

# A Deep Learning-Based Accident Detection in Emergency Situation from CCTV Camera Using H-RNN Algorithm

<sup>1</sup>Dr. S. Parthiban, <sup>2</sup>Dr. P. Santhosh Kumar, <sup>3</sup>Venkatesh. K, <sup>4</sup>Saravana Gokul G, <sup>4</sup>Siva. N

<sup>1,2,4</sup>Assistant Professor, Dept. of Computer Science and Engineering, <sup>3</sup>Research Scholar,

<sup>1</sup>Jain (Deemed to be University), Bangalore, India,

<sup>2</sup>SRM Institute of Science and Technology, Ramapuram, Chennai, India.

<sup>3</sup>Annamalai University, Chidambaram, India.

<sup>4</sup>Siddharth Institute of Engg. and Tech., Puttur, Andhara Pradesh, India

e-mail: [parthibans1983@gmail.com](mailto:parthibans1983@gmail.com), [psanthoshkumar.id@gmail.com](mailto:psanthoshkumar.id@gmail.com), [venkiur@gmail.com](mailto:venkiur@gmail.com)  
[sharavana.ssg@gmail.com](mailto:sharavana.ssg@gmail.com), [nsiva5809@gmail.com](mailto:nsiva5809@gmail.com)

## ABSTRACT

Accidents have been a significant reason for deaths in India. Over 80% of accident related incident happen not because of the mishap itself yet the absence of ideal assistance arriving at the mishap casualties. In interstates where the traffic is truly light and speedy a mishap casualty could be left unattended for quite a while. The purpose is to make a framework which would distinguish a mishap dependent on the live feed of video from a CCTV camera introduced on an interstate. The thought is to take each casing of a video and run it through a profound learning - Recurrent neural organization model which has been prepared to arrange edges of a video into mishap or non-mishap in progressive way. Progressive Recurrent Neural Networks has demonstrated to be a quick and precise way to deal with arranges pictures. H-RNN based picture classifiers have given precisions of over 95% for relatively more modest datasets and require less preprocessing when contrasted with other picture characterizing calculations. Contrasted with conventional RNNs, H-RNN is more reasonable to video extraction, since it can misuse long fleeting reliance among outlines, then; the calculation activities are fundamentally decreased.

**Keywords:** *Accident Victims, Deep Learning, RNN, H-RNN, Temporal dependency, Video extraction.*

## 1 INTRODUCTION

These days, the interest in common, military and business reconnaissance is growing up because of the expanding request of security. A huge number of camcorders can be found at public spots, public vehicle, banks, air terminals, and so forth bringing about colossal number of data which is hard to measure progressively. To productively put together developing loads of observation recordings it is important to naturally arrange information utilizing signal based portrayal. Video outline procedures can be exceptionally helpful instrument when applied to the reconnaissance recordings. Principle objective of video synopsis is to distinguish fascinating fragments with regards to the video and present them to the client. Applied to the reconnaissance space, outline strategies can furnish client both with review of the occasions that happened and quicker per using abilities. By identifying and coordinating occasions, embodiment of the observation video is caught in the synopsis diminishing time required for perusing the substance. Despite the fact that reconnaissance frameworks are being used for quite a long time, number of distributions identified with observation space has recently been written in most recent couple of years. Recognition and grouping of occasions is utilized frequently in the writing. Article discovery procedure based non wavelet coefficients is utilized to distinguish frontal and back perspective on walkers. In two distinct models that utilize rundown methods in the reconnaissance space are depicted [1]. Video outline dependent on the advancement of review time, outline skipping and nibbled rate requirement is introduced. For a given transient rate requirement the ideal video synopsis issue is characterized as finding a predefined number of casings that limit the fleeting contortion. In instrument that uses MPEG-7 visual descriptors and produces a video list for outline creation. The subsequent file creates a review of the film and permits non-direct admittance to the substance. This methodology depends on progressive bunching for consolidating shot sections that have comparative highlights and neighbor each other in the time area. In comparability diagram, and use chart dividing standardized cut for grouping shots into scenes. Video movement investigation can be utilized for making video synopses.

