sergey, V. Vincent, A. Alexander, Inception-v4 inception-resnet and the impact of residual connections on learning, *Comput. Sci.*, 2017.
28. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, et al., Imagenet large scale visual recognition challenge, *Int. J. Comput. Vis.*, (2015), 211–252
29. C. Lum, J. Schallhorn, F. Lum, S. Ramanathan, J. Oatts, A. G. de A. Campomanes, et al., Ocular injuries and blindness caused by crowd control measures in civil protests, *Ophthalmology*, 2020.
30. D. Won, Z. C. Steinert-Threlkeld, J. Joo, Protest activity detection and perceived violence estimation from social media images, in *Proceedings of the 25th ACM international conference on Multimedia*, 2017.
31. Y. Wang, Y. Li, J. Luo, Deciphering the 2016 μs presidential campaign in the twitter sphere: A comparison of the trumpists and clintonists, in *Tenth International AAAI Conference on Web and Social Media*, 2016.

32. D. Chen, K. Park, J. Joo, Understanding Gender Stereotypes and Electoral Success from Visual Self-presentations of Politicians in Social Media, *Adv. Comput. Sci. Profess.*, 2020.
33. D. Ganguly, M. H. Mofrad, A. Kovashka, Detecting sexually provocative images, in *IEEE Workshop on Applications of Computer Vision (WACV)*, 2017.
34. H. Peng, J. Li, Y. Song, R. Yang, R. Ranjan, P. S. Yu, Streaming Social Event Detection and Evolution Discovery in Heterogeneous Information Networks, *ACM Trans. Knowl. Discov. Data*, (2021), 1–33
35. G. Tripathi, K. Singh, D. K. Vishwakarma, Detecting Arson and Stone Pelting in Extreme Violence: A Deep Learning Based Identification Approach, in *International Conference on Intelligent Human Computer Interaction*, 2020.
36. H. Zhang, J. Pan, Casm: A deep-learning approach for identifying collective action events with text and image data from social media, *Sociol. Methodol.*, (2019), 1–57.
37. D. Thenmozhi, C. Aravindan, A. Shyamsunder, A. Viswanathan, A. K. Pujari, Extracting Protests from News Using LSTM models with different Attention Mechanisms, in *CLEF (Working Notes)*, 2019.
38. A. Hanson, K. Pnvr, S. Krishnagopal, L. Davis, Bidirectional Convolutional LSTM for the Detection of Violence in Videos, in *European Conference on Computer Vision*, 2018.
39. S. A. Sumon, M. T. Shahria, M. R. Goni, N. Hasan, A. Almarufuzzaman, R. M. Rahman, Violent Crowd Flow Detection Using Deep Learning, in *Asian Conference on Intelligent Information and Database Systems*, 2019.
40. W. Song, D. Zhang, X. Zhao, J. Yu, R. Zheng, A. Wang, A Novel Violent Video Detection Scheme Based on Modified 3D Convolutional Neural Networks, *IEEE Access*, (2019), 39172–39179
41. X. Xu, X. Wu, G. Wang, H. Wang, Violent Video Classification Based on Spatial-Temporal Cues Using Deep Learning, in *2018 11th International Symposium on Computational Intelligence and Design (ISCID)*, 2018.
42. C. Li, L. Zhu, D. Zhu, J. Chen, Z. Pan, X. Li, et al., End-to-end Multiplayer Violence Detection based on Deep 3D CNN, in *Proceedings of the 2018 VII International Conference on Network, Communication and Computing*, 2018.
43. D. Tran, L. Bourdev, R. Fergus, L. Torresani, M. Paluri, Learning spatiotemporal features with 3d convolutional networks, *Proc. IEEE Int. Conf. Comput.*, 2015.
44. Z. Zhou, M. Zhu, K. Yahya, Violence Behavior Detection Based on 3D-CNN, *Comput. Syst. Appl.*, (2017), 34.
45. B. Mandal, J. Fajtl, V. Argyriou, D. Monekosso, P. Remagnino, Deep residual network with subclass discriminant analysis for crowd behavior recognition, in *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018.
46. S. Ammar, M. R. Anjum, T. Rounak, M. Islam, Using deep learning algorithms to detect violent activities, *Semant. Schol.*, 2019.
47. Z. Meng, J. Yuan, Z. Li, Trajectory-Pooled Deep Convolutional Networks for Violence Detection in Videos, in *International Conference on Computer Vision Systems*, 2017.
48. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016.
49. R. E. Fan, K. W. Chang, C. J. Hsieh, X. R Wang, C. J. Lin, LIBLINEAR: A library for large linear classification, *J. Mach. Learn.Res.*, **9** (2008), 1871–1874.
50. J. Duchi, E. Hazan, Y. Singer, Adaptive Subgradient Methods for Online Learning and Stochastic Optimization, *J. Mach. Learn. Res.*, **12** (2011), 2121–2159.

