Machine Learning and Vision for Front Load Trucks in Waste Management System

Project Stakeholders

Project Instructor :

Prof. Muhammad Shahid

Data Analytics for Business

St. Clair College, Windsor

Project Sponsor :

Mr. Chris HO

General Project Manager

Metric Masters Ltd, BC

Project Executed By:

Prathap David Carlo

Data Analytics for Business

St. Clair College, Windsor

Agenda

- Problem Statement
- Factors Contributing to Problem Statement
- Analytical Problem
- Goals, Objectives and Metrics of the Project
- Data Source
- Tools and Technology
- Exploratory Data Analytics
- Machine Learning Models
- Results
- Future Recommendations
- Conclusions
- References

Problem Statement

The need for having an efficient Waste Management Process is very essential to save our planet due to the enormous growth in human population and overconsumption of resources. Currently the garbage collection trucks are collecting the waste bins located across the cities with human assistance and manual interventions. However, it is impractical for the management to verify the accuracy of the executed data in terms of number of actual number of bins collected from the reports that is submitted by the truck operators to management upon completion of their daily work. Due to this, Waste management organizations are spending a lot of time and money in human resources, besides knowing the fact that it creates the room for malpractices through inaccurate data which in turn creates grievance/complaints from the customers claiming that there is no service is done to them.

CONTRIBUTING FACTORS TO PROBLEM STATEMENT

From customer perspective

A customer is charged 3 ways:

- ❑ Rental of a bin.
- ❑ A per lift/tip service.
- ❑ Approximate weight disposed of at the landfill.

Why inaccurate data is produced by truck operators?

- ❑ Moving the bin out of the way to get to a bin behind it
- ❑ Relocating a bin to a different location on the same site
- ❑ Lifting it to clear snow
- ❑ Load is contaminated.
- ❑ Driver spots material that's not allowed in bin.

These aspects create room for malpractices that results in bad lift/No-lift/incomplete lift.

To avoid this, we need to have a lift detection technique to produce accurate results.

Analytical Problem

Detection of Real-Time incidents recorded at every 3 seconds timestamp by cameras installed on front of the trucks using Deep Learning Architecture

- **Good Incidents: Complete Lift**

❑ Front Load Garbage Truck lifts the bin

❑ Dump garbage

❑ Return to its initial stage/any other place.

- **No Incidents : No Activity**

❑ If the vehicle is at the traffic

❑ Obstacles in between

❑ Not at the bin location

- **Bad incidents: Not Complete Lift**

❑ Front Load Garbage lifts the bin

❑ Doesn't dump garbage

❑ Return to its initial state/ any other place.

Related Work

Paper 1: Human Walking Activities Recognition [5]

This research work presents the multitask human walking activities recognition using human gait patterns. The data are collected for 50 subjects in a controlled laboratory environment using inertial measurement unit (IMU) sensors for 7 different activities. Four deep learning models are used. CNN–LSTM, CNN–GRU, LSTM–CNN and LSTM–GRU and their respective accuracies 97.26%, 90.67%, 77.38%, 97.83%, 94.35%, 97.64%, 96.98%.

Paper 2 : LSTM-CNN network for human activity recognition using Wi-Fi CSI data [6]

Wi-Fi based human activity recognition system that can identify different activities via the channel state information from Wi-Fi devices. A special deep learning framework, Long Short-Term Memory. Convolutional Neural Network (LSTM-CNN), is designed for accurate recognition. LSTMCNN is going to be compared with the LSTM network and the experimental results demonstrate that LSTM-CNN outperforms existing models and has an average accuracy of 94.14% in multiactivity classification

Paper 3 : Classification of Mental Stress Using CNN-LSTM Algorithms with Electrocardiogram Signals [7]

The mental stress faced by many people in modern society is a factor that causes various diseases. The person's stress can be analyzed using Electrocardiogram (ECG) Signal and some parameters. Using Convolutional Neural Network (CNN) and Long Shot-Term Memory (LSTM) gained an accuracy of 98.3%. The performance evaluation is done by using Precision recall (PR), Confusion Matrices, and Receiver Operating Characteristics (ROC). Therefore, Model can help to manage and monitor the mental health of people exposed to stress. Which may have potential for development in various healthcare systems.

Paper 4 : Driver Fatigue Detection using LSTM [8]

In recent years, fatigue driving is the serious threat to traffic safety. Research on fatigue driving recognition has a great significance to improve traffic safety. To improve the accuracy and efficiency of detecting driver drowsiness multitask Convolutional Neural Network (MCNN) is used for detecting human face whereas, recognized facial images are passed as sequence to and Long Shot-Term Memory (LSTM) to gain feature value. As a result, this model gave 93% accuracy. They proposed a new, high-precision, real-time fatigue detection method based on the computer vision.

Project Goals

- To detect and report events of interest that are recorded by cameras installed on trucks during the process of garbage collection.
- To design an optimized ML algorithm that brings out better accuracy to the classification problem for activities captured under low image resolutions.

Objectives

- To conduct background studies on various waste collection processes, waste management systems, applied AI concepts in real world for waste management process improvements.
- To identify the potential ethical concerns related to the dataset
- To explore required images to the solve the business problem considering number of images, their structure & sequence.
- To build the machine learning algorithm that performs feature extraction, classification with training images in sequence and their testing.
- To evaluate the performance of the machine learning algorithm.
- To optimize the model with fine tuning.
- To create final project report and the presentation.