In this methodology demonstrated that by examining worldwide/camera movement and item movement is conceivable to separate valuable data about the video structure. More complete diagram of existing strategies and accessible writing on astute reconnaissance frameworks can be found. We present in this paper occasion identification and grouping approach for building both static and dynamic rundown. Primary thought of our methodology is to join video skim with set of key edges coordinated in groups to empower quick perusing of entire video. To make the synopsis we initially distinguish occasions utilizing energy distinction between outlines. At that point we group occasions dependent on their visual appearance, lastly dependent on the bunches structure we assemble the synopsis and present it to the client. Video rundown is utilized to lead client to the key casing group with a particular occasion. Key casing rundown at that point sums up groups containing fascinating occasions. In any case, every one of these methodologies are not able enough to display the rich data in video content. As of late, roused by the incredible accomplishment of profound learning, Convolution Neural Network (CNN) and Recurrent Neural Network (RNN) are acquainted with video outline, where CNN is used to separate profound visual highlights and RNN is utilized to anticipate the likelihood of one subshot to be chosen into the synopsis [2]. This design has accomplished the cutting edge brings about video synopsis. Aside from the incredible capacity of CNN in visual component extraction, this chiefly profit by the ability of RNN in misusing the fleeting reliance among outlines. Tragically, RNN just functions admirably for short edge succession. In any event, for LSTM, one sort of RNN that is the most brilliant in long edge arrangement displaying, the ideal video length is under 80 casings. While to video rundown, a large portion of the recordings contain a great many casings. For this situation, it is hard for RNN to catch this long-range transient reliance of recordings. In this way, current methodologies that apply RNN straightforwardly to video rundown may confine the nature of video outline.

To address this issue, we propose a progressive structure of RNN. As portrayed in Figure 1, the progressive RNN is made out of multi-layers, and each layer is with at least one short RNNs, by which the since quite a while ago info grouping is handled progressively. As a matter of fact, the progressive RNN is an overall design which differs as per explicit errands. In this paper, a particular various leveled RNN is intended for the errand of video rundown, called as H-RNN. Detailed, it contains two layers. The main layer is

a LSTM, which is used to handle video subshots produced by cutting the video uniformly, and the intra-subshot transient reliance is encoded in the last covered up. At that point, the last covered up of each subshot is contribution to the subsequent layer. In particular, the subsequent layer is a bi-directional LSTM, which is made out of a forward and a retrogressive LSTM. It is utilized to abuse the between subshot fleeting reliance and decide if a certain subshot is significant to be a key subshot[3].

## 2. RELATED WORKS

Most people associated with car crashes get help from drivers, travelers, or others. In any case, when an auto collision happens in a scantily populated zone or the driver is the solitary individual in the vehicle and the accident brings about loss of cognizance, nobody will be accessible to send a misery message to the appropriate specialists inside the brilliant window for clinical treatment. Thinking about these issues, a strategy for distinguishing fast head-on and single-vehicle crashes, investigating the circumstance, and raising an alert is required. To address such issues, this paper proposes a profound learning-based Internet of Vehicles (IoV) framework called Deep Crash, which incorporates an in-vehicle infotainment (IVI) telematics stage with a vehicle self-crash discovery sensor and a front camera, a cloud-based profound learning worker, and a cloud-based administration stage[4].

Occasion mining is a helpful method to comprehend PC framework practices. The focal point of later deals with occasion mining has been moved to occasion outline from finding continuous designs. Occasion synopsis looks to give an intelligible clarification of the occasion grouping on specific angles. Past strategies have a few impediments, for example, high time unpredictability, a low accuracy particularly with the presence of commotion and stage shifts, and giving a synopsis which is hard for a human to comprehend. Mishaps have been a significant reason for passings in India. Over 80% of mishap related passings happen not because of the mishap itself yet the absence of opportune assistance arriving at the mishap casualties. In roadways where the traffic is truly light and speedy a mishap casualty could be left unattended for quite a while. The plan is to make a framework which would identify a mishap dependent on the live feed of video from a CCTV camera introduced on a roadway. The thought is to take each edge of a video and run it through a profound learning convolution neural organization model which has been

prepared to characterize edges of a video into mishap or non-mishap. Convolutional Neural Networks has demonstrated to be a quick and exact way to deal with group pictures. CNN based picture classifiers have given precisions of over 95% for relatively more modest datasets and require less preprocessing when contrasted with other picture arranging calculations [5].

