IJCRT.ORG ISSN : 2320-2882
INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

# DEVELOP A SYSTEM FOR ELECTRIC VEHICLE RECOGNITION USING MACHINE LEARNING

*Implementation and Results*

**[1]Riya Singh, [2]Prarthna P Chhabria, [3]Mrs. Amitha S**

[1]Student, [2]Student, [3]Asst. Professor
[1]Department of Computer Science and Engineering,
[1]K.S. School of Engineering and Management, Bangalore, India

***Abstract:*** The future of the automobile industry is electric vehicles which are advantageous to us in many ways. Electric vehicles do not emit pollutants which make them eco-friendly and are generally differentiated with the help of a green license plate. The proposed system aims at building an intelligent model that can recognize electric vehicles such as cars, buses, autos and two-wheelers from its green license plate. The technique used here is image processing which works by processing the images of vehicle license plates. The final objective is to create an automatic authorized electric vehicle recognition machine. The data can be used by the government to conduct surveys and monitor the growth of electric vehicles in society. The data can also be further analyzed by environmentalists to understand the level of impact the electric vehicles have on the environment.

***Index Terms*** **- Electric Vehicles, Recognition, Machine Learning, OpenCV, Color detection, License plate detection.**

## I. INTRODUCTION

Electric vehicles run partially or fully on electricity. These vehicles produce less pollution when compared to vehicles that run on petrol and diesel. Due to this, the number of electric vehicles is increasing day by day.

Since the shift to electric vehicles is recent, the current license plate detection systems do not detect the green license plates of these vehicles. Therefore, there is a need for a system that detects the license plates of electric vehicles.

The systems that are existing today cannot distinguish between conventional vehicles and electric vehicles. The proposed system aims at eradicating this by adding a feature that allows the system to distinguish between the vehicles.

Keeping a track of the increasing number of vehicles can be complicated and difficult. With this in mind, identification of vehicles that cause violations such as breaking the signals, accidents and so on also becomes difficult. The proposed system can be used to help eliminate these problems and make the process of identification easier.

Since electric vehicles cause little to no pollution, the data from the proposed system can be further analyzed by environmentalists to help them understand the levels of pollution on the road. Thereby helping them formulate environmental plans accordingly.

## II. RELATED WORK

Recognition of license plates have become an important part of ensuring everyday safety of the vehicles and people on the road, as well as making sure to identify those causing problems by breaking traffic rules on the vehicle. The proposed project takes inspiration from this system to be able to recognize the license plates.

Apart from the license plate recognition, the project also uses the ideas behind a color recognition feature with the help of computer vision using OpenCV libraries. This is used to help identify the green color that is present in the license plates of electric vehicles.

## III. THE PROPOSED METHOD

The proposed system has five objectives to accomplish. They are:

1. To collect data and create a functional dataset.
2. To pre-process the data frame and perform exploratory data analysis.
3. To scale the data and fit it into the model.
4. To include optical character recognition for detection of license plate number.
5. To add color recognition to the model for detection of green license plate.

### 3.1 To collect data and create a functional dataset

Collecting data and creating a functional dataset involves several steps, including identifying the data sources, gathering the data, cleaning and pre-processing the data, and organizing the data into a usable format.

The first step in collecting data is to identify the sources of data. This may include public databases, surveys, or other sources of information that are relevant to the problem being studied. Once the sources of data are identified, the data must be gathered. This may involve downloading data from a website or extracting data from a database.

After the data is gathered, it may need to be cleaned and pre-processed to remove any errors, inconsistencies, or missing values. This can involve standardizing data formats, removing duplicates, and filling in missing values with estimates or averages.

The next step is to organize the data into a usable format. This may involve creating a spreadsheet or database to store the data, or using a specialized software tool to manage the data. The data must also be properly labeled and categorized to make it easy to analyze and interpret.

Once the data is organized, it can be used to create a functional dataset. This involves selecting the appropriate features or variables that are relevant to the problem being studied, and transforming the data into a format that can be used for analysis. This may involve normalization, feature selection, or feature engineering to create new features that are more informative.

Finally, the functional dataset can be used to train machine learning models or perform statistical analyses to extract insights and make predictions. The accuracy and reliability of the results will depend on the quality of the data and the effectiveness of the pre-processing and feature engineering techniques used.

### 3.2 To pre-process the data frame and perform exploratory data analysis

Pre-processing data and performing exploratory data analysis (EDA) are two important steps in the data analysis process. These steps are essential for identifying patterns, trends, and relationships in the data, and for preparing the data for further analysis.