Metrics

- Literature review: Reviewing of 8 required Articles
- 5c’s : consent, clarity, consistency, control (and transparency), and consequences (and harm).
- Number of incidents, images, format, labels, image resolutions.
- Feature Extraction using Cov3D, Max-Pooling Layers , Padding, Dense and Dropout layers
- Model Performance Evaluation Metrics : Confusion matrix, Accuracy score, Recall score, Precision score, F1 score.
- Adaptive learning rate, optimization techniques(decay rate, learning rate, weight initiation, activation functions, batch size and number of epochs).
- Content(No. of pages), organized (References), Time spent and structured report.

Data Source

The dataset has been provided by client and stored in a secure cloud storage.

1. Json File for **speed of the vehicle** being collected using the GPS system of the vehicle.
2. The Incident data consists of **20,000 images with 3 folders.**

(Sequences of **True lifts/Good Incidents** , **False Lifts/Bad Incidents** and **No Incidents**)

3. Continuous Running Time dataset for 4 hours with **approximately 6,000 images**

Quality & Fitness

Images are found to be with uneven sequences. It makes sense as it is purely depending on the truck operators.

5 C's

We have built a product that uses data ethically and responsibly following all the 5 C's: consent, clarity, consistency, control (and transparency), and consequences (and harm)

Tools and Technology

Project Development Tools

Google Drive

jupyter

GitHub

Programming Tool

python™

Communication Tools

Outlook

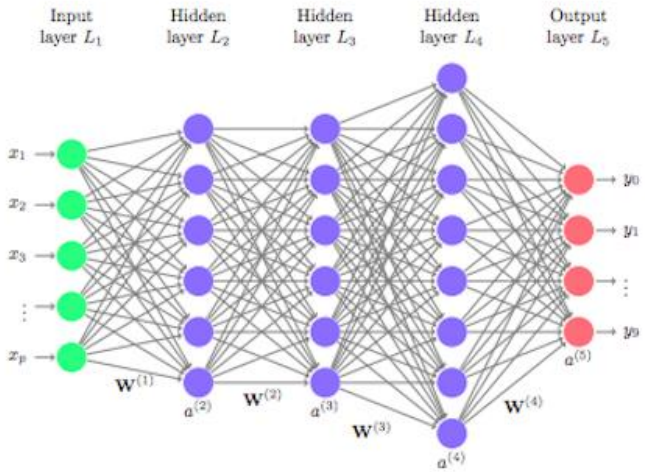
Microsoft Teams

Tools and Technology

Technological skills adapted

Deep leaning Architectures

CNN + LSTM



TensorFlow

Keras

Python Libraries

matplotlib

scikit learn

OpenCV

pandas

NumPy

Exploratory Data Analysis

Incidents are labeled

```
# Retreive the class names/class of incidents present in that directory using the specified path
labels = os.listdir(p)
print(labels)
```

```
/content
/content/Data Set/train
['Bad Incidents Data Set', 'No Activity Data Set', 'Good Incidents Data Set']
```

Convert into Categorical

- **Good Incidents: True Positives -1**
- **Bad incidents: True Negatives -0**
- **No incidents: True Negatives -2**

Good incidents : 174

Bad Incidents : 160

No Incidents : 136

Total : 470

```
Training Data Set
(376, 90, 112, 112, 3)
[1 1 0 1 1 0 1 1 0 1 1 1 1 0 0 1 1 1 1 0 1 0 1 1 2 1 0 1 1 1 0 0 1 0 0 1 0
 0 0 0 0 1 0 1 1 0 0 1 1 1 1 0 0 1 0 0 1 0 0 1 1 1 1 1 0 2 1 0 1 0 1 2 0 1
 0 0 0 0 1 1 0 1 1 0 1 0 0 2 0 1 1 1 0 0 0 2 1 1 0 0 0 1 1 1 1 1 1 0 1 1 0
 0 0 2 1 0 1 0 0 1 1 1 1 0 2 0 0 0 0 1 0 1 0 0 0 1 1 1 2 1 1 1 2 1 0 1 0 1
 1 0 0 0 1 0 0 0 0 1 1 2 1 1 0 2 2 1 1 0 0 0 0 1 0 1 1 1 2 2 0 0 0 0 1 1 1
 0 0 1 2 1 1 1 0 1 1 0 1 2 1 0 0 0 0 2 0 2 1 1 0 0 0 1 0 2 0 2 1 1 0 1 0 1
 1 0 0 1 1 1 2 1 1 0 1 1 0 1 2 1 0 1 0 2 0 0 2 0 1 0 0 0 0 0 2 0 0 1 2 0 0
 1 1 2 1 1 0 2 1 1 0 1 0 1 1 2 1 0 1 0 1 2 2 1 2 1 1 1 1 1 1 0 0 1 1 2 0 0
 1 0 1 1 1 1 0 1 0 0 0 0 0 0 1 1 1 0 1 0 1 0 0 0 0 1 1 1 1 1 2 1 0 1 0 1 0
 1 1 0 1 1 1 1 1 1 1 0 1 0 0 1 1 0 0 0 1 0 0 1 2 0 1 1 0 1 1 1 0 0 0 0 0 2
 1 0 1 1 0 0]
(376,)

Validation Data Set
(94, 90, 112, 112, 3)
[1 1 1 0 1 1 1 0 1 0 1 1 0 0 0 1 2 1 0 0 0 1 1 1 1 0 1 1 0 0 2 0 1 1 1 1 0
 0 0 2 0 1 1 2 1 0 1 2 1 1 0 1 1 1 1 1 0 2 1 0 2 1 2 1 1 1 2 1 1 0 0 1 1 1
 1 1 1 0 0 0 0 1 2 1 0 0 0 1 1 1 1 0 1 1]
(94,)
```