Crash Detection is fundamental in giving opportune data to traffic the executives communities and the general population to decrease its unfavorable impacts. Expectation of crash hazard is imperative for staying away from auxiliary crashes and shielding parkway traffic. For a long time, analysts have investigated a few methods for ahead of schedule and exact recognition of collides with help in rush hour gridlock occurrence the executives. With late headways in information assortment procedures, bountiful ongoing traffic information is accessible for use. Large information foundation and AI calculations can use this information to give appropriate answers for the roadway traffic security framework [6]. This paper investigates the plausibility of utilizing profound learning models to recognize crash event and anticipate crash hazard. Volume, Speed and Sensor Occupancy information gathered from side of the road radar sensors along Interstate 235 in Des Moines, IA is utilized for this investigation. This genuine traffic information is utilized to configuration highlight set for the profound learning models for crash identification and crash hazard forecast. The outcomes show that a profound model has better accident location execution and comparative accident expectation execution than best in class shallow models. Also, an affectability investigation was led for crash hazard forecast utilizing information 1-minute, 5-minutes and 10-minutes preceding accident event. It was seen that is difficult to foresee the accident danger of a traffic condition, 10 min preceding an accident. Abusing the transient reliance among video casings or sub shots is significant for the errand of video synopsis. Essentially, RNN is acceptable at worldly reliance displaying, and has accomplished overpowering execution in numerous video-based undertakings, for example, video inscribing and characterization. Notwithstanding, RNN isn't adequately skilled to deal with the video rundown task, since customary RNNs, including LSTM, can just arrangement with short recordings, while the recordings in the synopsis task are ordinarily in longer term. To address this issue, we propose a progressive intermittent neural organization for video synopsis, called H-RNN in this paper. In particular, it has two layers, where the primary layer is used to encode short video sub shots cut from the first video, and the last concealed condition of each

sub shot is contribution to the second layer for computing its certainty to be a key sub shot[7].

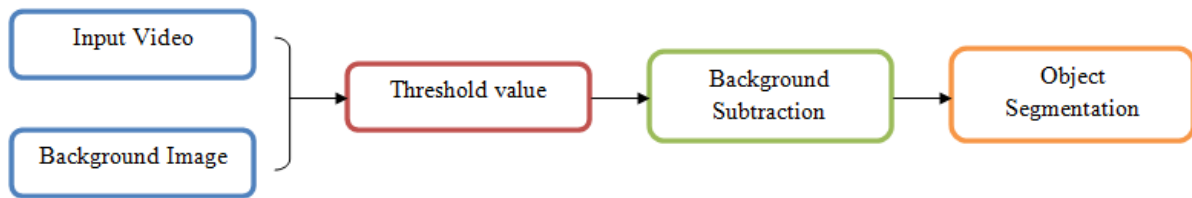
In recent years surveillance video analysis has become an emerging area of research as video surveillance system is used in all places for monitoring the happenings to ensure safety. Automatic video surveillance system is useful for all applications like military surveillance, traffic monitoring, health monitoring of elder people at home, street surveillance. It is essential to represent video events for the further processing like video browsing, retrieval, and summarization. Hence this paper presents a novel method for representing visual events in a video for event detection. For efficient visual event representation shape and motion can be used as features. Hence, pyramid of HOG (PHOG) feature is extracted for shape information and then it is combined with Histogram of magnitude, orientation and entropy of Optical Flow (HMOEOF) to have motion information [8]. The extracted features are used to train the SVM classifier in order to detect and classify the events in the video as normal or abnormal. The proposed method results an equal error rate of 16.71% which shows that it outperforms well compared to other event detection approaches.