Pre-processing the data involves several tasks, including cleaning the data, transforming the data, and reducing the data. Cleaning the data involves removing or correcting any errors, inconsistencies, or missing values in the data. This can involve using statistical techniques to impute missing values or removing outliers that may be affecting the analysis. Transforming the data involves converting the data into a more suitable format for analysis. This can include standardizing variables or transforming variables into more meaningful units. Reducing the data involves selecting a subset of the data that is relevant to the analysis.

Exploratory data analysis involves using statistical techniques and visualization tools to gain insights into the data. This includes identifying patterns, trends, and relationships in the data. Common EDA techniques include computing summary statistics, such as mean and standard deviation, and creating visualizations, such as histograms, scatterplots, and box plots. These techniques can help identify outliers, skewness, and other features of the data that may impact the analysis. EDA can also help identify potential variables to include in the analysis and suggest appropriate data transformations.

By pre-processing the data and performing exploratory data analysis, analysts can gain a better understanding of the data and identify potential issues that need to be addressed before proceeding with further analysis. This can help ensure that the data is appropriate for the intended analysis and improve the accuracy and reliability of the results.

### 3.3 To scale the data and fit it into the model

Scaling data and fitting it into a model are important steps in the data analysis process, particularly when using machine learning or statistical models to make predictions or classifications based on the data.

Scaling the data involves transforming the data so that all variables have the same scale or unit of measurement. This is important because many machine learning algorithms and statistical models rely on the assumption that all variables are on the same scale. If the variables are not on the same scale, some variables may dominate the analysis, leading to biased or inaccurate results. Scaling can be done using various techniques, such as min-max scaling, standardization, or normalization.

After scaling the data, the next step is to fit it into a model. This involves selecting an appropriate model and training it on the scaled data. The choice of model will depend on the type of analysis being performed and the characteristics of the data. For example, linear regression models are often used to predict a continuous outcome variable, while decision trees or random forests may be used for classification problems.

Fitting the data into a model involves estimating the model parameters or coefficients that best describe the relationship between the input variables (predictors) and the output variable (target). This is done by minimizing a cost function or maximizing a likelihood function using an optimization algorithm. The model is then evaluated on a separate dataset (validation set) to assess its performance and determine if it is suitable for making predictions or classifications on new data.

Overall, scaling the data and fitting it into a model are critical steps in the data analysis process, as they ensure that the data is properly prepared for analysis and that the model is trained on data that is appropriate for the intended analysis.

### 3.4 To include optical character recognition for detection of license plate number

Including a detection algorithm for license plate detection involves using computer vision techniques to locate and extract license plate regions from images or videos. The goal is to develop an algorithm that can accurately and efficiently detect license plates in a variety of scenarios, including varying lighting conditions, different plate sizes and shapes, and occlusions.

The detection algorithm typically involves several steps. The first step is to preprocess the image or video to enhance the license plate regions and remove noise or other irrelevant information. This can involve applying filters, adjusting the brightness and contrast, or using image segmentation techniques to separate the license plate from the background.

Once the image is preprocessed, the next step is to identify candidate license plate regions. This is often done using edge detection algorithms, which can identify the edges of objects in the image or video. The candidate regions are then further processed to identify the most likely license plate regions based on various criteria, such as size, aspect ratio, and color.

After the license plate regions are identified, the algorithm can perform character recognition to read the license plate number. This can be done using optical character recognition (OCR) algorithms, which can identify the individual characters in the license plate region and recognize them based on a database of known characters.

The accuracy and efficiency of the license plate detection algorithm depends on the quality of the pre-processing, edge detection, and character recognition techniques used. It may also depend on the training data used to develop the algorithm and the performance evaluation metrics used to assess its effectiveness. Overall, a well-designed license plate detection algorithm can play a crucial role in applications such as traffic management, law enforcement, and parking management.

### 3.5 To add color recognition to the model for detection of green license plate

To add color recognition to the model for detection of green license plate involves incorporating color-based features into the license plate detection algorithm. The goal is to enhance the accuracy and reliability of the algorithm in detecting green license plates specifically, as green license plates are often used to indicate electric or low-emission vehicles.

Color recognition can be achieved using various techniques, such as color filtering or histogram analysis. Color filtering involves setting a threshold for each color channel (red, green, and blue) and only retaining pixels that fall within the specified range. This can be used to isolate green pixels and remove other colors from the image or video. Histogram analysis involves analyzing the distribution of color values in the image or video and identifying peaks or clusters corresponding to specific colors, such as green.