Exploratory Data Analysis

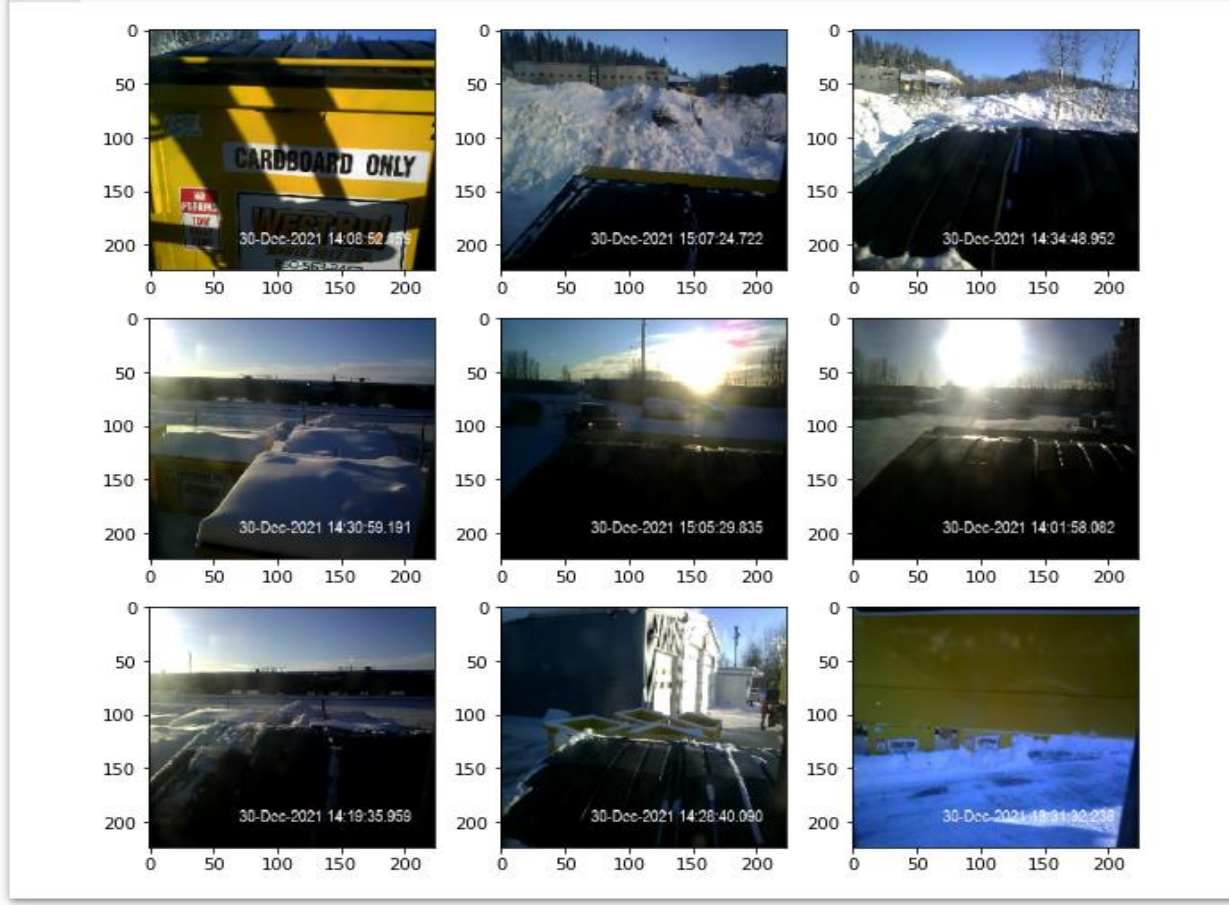
Image data representation is in a sequence as an incident

1.Convert image into 2D Array(matrix)
2. Convert sequences of images into list of array of incidents
3.Convert all incidents into 1 dimensional

Image Resize = 112*112
Number of incidents = 470
Sequence of images per incident = 90
Number channels = 3 (RGB)

Final input shape - (470,90,112*112,3)

