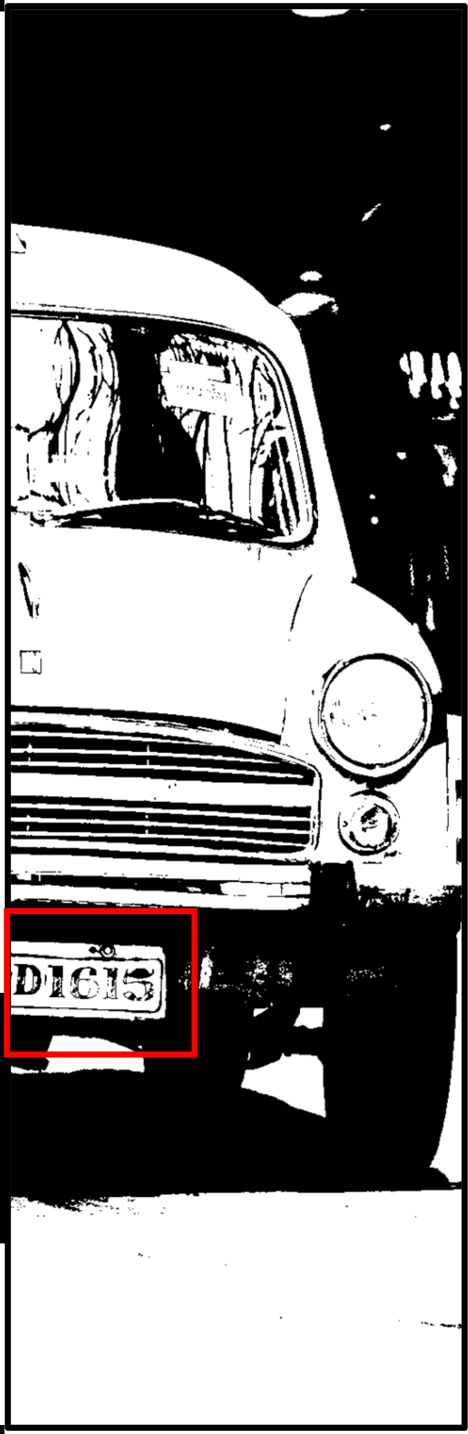
TATA CONSULTANCY SERVICES

**HumAIn 2019**

# VEHICLE LICENSE PLATE DETECTION

*Object detection and text extraction using machine intelligence.*



# BACKGROUND

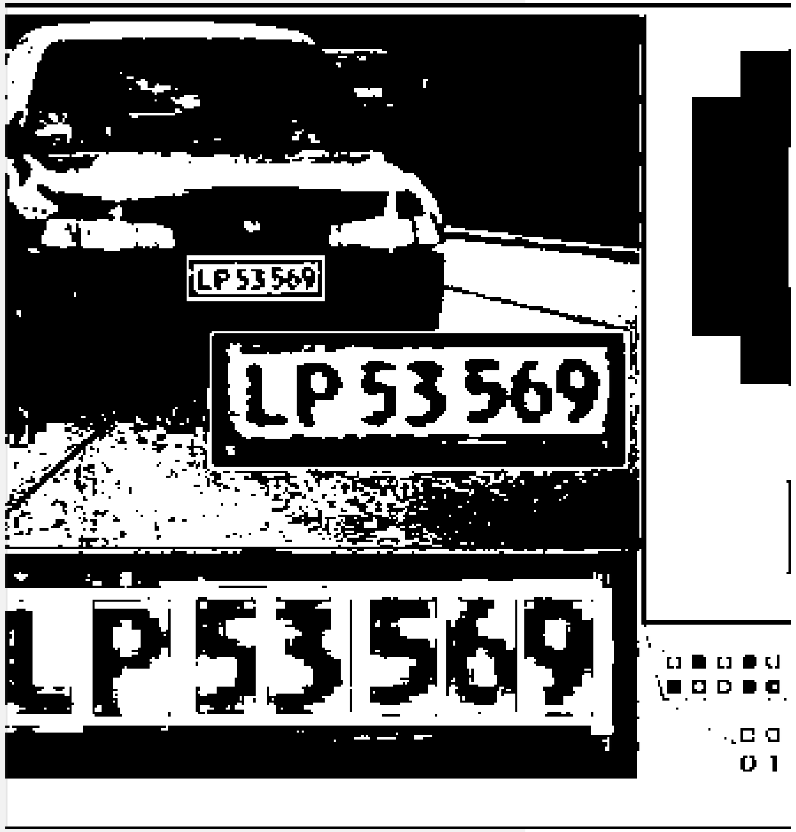
- *Vehicle number plate detection and recognition plays an important in the transportation management system.*
- *An accurate license extraction can enable smooth traffic law enforcement, criminal detection and flow management to enable automation of finances in the traffic space like toll collection.*
- *The problem presents itself in two parts:*
    - The license plate region detection in the image.
    - Extraction of text from the plate.