### **3 METHODOLOGIES**

This paper applies profound learning calculations including a few model variations to tackle traffic security issues through accident identification and crash hazard assessment on a metropolitan Interstate roadway. The recognition and expectation issue are taken care of as a double characterization issue, with TMC crash reports with reality data as marks, and the encompassing traffic information as tests. High-goal traffic information including volume, speed and sensor inhabitation are utilized in this examination. In the wake of adjusting the classes, distinctive organization structures and preparing activities have been investigated. It tends to be discovered that for identification, Hierarchical Recurrent neural organizations (H-RNN) with drop-out activity performs in a way that is better than some shallow models as far as characterization with a steady preparing measure bereft of any over fitting issue. For forecast, information from various time allotments are tried to additionally explore the model expectation power.

#### **3.1 Background subtraction**

The goal of the subtraction is to section the moving articles that are available in the scene. It is utilized for recordings caught by fixed cameras. To accomplish the movement division, a model of the foundation scene is vital. Each picture in the video succession is contrasted with the foundation model. The pixels having a huge distinction from the model are considered as having a place with moving articles. The yield of a BS calculation is a parallel picture where the moving articles are introduced as a gathering of white pixels (masses). The most fundamental strategy for BS is to take a picture of the scene with static articles as a foundation model. The pixels of moving items are distinguished by applying an edge on the total distinction between the foundation picture and the approaching edge as introduced in Figure 1.



**Figure 1: A flowchart of a basic background subtraction**

### 3.2 Recurrent Neural Network

A standard RNN is built by broadening a feed forward organization with an additional input association, so it can display arrangement. Basically, it can decipher the information arrangement  $x_1, x_2, \dots, x_n$  into another succession  $y_1, y_2, \dots, y_n$  iteratively by the accompanying conditions:

$$a_n = \phi(t_a x_n + w_n a_{n-1} + b_x) \text{-----(1)}$$

$$b_n = \phi(w_n a_n + b_y) \text{-----(2)}$$

where  $a$  is the hidden state,  $n$  signifies the  $n$ -th time step,  $\phi$  stands for the activation function, and  $t, w$  and  $b$  are the training weights and biases.

### 3.3 Hierarchical Recurrent Neural Network

The inspiration for planning progressive RNN is to improve its ability to abuse long-go worldly reliance of the recordings. As a matter of fact, it

is initially enlivened by the activity of one-dimensional convolution. As portrayed in the main layer, a one dimensional channel  $w$  is used to misuse the successive data by performing repetitive procedure on the info arrangement  $a$ :

$$a = w * b; \text{-----}(3)$$

where  $a$  denotes the output sequence, and  $*$  stands for the recurrent operation.

By and large, contrasted with current RNN-based methodologies in video synopsis, H-RNN has the accompanying focal points:

1. H-RNN can show long-run worldly reliance with a brief timeframe step. Subsequently, it diminishes the data misfortune in edge arrangement displaying in the interim the calculation activities are decreased altogether.
2. The progressive construction of H-RNN builds the nonlinear fitting capacity of customary RNN, which has been exhibited very supportive for visual errands.
3. H-RNN abuses the intra-subshot (i.e., among outlines in the subshot) and between subshot worldly reliance in the two layers, separately. This progressive construction is more reasonable for video information, since video worldly design is characteristically layered as casings and subshots.

Every video is an information point that either does or doesn't contain a fender bender. In any case, every video is a bunch of individual pictures that are time-subordinate groupings. The calculation I've picked - a progressive intermittent neural organization - can regard every video as a period subordinate arrangement, yet at the same time permit every video to be a free information point.

The calculation utilizes two layers of long transient memory neural organizations. The primary neural organization (NN) is an intermittent organization that investigates the time-



subordinate grouping of the pictures inside every video. The second takes the encoding of the primary NN and constructs a second NN that reflects which recordings contain mishaps and which don't. The subsequent model empowers an expectation of whether new dashcam film has a mishap.