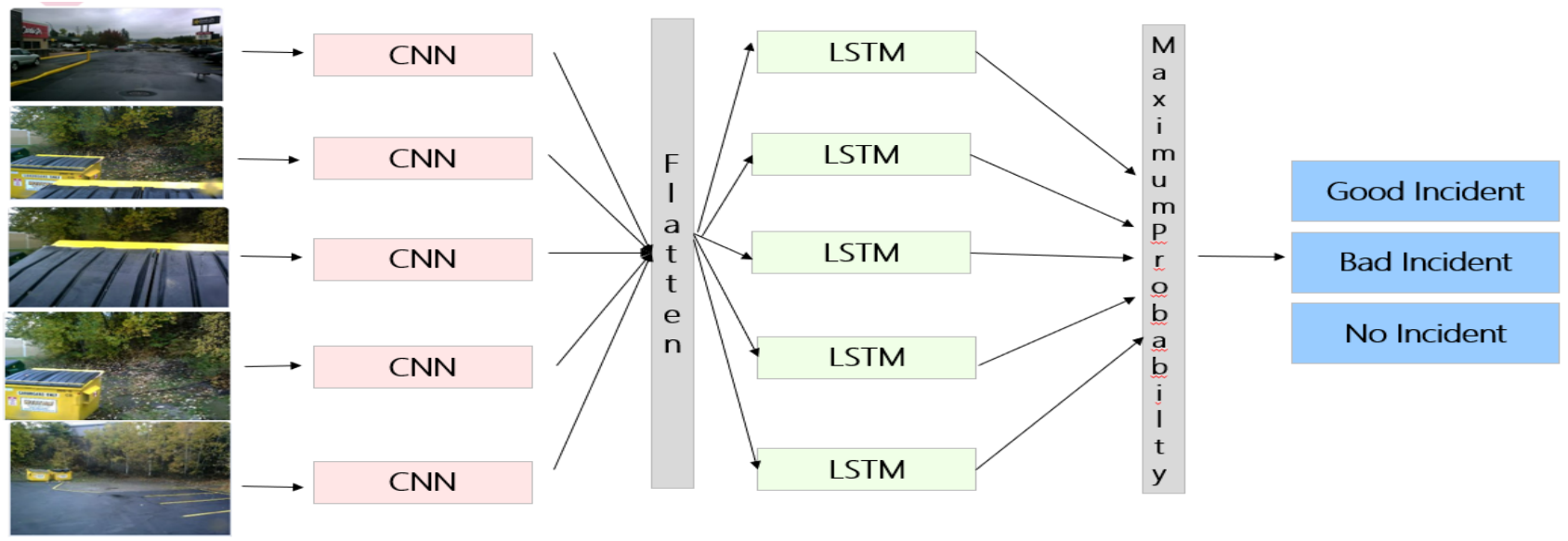
```
plt.figure(figsize=(10,10)) # specifying the overall grid size

for i in range(9):
    plt.subplot(3,3,i+1)    # the number of images in the grid is 5*5 (25)
    plt.imshow(incident_seq_images_data[i])

plt.show()
```



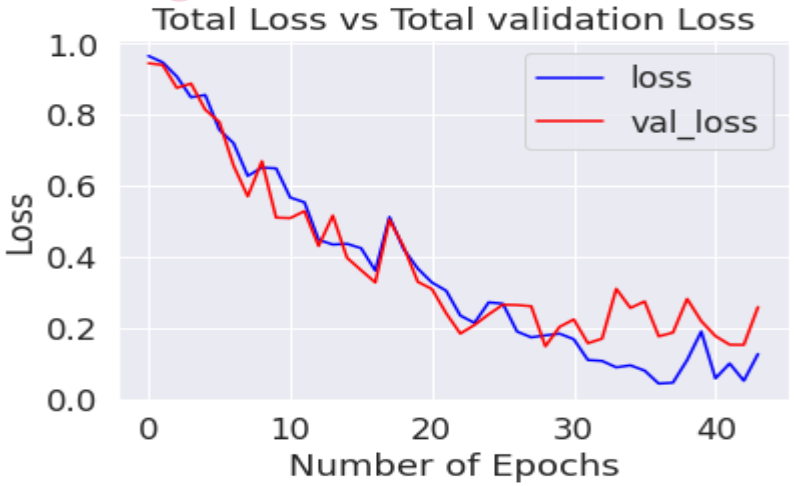
Deep Learning Architectures

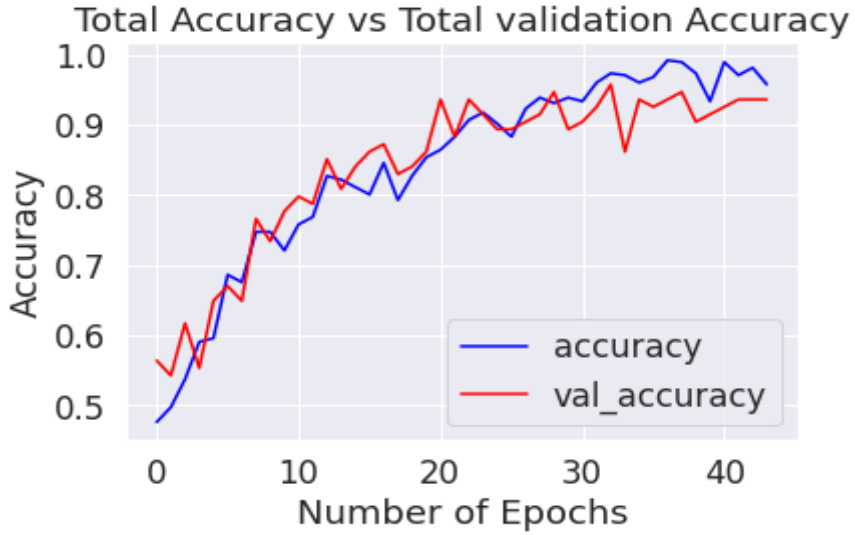
LRCN Model
(CNN with time distribution of sequence + LSTM)