MH15-TC-553

# UNDERSTANDING THE PROBLEM

- A practical view of the problem :
  - The webcams and CCTVs used to capture the images in the real scenario are low spec. Using very high resolution cameras can cost a heavy amount to the state's budget.
  - The capture may occur when the vehicle is in motion causing a slight blur.

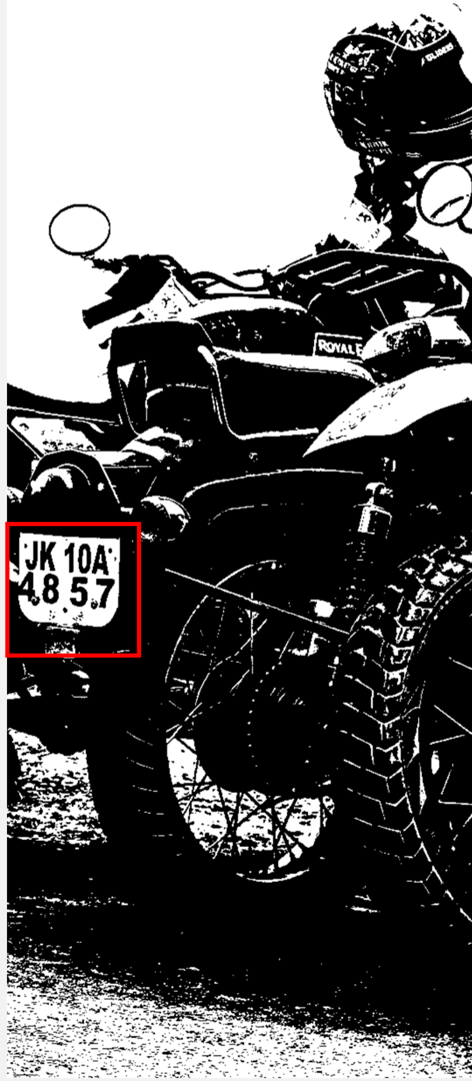
- The problem incorporates the barrier of detecting license plate from a media lacking high specs and then performing OCR on the detection to obtain the license plate number.



# SCOPE OF THE SOLUTION

*Elements being considered for the sake of production ready model development*

- It is being considered that the model is to be presented to suit practical need of the people.
- The dataset provided therefore, was operated upon to make it better fitting to practical scenario.
- Since the dataset has only images with a single vehicle as a time. The product would mainly deliver to OCR a single license plate in a given scenario. Results over dataset with multiple vehicles maybe unjustified.
- Out of scope:
  - The solution is not being developed for in-video detection over extremely hazy and greyscale camera.
  - Multiple detection may be included with a different dataset training. Currently, the accurate detection is being considered out of scope.



# ASSUMPTIONS

*These are made keeping in mind the possibility of practical use.*

- The hardware which would be used to detect the images for the model to work on would be an affordable low cost camera (eg. Webcam, CCTV camera) which would record between 360-480 pixels resolution images.
- The usage of pre-trained models is allowed for the purpose of speeding up the training procedure with simple weight tweaking or transfer learn.
- One is allowed to incorporate techniques that may not have anything to do with the problem statement directly but are essential to tamper the input for better accuracy and results.
- Knowledge about the limitations of alphanumeric used in a regular license plates maybe used to create models.