Once the color-based features are identified, they can be incorporated into the license plate detection algorithm. This can be done by assigning weights or probabilities to the color-based features based on their importance or relevance to the detection of green license plates. For example, the color-based features may be given a higher weight if they strongly indicate the presence of a green license plate.

The addition of color recognition to the license plate detection algorithm can improve its accuracy and reliability in detecting green license plates, especially in scenarios where other factors, such as lighting or occlusion, may affect the detection of the license plate. This can be particularly important in applications such as traffic management or parking management, where the identification of electric or low-emission vehicles is of interest.

## IV. Workflow Of The System

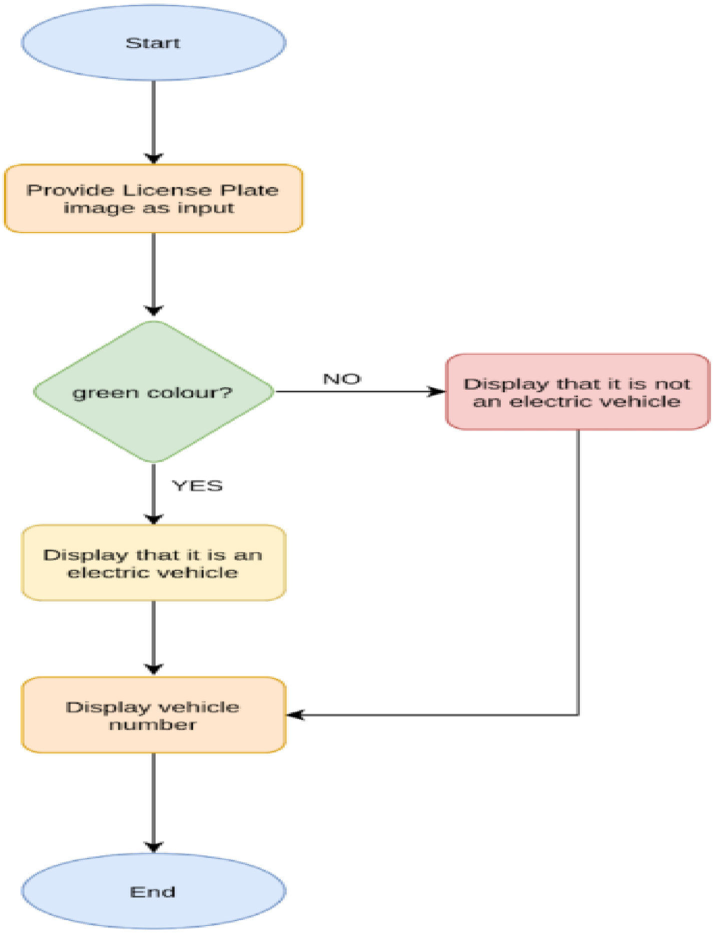


Fig 1. Steps in detection of electric vehicle with the license plate

The fig 1 above shows the flow of our system. Firstly, the application starts by providing the image of a license plate as input. Then the model checks if the image is green or not to identify whether the given vehicle is electric or not. If it is green, then the system displays an output saying the vehicle is electric. If it is not green, an output displaying that the given vehicle is not electric is generated. After this step, the model goes on to read the character segments from the license plate and generate the license plate number.