- LRCN is a combination of CNN and LSTM
- The architecture enables to extract the Spatial information in the temporal structure using distributed layer- Spatiotemporal sequences
- These spatiotemporal sequences are passed to the LTSM model as in input to generate three label classification



Results - LRCN Approach






```
model_evaluation_history = LRCN_model.evaluate(features_test, labels_test)
```

```
1/1 [==============================] - 0s 403ms/step - loss: 0.4984 - accuracy: 0.8400
```

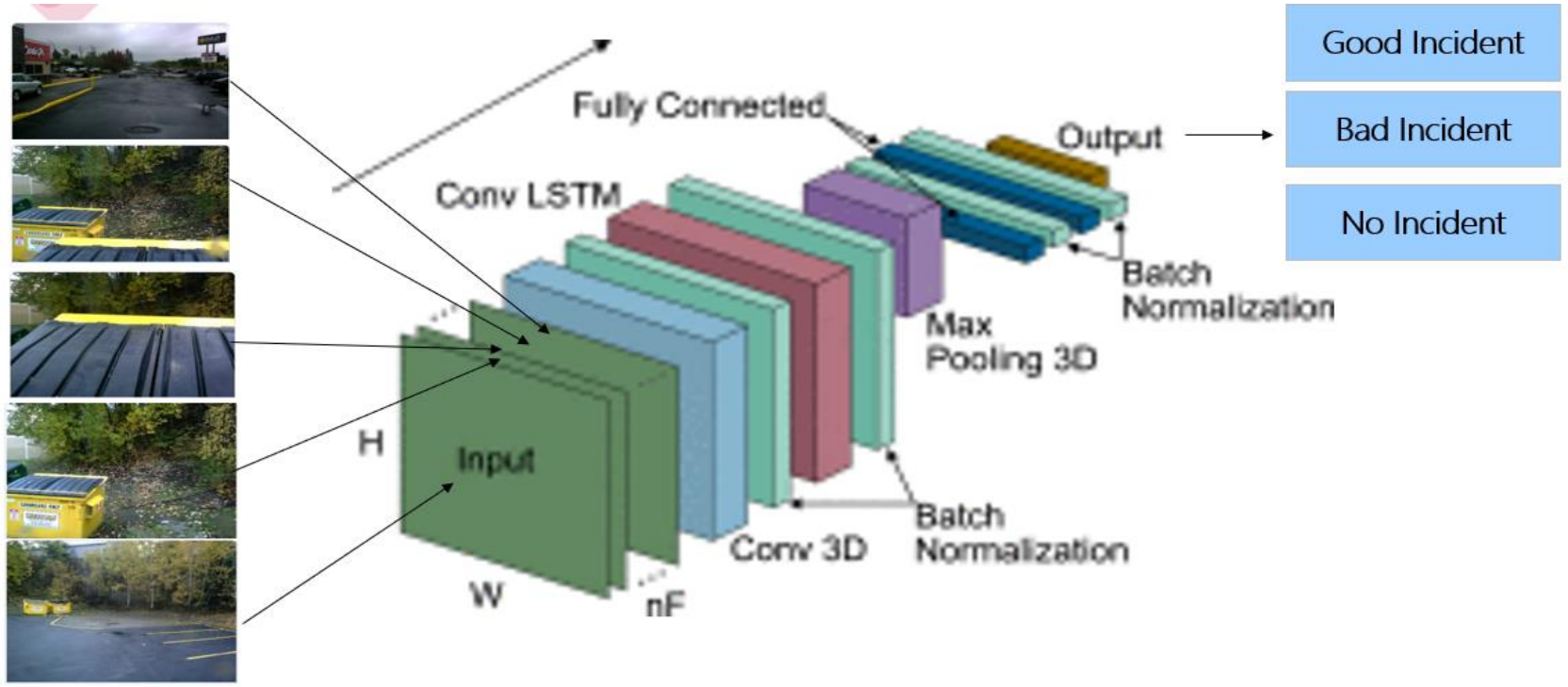
Classification Report

	precision	recall	f1-score	support
0	0.75	0.90	0.82	10
1	0.88	0.70	0.78	10
accuracy			0.80	20
macro avg	0.81	0.80	0.80	20
weighted avg	0.81	0.80	0.80	20