# SOLUTION APPROACH

*A Broad Overview*

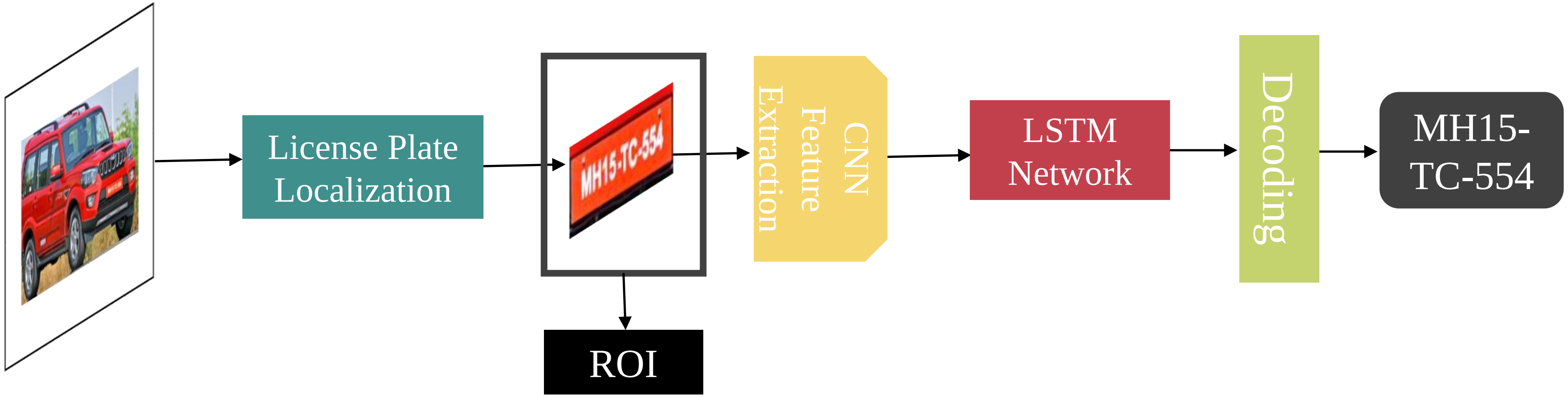
<u>The problem will be tackled in two parts:</u>

Object Localization :
At first the license plate needs to be detected in the image provided. This task will be done using YOLOv3 object detection model [1]. The model will be retrained over the original weights to detect the license plate in the image. With the region of interest extracted, the model output would then be passed on for text extraction.

Text Extraction:
The input license plate image would then be passed through a CNN for extracting image features. The feature map would then be reshaped to pass through an LSTM network (10-11 parallel LSTMs to predict the license characters). The output of each LSTM would be passed to a activation function to obtain a probability distribution over the library of characters that are used in license plates. This output of type, hot encoded, will then be decoded to give the license plate number.[2]

# SOLUTION APPROACH

*A Broad Overview*




The solution discussed has been proposed after analysing:

1. YOLOv3 is a fast object detection module. The trade-off occurs against accuracy. Speed here is more important as for production use the results would be required instantaneously. The accuracy of bound detected may lag as the region of interest extract would have character contents of the license place only even if the detection area is a little more than required. [A1]
2. This method of text extraction is useful but bounds the number of characters that can be extracted. However, for our use the limitation is not an issue as the license plate as limited number of characters only.