Through this technique, the HRNN joins a period subordinate part of the casings inside every video to foresee how likely it is another video contains an auto crash.

### **H-RNN Algorithm:**

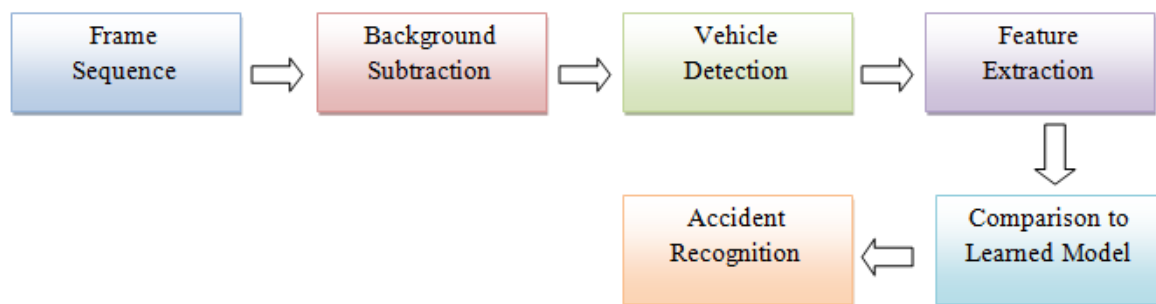
- ✓ Initialize the parameters of training data size  $s$ .
- ✓ Calculate the threshold value by background image.
- ✓ Calculate the background subtraction to segment object.
- ✓ Update Background image  $b_i$ .
- ✓ Compare the given input with training set  $t_s$ .
- ✓ Update the fitness position value  $f_i$ .
- ✓ If the given condition is not meet, then go to step 3.

Every video is a bunch of individual pictures that are time-subordinate groupings. The hierarchical recurrent neural network can regard every video as a period subordinate grouping, yet permit every video to be an autonomous information point. The calculation utilizes two layers of long transient memory neural organizations. The principal neural network (NN) is an intermittent organization that dissects the time-subordinate arrangement of the pictures inside every video. The second takes the encoding of the main NN and assembles a second NN that reflects which recordings contain mishaps and which don't. The subsequent model empowers an expectation of whether new scramble cam film has a mishap. Through this technique, the H-RNN fuses a period subordinate part of the edges inside every video to foresee how likely it is another video contains a fender bender.

### **3.4 The general steps of the video based accident detection techniques**

The situation of road accident utilized in writing concerns the accident between two or numerous vehicles in various sort of streets including

crossing point, metropolitan regions and parkways. Equivalent to irregular conduct identification, the various strides of video based street mishap discovery are: 1) Motion location, 2) Feature Extraction and afterward 3) Features Analysis and Accident Recognition. A square outline of the arrangement of steps to identify a mishap is appeared in Figure with more subtleties. The most every now and again utilized calculations for movement recognition are BS and OF. The highlights identified with the traffic movement and vehicles directions can be extricated utilizing diverse handling methods, for example, following. At long last, the examination of these traffic highlights permits the separation of ordinary movement versus strange as appeared in figure 2.



**Figure 2 : The block diagram of the video based accident detection techniques.**

The target of this progression is the restriction of all moving articles in the picture. It tends to be finished utilizing various techniques, for example, BS, OF and SIFT (the Scale Invariant Feature Transform descriptor).