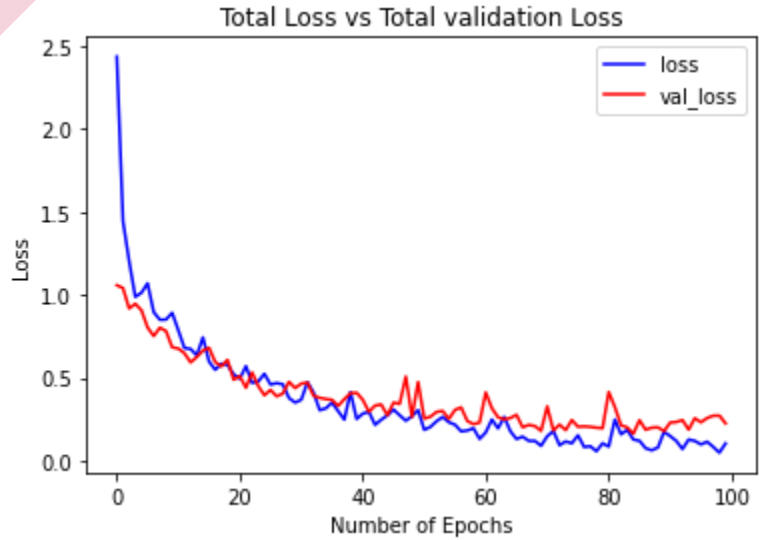
Deep learning Architectures

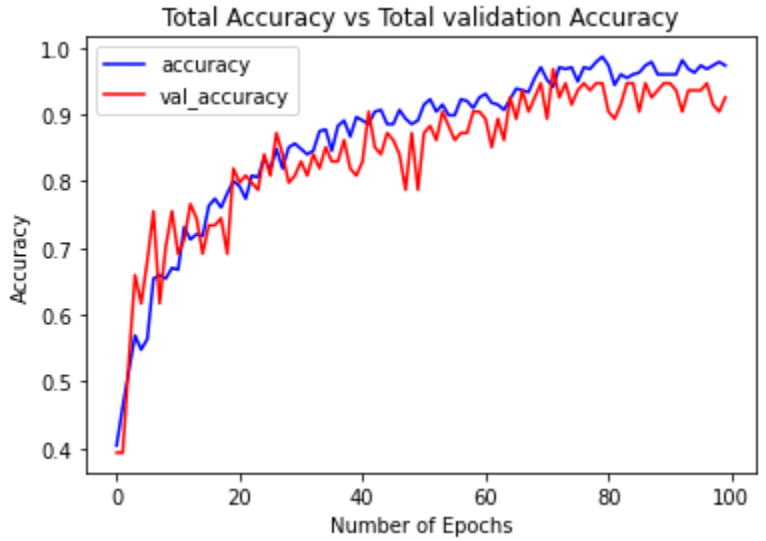
Conv3D Model

- Applies a 3D convolution over an input signal composed of several input planes.
- Softmax Activation Function for 3 label classification