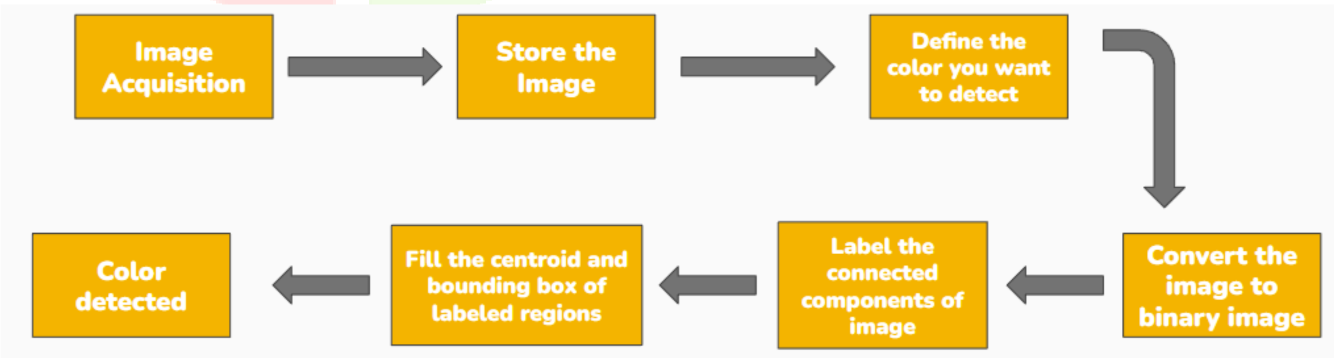


Fig 2. Steps in color detection

The fig 2 shows the steps that are involved in the detection of a color in an image. Upon providing the image as an input, the image is first loaded in a readable format in a variable. The range of color can be then defined to be used for detection. The image is then converted to a binary image and the color is then detected by searching for the nearest RGB values.

## V. Methodology

The methodologies used to accomplish the previously mentioned objectives are as given below.

### 5.1 Methodology for Objective 1

The methodology for collecting data and creating a functional dataset for the detection of electric vehicles with green license plates may involve the following steps:

1. Identify the data sources: The first step is to identify the sources of data that can be used to create the dataset. This may include publicly available datasets, data from government agencies or transportation authorities, or data collected from sensors or cameras installed at various locations.

2. Define the data variables: Once the data sources have been identified, the next step is to define the variables that will be collected and used in the analysis. This may include variables such as the license plate number, vehicle make and model, vehicle type (electric, hybrid, or gas-powered), date and time of detection, and location of the detection.

3. Collect the data: The data can be collected using various methods, such as manual data entry, automated data collection using sensors or cameras, or web scraping from publicly available sources.

4. Preprocess the data: The collected data may need to be preprocessed to ensure that it is in a usable format. This may involve cleaning the data to remove any errors or inconsistencies, transforming the data into a common format, and handling missing or incomplete data.

5. Label the data: In order to train a machine learning model for the detection of green license plates, the data must be labeled to indicate which vehicles have green license plates. This can be done manually or using automated methods, such as color recognition algorithms.

6. Create the functional dataset: Once the data has been preprocessed and labeled, it can be organized into a functional dataset that can be used for analysis and model training. This may involve merging data from different sources, selecting relevant variables, and splitting the data into training, validation, and testing sets.

7. Document the dataset: It is important to document the dataset by providing a detailed description of the data sources, variables, preprocessing steps, and labeling methods used. This documentation can help ensure that the dataset is reproducible and can be used by other researchers or practitioners in the future.

Overall, the methodology for collecting data and creating a functional dataset for the detection of electric vehicles with green license plates requires careful planning and attention to detail to ensure that the dataset is accurate, reliable, and suitable for the intended analysis.

### 5.2 Methodology for Objective 2

The methodology for preprocessing the data frame and performing exploratory data analysis for the detection of electric vehicles with green license plates may involve the following steps:

1. Remove duplicates and irrelevant data: The first step is to remove any duplicate records and data that is not relevant to the analysis. This may include removing records with missing or incomplete data or removing columns that are not needed for the analysis.

2. Handle missing values: If there are missing values in the data, they need to be handled appropriately. This may involve imputing missing values with a reasonable estimate or removing records with missing values.

3. Data transformation: The data may need to be transformed to prepare it for analysis. This may include normalizing the data, scaling it, or transforming it using methods such as logarithmic or exponential transformation.

4. Exploratory data analysis (EDA): EDA is a critical step in understanding the data and identifying patterns or trends that may be useful for analysis. This may involve visualizations such as histograms, scatter plots, or box plots to identify data distributions, outliers, or correlations between variables.

5. Feature engineering: Feature engineering involves creating new variables or features from the existing data to improve the accuracy of the model. This may include creating new variables based on time or location data, or combining variables to create new features that may be more useful for analysis.

6. Data sampling: Depending on the size of the data set, it may be necessary to sample the data to reduce computational time or to balance the classes if there is a class imbalance.

7. Document the pre-processing steps: It is important to document the pre-processing steps taken to prepare the data for analysis. This documentation can help ensure that the analysis is reproducible and can be used by other researchers or practitioners in the future.

Overall, the methodology for preprocessing the data frame and performing exploratory data analysis requires a thorough understanding of the data and the ability to manipulate it in a way that is useful for analysis. This step is critical in ensuring that the data is accurate, reliable, and suitable for analysis.