# SOLUTION APPROACH

*Models and algorithms*

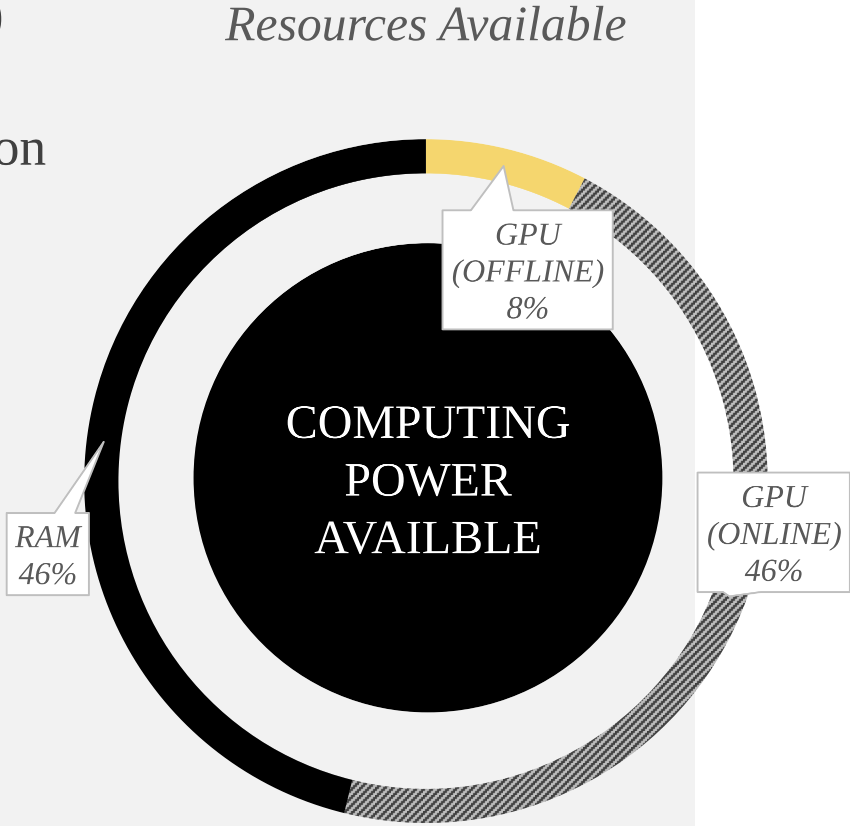
Object Detection Task:
The training would be done over a dataset which has images from both clean (as provided) and low resolution (as fabricated). However the test and validation set would contain low resolution images in slight majority to improve the models standing for production use [4]. K-fold cross validation scheme shall be applied keeping the limited size of dataset in view and the scoring shall be done using mean average precision (mAP).

Text Extraction Task:
The task shall be performed using parallel LSTM networks, followed by fully connected layers with softmax activation function [A2].The predictions would come from a library of one-hot encoded alphanumeric that appear in licenses. Further the encoded predictions would be decoded using a custom algorithm and result would be output. For this task we shall work with CTC (Connectionist Temporal Classification) Loss to deal with unsegmented license text.[2]

# IMPLEMENTATION FRAMEWORK

- The project was undertaken on a personal computer with intel - i5 core (7th generation) and NVidia GEFORCE 940 MX graphics (2 gb). Google Colab would be used in case of extra GPU power requirement (12 gb limit). The solution would be tested using a 13MP mobile camera capture.

- Softwares and libraries used are as follows:
  - Visual Studio Code (1.37.1)
  - Anaconda3 Distribution with python 3.6 support
  - Jupyter notebook (Visual analysis)
  - OpenCV (3.3.1)
  - Keras (Tensor flow Backend)
  - Pandas (0.23.4)
  - Git Bash