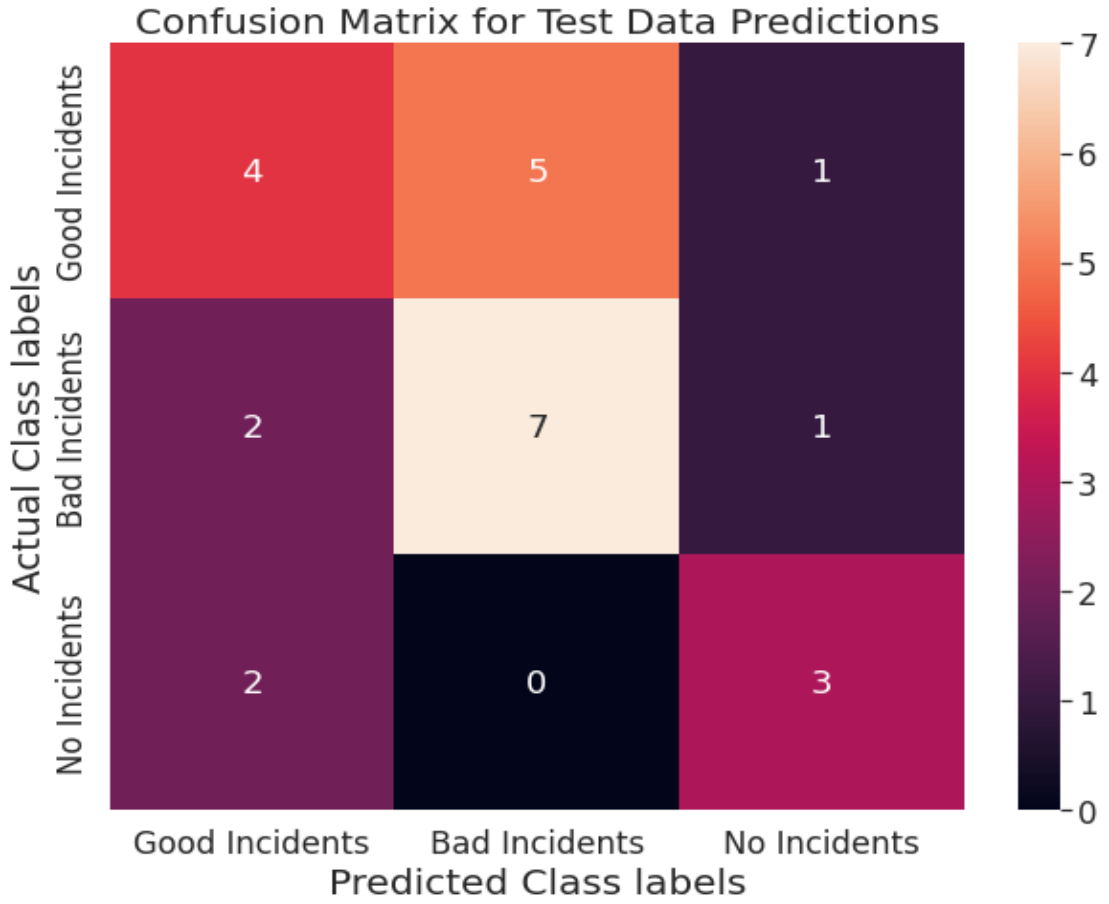


Results - CONV3D Model

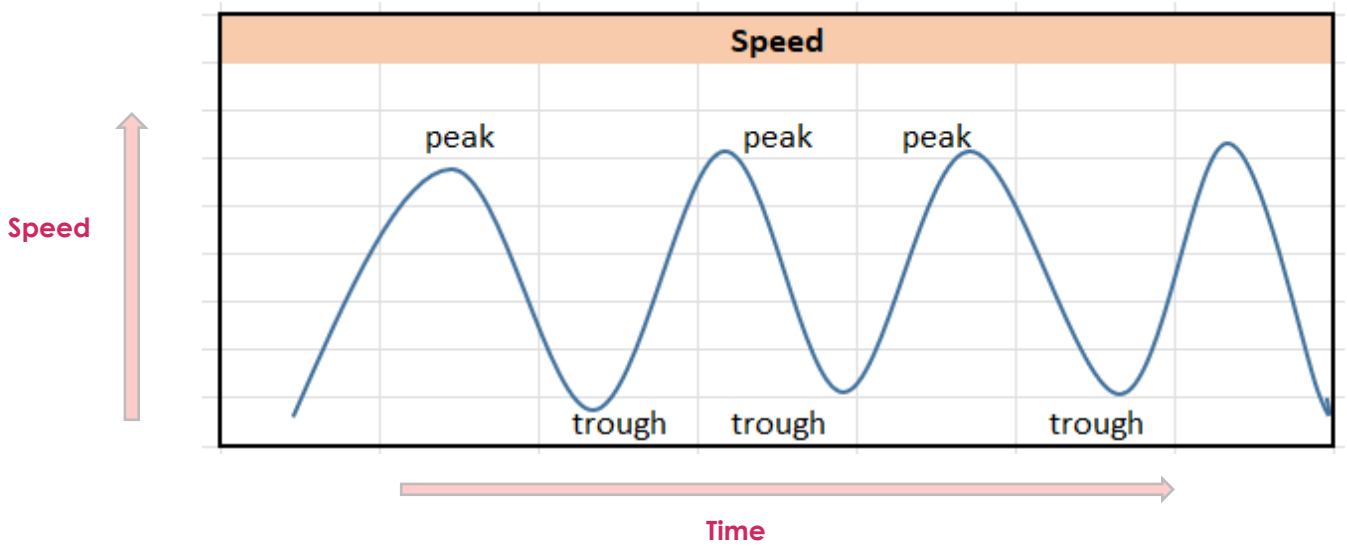




Test Accuracy Score : 0.80



Real-Time Incident Prediction



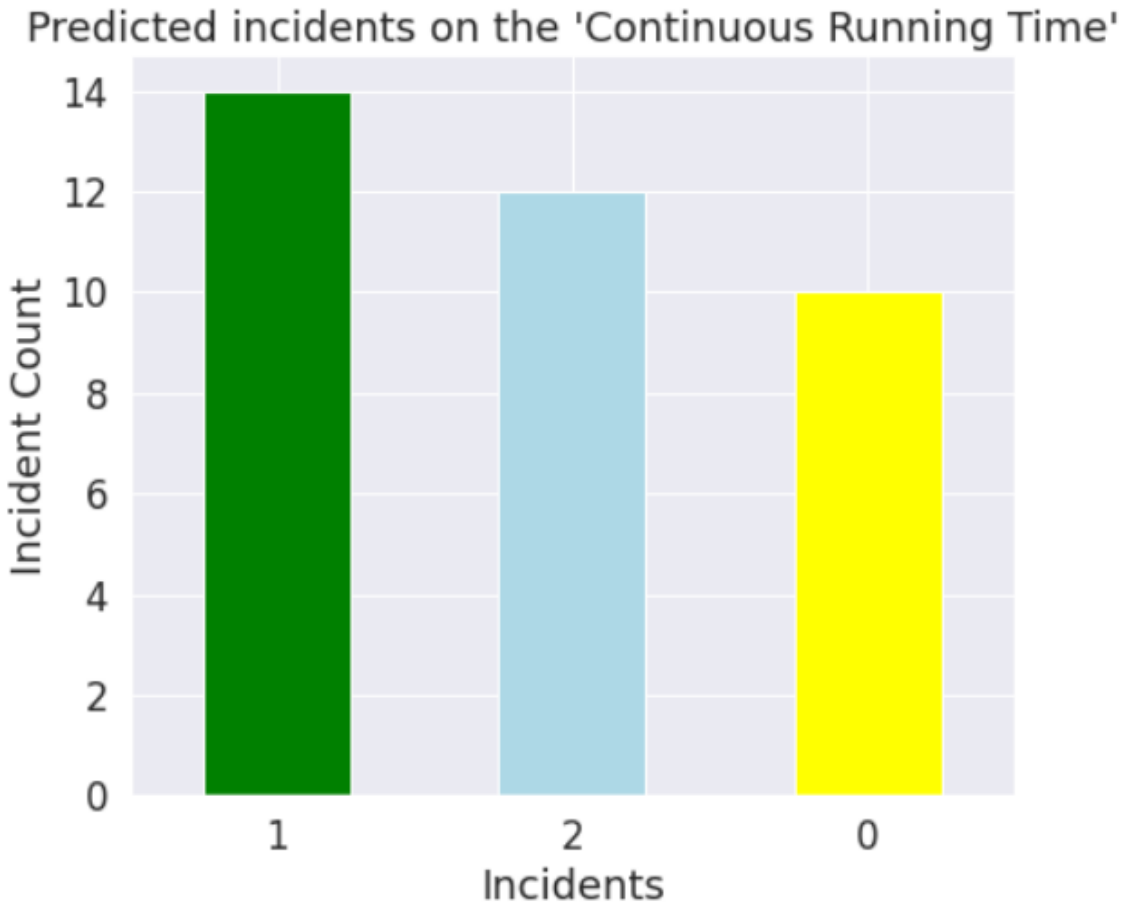
Time

Speed Cutoff at 8.00	Speed cutoff @ 6.00	Speed Cutoff at 5.00	Event
[72,	72,	68,	Incident
3,	3,	1,	
3,	3,	2,	
3,	1,	1,	
141,	137,	130,	Incident
1,	2,	1,	
2,	2,	2,	
2,	20,	19,	Traffic
20,	7,	7,	
8,	2,	2,	
1,	3,	2,	
2,	144,	143,	Incident
2,	4,	4,	
11,	2,	1,	
156,	2,	2,	
4,	9,	9,	
2,	3,	3,	
9,	98,	2,	
4,	67,	87,	Incident
98,	1,	1,	
72,	2,	55,	Incident
1,	5,	3,	
2,	3,	1,	
7,	4,	2,	

Real-Time Incidents

- Realtime Good Incidents -14
- Realtime Bad Incidents – 10
- Realtime No incidents – 12

No Incidents and Bad Incidents can be ignored as they are taken into consideration as we are interested in Good Incidents only



Conclusion

- Both the models are able to produce desirable results as expected.
- Test Accuracy Comparison of both models tells that LRCN Architecture works the best for our data

LRCN Model Test Accuracy : 92%

Conv3D Model Test Accuracy : 80%

- Model can be Deployed for the Real-Time Scenario and then it needs continuous training

- Building the model for all weather condition incidents which needs high level training with all condition sequence image data.
- Fine Tuning it further to increase the test accuracy from 92%.
- Use of bin detections tags/Geo locations ID to recognize the start and stop of the activity for better performances.
- Deployment of the model to the real-time set up

References

[1] A. -M. R. Abdali and R. F. Al-Tuma, "Robust Real-Time Violence Detection in Video Using CNN And LSTM," *2019 2nd Scientific Conference of Computer Sciences (SCCS)*, 2019, pp. 104-108, doi: 10.1109/SCCS.2019.8852616.

[2] Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., & Woo, W. C. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, *28*.

[3] Xia, K., Huang, J., & Wang, H. (2020). LSTM-CNN architecture for human activity recognition. IEEE Access: Practical Innovations, Open Solutions, 8, 56855–56866. https://doi.org/10.1109/access.2020.2982225