51. H. Iqbal, Harisiqbal88/plotneuralnet v1. 0.0, URL: https://doi. org/10.5281/Zenodo, 2018.
52. I. Amnesty, Police-brutality, *Amnesty Int.*, 2021.
53. S. Venkatesan, The Riot Trauma: What Injuries Should You Expect From Non-Lethal Police Weapons and Protests?, *Trauma, Soc. EM, EMS*, 2020.
54. M. A. Kadivar, N. Ketchley, Sticks, stones, and Molotov cocktails: Unarmed collective violence and democratization, in *Socius: Sociological Research for a Dynamic World*, **4** (2018), 2378023118773614.
55. B. E. Mangus, L. Y. Shen, S. D. Helmer, J. Mahe, R. S. Smith, Taser and Taser associated injuries: a case series, *Am. Surg.*, **74** (2008), 862–865.
56. K. M. Buchanan, L. J. Elias, G. B. Goplen, Differing perspectives on outcome after subarachnoid hemorrhage: the patient, the relative, the neurosurgeon, *Neurosurgery*, **46** (2000), 831–840.
57. D. K. Raja, D. R. Raja, V. N. Shakul, The role of maxillofacial surgeon in the management of skull base tumors," *IP Int. J. Maxillofac. Imaging*, **2** (2021), 107–109.
58. A. Younas, I. Shah, T. C. Lim, M. Figari, G. Louie, D. Matic, et al., Evaluating an International Facial Trauma Course for Surgeons: Did We Make a Difference?, in *Craniomaxillo-facial Trauma Reconstruction Open*, **6** (2021), 24727512211019245.
59. E. Y. Y. Chan, K. K. C. Hung, H. H. Y. Hung, C. A. Graham, Use of tear gas for crowd control in Hong Kong, *The Lancet*, **394** (2019), 1517–1518.
60. K. Pauley, R. Flin, A. Azuara-Blanco, Intra-operative decision making by ophthalmic surgeons, *Br. J. Ophthalmol.* **97** (2013), 1300–1307.
61. J. W. Eikelboom, G. Karthikeyan, N. Fagel, J. Hirsh, American Association of Orthopedic Surgeons and American College of Chest Physicians guidelines for venous thromboembolism prevention in hip and knee arthroplasty differ: what are the implications for clinicians and patients?, *Chest*, **135** (2009), 513–520.
62. S. Budd, E. C. Robinson, B. Kainz, A survey on active learning and human-in-the-loop deep learning for medical image analysis, *Med. Image Anal.*, (2021), 102062.
63. M. Kuhn, K. Johnson, Applied Predictive Modelling, 2013.
64. S. H. Wang, K. Wu, T. Chu, S. L. Fernandes, Q. Zhou, Y. D. Zhang, et al., SOSPCNN: Structurally Optimized Stochastic Pooling Convolutional Neural Network for Tetralogy of Fallot Recognition, *Wireless Commun. Mobile Comput.*, **2021** (2021).
65. Y. D. Zhang, S. C. Satapathy, D. S. Guttery, Improved breast cancer classification through combining graph convolutional network and convolutional neural network, *Inf. Process. Manage.*, **58** (2021),102439.
66. S. Narayanan, Understanding farmer protests in India, *Academ. Stand. Against Povert.*, **1** (2020), 1.
67. B. Bhushan, South atlantic quarterly, *S. Atl. Q.*, **120** (2021), 201–208.

AIMS Press

©2021 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0)