### **3.5 Vehicle Detection**

After motion detection, the most testing task is vehicle acknowledgment. Moving targets can be spoken to by focuses or mathematical regions like square shapes. Anyway to recognize vehicles from different articles, there is two classifications of approaches: The initial ones are the delegate approaches dependent on the shadings, the edges, the state of vehicles or their parts like windshield and lights. The subsequent one are more perplexing dependent on preparing methods like Haar or descriptors like the Speeded Up Robust Features (SURF) utilized in vehicle

discovery and the Histogram of Oriented Histogram (HOG). The Haar-like course classifier is a classifier prepared with many examples of the item to distinguish the called positive examples from other subjective articles called negatives tests. Subsequent to preparing, a classifier can be applied to a zone important to play out the location. In any case, if there should be an occurrence of HOG, the picture is separated into cells which are assembled into covering blocks. The slope is figured for every pixel and the mean direction of the angle is related with every cell. The vector of highlights is then shaped, for each square, to frame the histograms of the situated angles. These vectors, standardized later, describe the state of the article. Streamed by an order strategy, the idea of item could be perceived. The most utilized classifier are the Support Vector Machine (SVM) and AdaBoost.

### **3.6 Features extraction**

A feature is an attribute of a given article moving. It very well may be speed, direction, directions, region, position and histograms, and so forth Normally joining more than one trademark gives more insights concerning the movement. A few analysts are intrigued on separating highlights of the vehicles moving, different scientists favored extricating highlights, everything being equal (pixels) to have thick movement data at a worldwide scale which we expect to be more exact. A correlation between the separated highlights and predefined conditions permit the discovery of unusual conduct. In this way, the exactness of every calculation of mishap location relies upon the nature of the removed highlights. To have adequate outcomes, highlights should be significant covering the limit of the part of the movement.

### **3.7 Accident recognition**

We characterize them in two classifications. The first depends on the examination of the removed highlights to predefined fixed limits. The subsequent one depends on the correlation of the separated highlights to learned models. These models could be the typical direction of a vehicle or a particular portrayal of its movement, for example, the histogram of speeds. The models of an ordinary traffic are acquired from an enormous dataset of recordings of typical traffic circumstance.

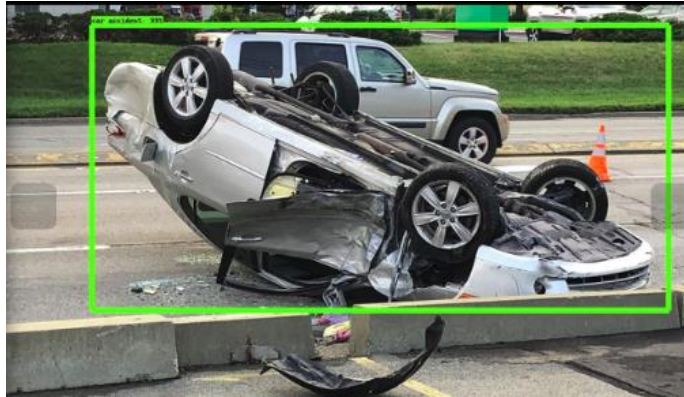
#### 4 EXPERIMENTAL RESULTS

We utilized the CADP dataset for recordings containing mishaps and the DETRAC dataset which was initially for object location of vehicles, as our recordings not containing mishaps. To extend our dataset we additionally download YouTube recordings that contain mishap. More than 380 recordings were gathered from the previously mentioned sources. For the last dataset, we had 188 recordings with vehicle, transport, bicycle and so on mishaps recorded in the CCTV camera at the sides of the road. We took a similar number of negative cases (without mishap) to keep up adjusted classes. Every video is separated into its individual edges to be broke down independently. Every one of these pictures is a two-dimensional cluster of pixels where every pixel has data about the red, green, and blue (RGB) shading levels. To diminish the dimensionality at the individual picture level, we convert the 3-D RGB shading clusters to grayscale. Furthermore, to make the calculations more manageable on a CPU, we resize each picture to (144, 256) - as a result decreasing the size of each picture to a 2-D cluster of 144x256.

$$\text{Precision\%} = \text{TP} / (\text{TP} + \text{FP}) \text{ -----(4)}$$

$$\text{Recall\%} = \text{TP} / (\text{TP} + \text{FN}) \text{ -----(5)}$$

In the open information discipline, informational index is the unit to quantify the data delivered in a public open information vault. The European Open Data gateway totals the greater part 1,000,000 informational indexes. In this field different definitions have been proposed, however presently there is anything but an authority one. Some different issues (continuous information sources, non-social informational collections, and so forth) build the trouble to arrive at an agreement about it. Kaggle rivalries are unfathomably fun and fulfilling, however they can likewise be scaring for individuals who are moderately new in their information science venture. Before, dispatched numerous Playground rivalries that are more receptive than our Featured rivalry, and in this manner more fledgling agreeable.