[4] Anh Khoa, T., Phuc, C. H., Lam, P. D., Nhu, L. M. B., Trong, N. M., Phuong, N. T. H., Van Dung, N., Tan-Y, N., Nguyen, H. N., & Duc, D. N. M. (2020). Waste management system using IoT-based machine learning in university. Wireless Communications and Mobile Computing, 2020, 1–13. https://doi.org/10.1155/2020/6138637

[5] Semwal, V. B., Gupta, A., & Lalwani, P. (2021). An optimized hybrid deep learning model using ensemble learning approach for human walking activities recognition. The Journal of Supercomputing, 77(11), 12256-12279.

[6] Shang, S., Luo, Q., Zhao, J., Xue, R., Sun, W., & Bao, N. (2021, April). LSTM-CNN network for human activity recognition using WiFi CSI data. In Journal of Physics: Conference Series (Vol. 1883, No. 1, p. 012139). IOP Publishing.

References

[7] Mingu Kang, Siho Shin, Jaehyo Jung, Youn Tae Kim, "Classification of Mental Stress Using CNN-LSTM Algorithms with Electrocardiogram Signals", Journal of Healthcare Engineering, vol. 2021, Article ID 9951905, 11 pages, 2021.

[8] Long Chen, Guojiang Xin, Yuling Liu, Junwei Huang, "Driver Fatigue Detection Based on Facial Key Points and LSTM", Security and Communication Networks, vol. 2021, Article ID 5383573, 9 pages, 2021.

[9] Brownlee, J. (2019, August 14). CNN Long Short-Term Memory Networks. Machine Learning Mastery. https://machinelearningmastery.com/cnn-long-short-term-memory-networks/

[10] Golubev, A. (2021, December 22). Long-term Recurrent Convolutional Network for Video Regression. Medium. https://towardsdatascience.com/long-term-recurrent-convolutional-network-for-video-regression-12138f8b4713 Golubev, A. (2021, December 22). Long-term Recurrent Convolutional Network for Video Regression. Medium. https://towardsdatascience.com/long-term-recurrent-convolutional-network-for-video-regression-12138f8b4713

[11] Chan, M. (2021, December 14). Pytorch: Step by Step implementation 3D Convolution Neural Network. Medium. https://towardsdatascience.com/pytorch-step-by-step-implementation-3d-convolution-neural-network-8bf38c70e8b3


thank
you!


Think Green,
Think Clean.