**5.3 Methodology for Objective 3**

The methodology for scaling the data and fitting it into the algorithm for the detection of electric vehicles with green license plates may involve the following steps:

1. Data scaling: Data scaling is an important step in preprocessing the data before fitting it into the algorithm. It involves transforming the data to a specific range or distribution to help improve the performance of the algorithm. Common scaling methods include standardization, normalization, and min-max scaling.

2. Data splitting: The data is typically split into training, validation, and testing sets before fitting it into the algorithm. This helps to prevent overfitting and allows for the evaluation of the model's performance on unseen data. The training set is used to fit the model, the validation set is used to tune the model's hyperparameters, and the testing set is used to evaluate the model's performance.

3. Selecting the algorithm: The choice of algorithm depends on the specific task and the characteristics of the data. For example, for the detection of green license plates, a classification algorithm such as logistic regression, decision tree, or support vector machine may be used.

4. Fitting the algorithm: Once the data is scaled and the algorithm is selected, the next step is to fit the algorithm on the training data. This involves using the training set to train the algorithm by adjusting its parameters to minimize the error between the predicted and actual outputs.

5. Tuning hyperparameters: Hyperparameters are parameters that are not learned by the algorithm and must be set before fitting the model. These parameters can have a significant impact on the performance of the model, and therefore, tuning them is an important step in the fitting process. This can be done using techniques such as grid search, random search, or Bayesian optimization.

6. Model evaluation: Once the algorithm is fitted and its hyperparameters are tuned, it is evaluated on the testing set to assess its performance. This involves using various performance metrics such as accuracy, precision, recall, and F1 score to evaluate the model's performance.

7. Documentation: It is important to document the entire process, including the scaling and fitting steps, hyperparameter tuning, and evaluation metrics used. This documentation can help ensure that the analysis is reproducible and can be used by other researchers or practitioners in the future.

Overall, the methodology for scaling the data and fitting it into the algorithm requires careful consideration of the data and the algorithm used. This step is critical in ensuring that the model is accurate and reliable for the detection of electric vehicles with green license plates.

**5.4 Methodology for Objective 4**

The methodology for Optical Character Recognition (OCR) involves the following steps:

1. Data collection: The first step is to collect a large dataset of images that contain text, including license plate images. This dataset should be diverse and representative of different lighting conditions, angles, fonts, and languages.

2. Data preprocessing: The collected images may need to be preprocessed before using them for OCR. This may involve resizing the images, enhancing the image quality, and converting them to a standardized format.

3. Text detection: Once the images are preprocessed, the next step is to detect the areas of the image that contain text. This may involve using techniques such as Edge Detection, Morphological Transformations, and Connected Component Analysis.

4. Character segmentation: Once the areas of the image containing text are identified, the characters need to be segmented. This involves separating each character in the license plate image to obtain individual character images.

5. Character recognition: Once the characters are segmented, the next step is to recognize each character. This may involve using machine learning techniques such as Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), or Convolutional Neural Networks (CNNs).

6. Post-processing: Once the characters are recognized, the OCR output may require post-processing to improve its accuracy. This may involve techniques such as spell checking, grammar checking, and dictionary-based correction.

7. Evaluation: Finally, the OCR system's performance is evaluated on a separate testing dataset to assess its accuracy. This involves using metrics such as precision, recall, and F1 score to evaluate the accuracy of the OCR system.

8. Deployment: Once the OCR system is trained and evaluated, it can be deployed for practical applications, such as automatic license plate recognition systems.

Overall, the methodology for OCR involves careful consideration of the dataset, preprocessing techniques, text detection, character segmentation, character recognition, post-processing, and evaluation. This methodology is critical to ensure that the OCR system is accurate and reliable for practical applications.

### 5.5 Methodology for Objective 5

To recognize the green color using RGB values, the following methodology can be used:

1. RGB values of the color green: In the RGB color model, the color green is represented by the combination of red, green, and blue values. The green color has a higher value of green component (G) as compared to the other components (R and B). The RGB values for green can vary depending on the specific shade of green being recognized.

2. Thresholding: Once the RGB values are determined, the next step is to set a threshold value for the green component. This threshold value determines the minimum value of green component required to classify a pixel as "green." Any pixel with a green component value above the threshold will be classified as green, and those below the threshold will be classified as non-green.

3. Image segmentation: After the threshold is set, the next step is to apply the threshold to the entire image to segment the green regions. This is done by assigning a binary value of 1 to all pixels above the threshold and 0 to all pixels below the threshold.

4. Morphological operations: After image segmentation, morphological operations such as dilation and erosion can be applied to the segmented image to remove noise and fill in gaps in the green regions.