**Figure 3: Accident Detection with Prediction of crash risk Percentage**

When the framework begins running it considers each casing of the video that it is catching from the Pi-camera and runs it through the proposed model. It likewise sends the edge at which it recognized a mishap and what level of mishap it is. It additionally shows the time stamp regarding when the mishap was distinguished. Figure 3 shows a accident outline alongside the subtleties.

#### **4.1 Data and Processing**

The ACCV dataset contains in excess of 1000 recordings without accidents and more than 600 with crashes. Nonetheless, a significant number of the negative (non-crash) recordings show a similar driving scene - this set was winnowed to have exceptional recordings and to eliminate recordings with logos, introduction scenes, and so forth this left 439 non-crash recordings of quiet driving.

Of the 600 recordings with mishaps, the kinds of mishaps were especially assorted, including any blend of walkers, bicyclists, mopeds, vehicles, trucks, and so on, a considerable lot of which were far somewhere out there of the dash cam film. Even in the wake of getting rid of non-auto and distance crashes, the wide assortment of the scenes made preparing a model troublesome. With 36 of these recordings indicating head-on crashes including vehicles, enhanced the information by going to YouTube, where found and separated 4-second clasps for 93 instances of head-on impacts.

For the last dataset, sum of 129 positive recordings (those with crashes), and arbitrarily tested an equivalent number of negative recordings (those without a collide with) guarantee adjusted classes. Extra preparing was investigated for this task, for example, middle deducting out the foundation of the pictures and featurizing each edge in the picture. The last was finished with the pre-prepared origin and imagenet models (these are models that generally depend on convolutional neural organizations to group pictures dependent on a decreased arrangement of removed highlights from the pictures). These methodologies didn't create a perceptible improvement and were avoided with regard to the end result.

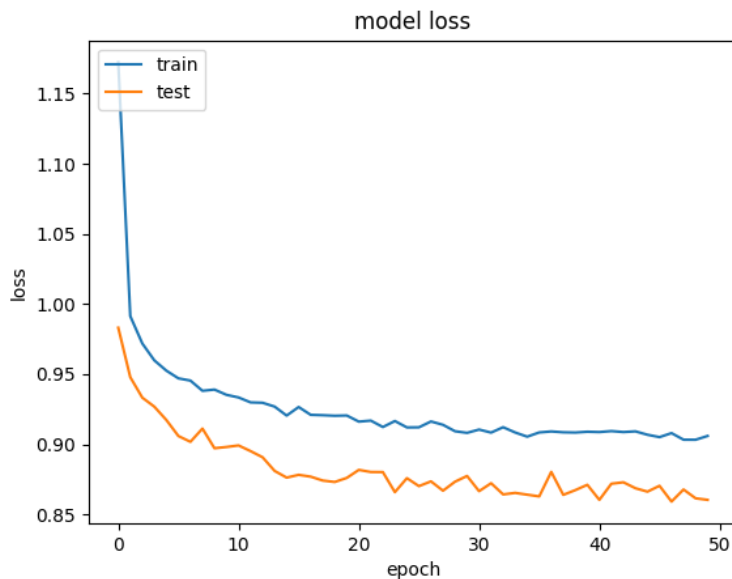
Test Loss: 0.177529

Test Accuracy of accident: 90% (468/517)

Test Accuracy of noaccident: 93% (890/947)