# SOLUTION SUBMISSION

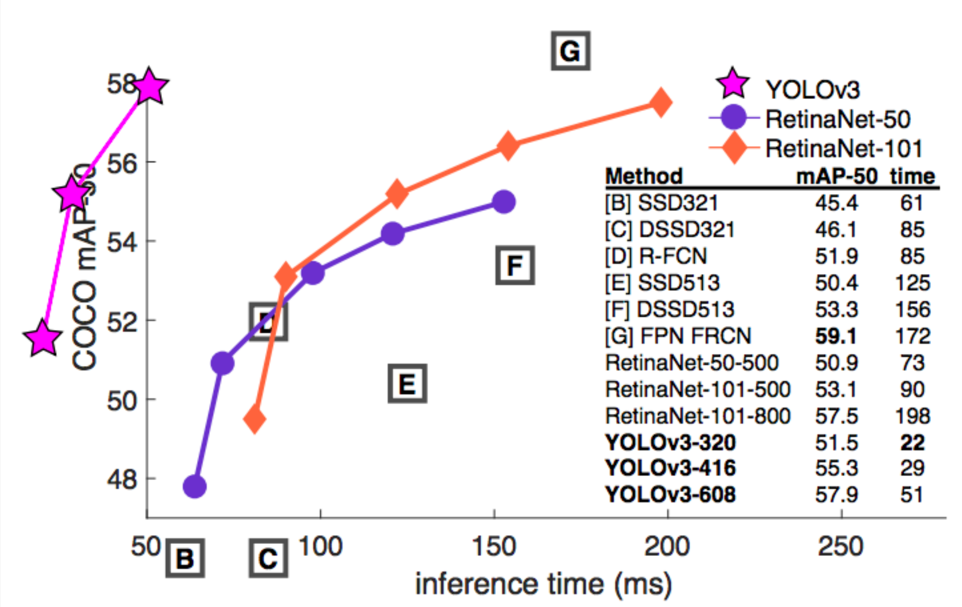
*Exploring the repository*

- The Repository can be divided in the following segments:
    - Data : Consists of code to obtain the data and perform augmentation over it. The code available also takes care of the annotation scaling according to augmentation performed for single vehicle images.
    - Main : Contains the model file and the implementation pseudo code pieces for achieving the complete task.
    - Notebook: Would contain notebooks with visuals of the implementation and dataset.
    - Presentation: Contains the explanation of the solution being proposed (This file itself).

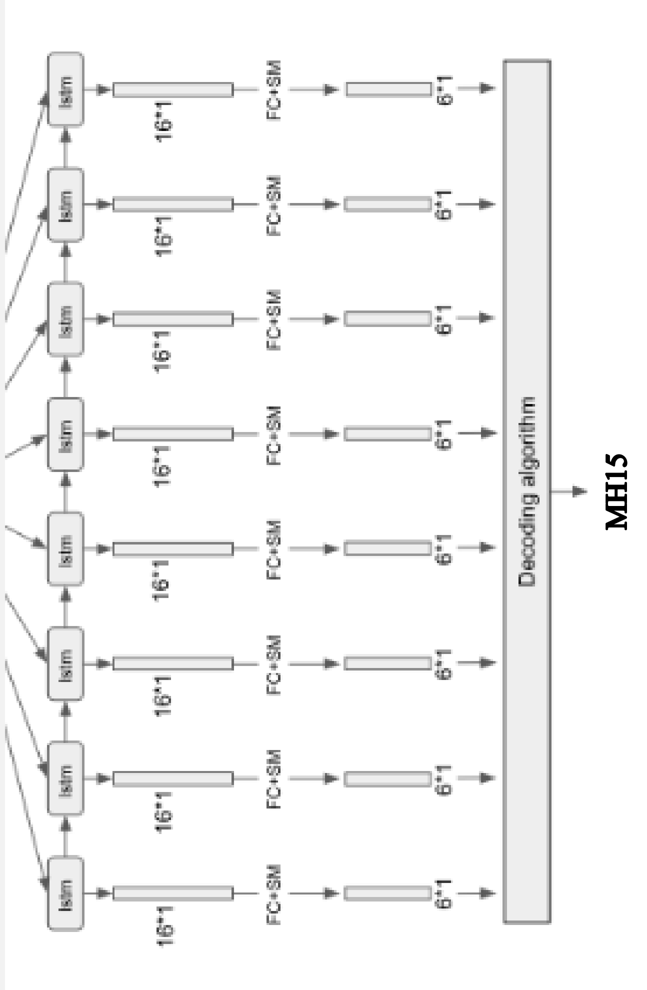
- Repository link: https://github.com/hrshtsharma17/VehiclePlateOCR

# APPENDIX

*For more information regarding the topics referred to in the solution. One can follow the following :*




Why YOLOv3 [A1]




LSTM Net Example [A2]

# REFERENCES

*The following literature was referred. Not all contribute to the solution but where essential to develop choices between approaches.*

[1] Redmon J, Farhadi A. Yolov3: An incremental improvement[J]. arXiv preprint arXiv:1804.02767, 2018.

[2] Z. Shi, M. Shi and C. Li, "The prediction of character based on recurrent neural network language model," in 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS), Wuhan, 2017.

[3]Alex Graves, Santiago Fernandez, Faustino Gomez and Jürgen Schmidhuber. Connectionist Temporal Classification: Labelling Unsegmented Sequence Data with Recurrent Neural Networks. ICML – 2006

[4] Machine learning yearning - Andrew Ng (developing practical AI products)

[5] M. B. Blaschko and C. H. Lampert. Learning to localize objects with structured output regression. In Computer Vision– ECCV 2008, pages 2–15. Springer, 2008

[6]C. L. Zitnick and P. Dollar. Edge boxes: Locating object proposals from edges. In Computer Vision–ECCV 2014, pages 391–405. Springer, 2014