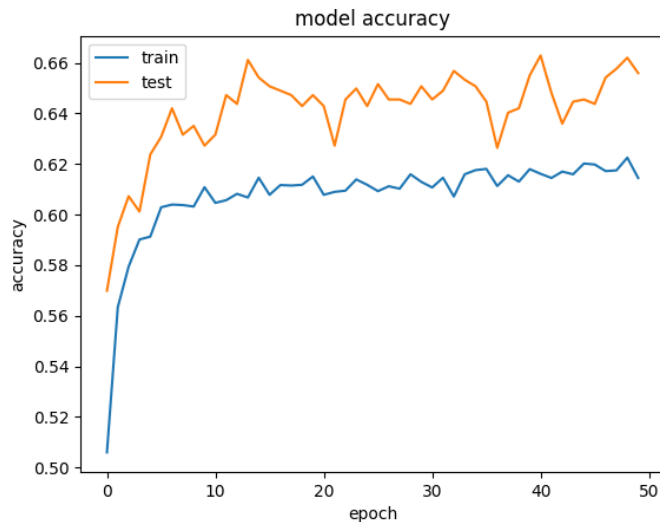
Test Accuracy (Overall): 92% (1358/1464)

**Figure 4: Accuracy Result and Loss Percentage**



**Figure 5: Train and Test Loss during Training**

It might likewise be recognized by an approval misfortune that is lower than the preparation misfortune. For this situation, it demonstrates that the approval dataset might be simpler for the model to foresee than the preparation dataset.



**Figure 6: Train and Test Accuracy during Training**

The running model gave us an exactness going from 82% to a greatest precision of 98.76%. On a normal the model gave us an exactness of over 92.38%.

## 5 CONCLUSIONS

In this paper, we propose a various leveled structure of H-RNN to upgrade the ability of conventional RNN in long-range fleeting reliance catching. Especially, for the undertaking of video rundown, we plan a specific two-layer RNN as per the layered video structure, called as H-RNN. Especially, the primary layer is a LSTM, which is used to abuse the intra-subshot transient reliance among outlines. The subsequent layer is a bi-directional LSTM that can catch both the forward and in reverse between subshot fleeting reliance, and the yield of the subsequent layer is used to anticipate whether a certain subshot is important to be chosen into the outline. Contrasted with current RNN-based methodologies, H-RNN is more appropriate to the undertaking of video synopsis, and the test results have checked its predominance.

## REFERENCES:

1. Chang, W. J., Chen, L. B., & Su, K. Y. (2019). Deep Crash: A Deep Learning-Based Internet of Vehicles System for Head-On and Single-Vehicle Accident Detection With Emergency Notification. *IEEE Access*, 7, 148163-148175.
2. Yang, L., & Yang, N. (2019). An Integrated Event Summarization Approach for Complex System Management. *IEEE Transactions on Network and Service Management*, 16(2), 550-562.
3. Lei, J., Luan, Q., Song, X., Liu, X., Tao, D., & Song, M. (2018). Action Parsing-Driven Video Summarization Based on Reinforcement Learning. *IEEE Transactions on Circuits and Systems for Video Technology*, 29(7), 2126-2137.
4. Singh, R., Saluja, D., & Kumar, S. (2020). Reliability Improvement in Clustering Based Vehicular Ad-hoc Network. *IEEE Communications Letters*.
5. Ghosh, S., Sunny, S. J., & Roney, R. (2019, March). Accident Detection Using Convolutional Neural Networks. In *2019 International Conference on Data Science and Communication (IconDSC)* (pp. 1-6). IEEE.
6. Dai, W., Wang, H., Song, Y., & Xin, Y. (2019, July). Performance Optimization of Wildlife Recognition with Distributed Learning in Ecological Surveillance. In *2019 IEEE 9th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)* (pp. 347-352). IEEE.
7. Huang, T., Wang, S., & Sharma, A. (2020). Highway crash detection and risk estimation using deep learning. *Accident Analysis & Prevention*, 135, 105392.
8. Bahbahani, M. S., & Alsusa, E. (2020, May). A Directional Clustering Protocol for Millimeter Wave Vehicular Ad hoc Networks. In *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)* (pp. 1-6). IEEE.