5. Feature extraction: Once the green regions are segmented and cleaned, features such as the size, shape, and texture of the green regions can be extracted to differentiate between green license plates and other green objects.

6. Classification: Finally, the ratio of the green color present in the image can be used to determine if the green belongs to a license plate or not.

Overall, this methodology involves determining the RGB values of the green color, setting a threshold for the green component, segmenting the image, applying morphological operations, extracting features, and performing classification. This methodology can be used for green license plate recognition, where the license plate's green color is one of the critical features to detect.

## VI. Implementation

The project was successfully implemented on Google Colab using pandas, OpenCV, NumPy and other python libraries. The online EasyOCR was installed on Colab notebook and used for character recognition in the license plate.

The output of the project included a string indicating if the vehicle was electric or not, an array that indicated the position of license plate, the number of the license plate and the confidence level of prediction of the number recognized by the OCR.

## VII. Results And Snapshots

The results indicate that the system has a good accuracy for the provided images. A total of 17 images were used for testing the working of the system, 16 of which recognized the characters perfectly and 1 image with a minor error.

Here are some outputs of the system tested:


```
The Vehicle is Not Electric

[([[80, 78], [230, 78], [230, 110], [80, 110]], '21 BH 2345 AA', 0.9927693560732336)]
```

Fig 3. Recognition of non-electric vehicle 1


The Vehicle is Not Electric

[([[232, 182], [326, 182], [326, 208], [232, 208]], 'MH15BD8877', 0.42953682245658387)]

Fig 4. Recognition of non-electric vehicle 2


The Vehicle is Electric

[([[78, 72], [194, 72], [194, 96], [78, 96]], 'KA 05 ME 5596', 0.7296955146655977)]

Fig 5. Recognition of electric vehicle 1


```
The Vehicle is Electric

[([[108, 78], [238, 78], [238, 104], [108, 104]], 'MH 01 CH  4839', 0.7191803984684223)]
```

Fig 6. Recognition of electric vehicle 2

## VIII. Applications

The basic approach of this project is to help identify electric vehicles on the road and use this acquired data in many ways. The benefits of this project are huge because it can help understand a lot of relationships between the increasing number of electric vehicles on the road and the changes occurring in the environment due to this.

Some of these applications are as listed below:

- To help traffic control authorities identify vehicles causing traffic violations
- To provide access of the data to environmentalists to analyze impact of electric vehicles on environment
- To enable the government to monitor the increasing number of electric vehicles on the roads
- Create datasets of license plate with their position and confidence level of detection

## IX. Future Scope And Enhancements

The scope of this project is vast as the era of vehicles run by fuels is slowly dying and people have started to prefer using electric vehicles.

With further advancements made in the technology, the need for a system to analyze the existing and upcoming changes would be necessary. The proposed system can be analyzed by various different kinds of authorities as well as be combined with new information to lead to meaningful insights.

Some ways to enhance the system are as given below:

- Accuracy of the system can be improved
- It can be customised to detect license plate in various countries
- The feature to detect license plates of different colours such as red, yellow, etc., can be added to classify vehicles into different categories

## X. Contribution To Society And Environment

The proposed system can contribute to society and the environment in many ways. Some contributions that can be idealized are as follows:

- To help traffic control authorities identify vehicles causing traffic violations
- To provide access of the data to environmentalists to analyze impact of electric vehicles on environment
- To enable the government to monitor the increasing number of electric vehicles on the roads

## XI. Conclusion

The proposed system detects electric vehicles by detecting the green color of the license plates which is unique to electric vehicles. It also returns the number and position of the license plate which can be used for creating datasets and further analysis.

## References

**[1]** Lee, Yang Yang, Zaini Abdul Halim, and Mohd Nadhir Ab Wahab. "License Plate Detection Using Convolutional Neural Network–Back to the Basic with Design of Experiments." IEEE Access 10 (2022): 22577-22585.

**[2]** Mufti, Naveed, and Syed Afaq Ali Shah. "Automatic number plate Recognition: A detailed survey of relevant algorithms." Sensors 21, no. 9 (2021): 3028.

**[3]** Tonge, Shubhra, Bhavisha Hemwani, Gulshan Thawani, Sonnal Katara, and Nikhil Dhanwani. "AUTOMATIC NUMBER PLATE RECOGNITION."

**[4]** T Joy, Deborah, Gitika Kaur, Aarti Chugh, and Shalini Bhaskar Bajaj. "Computer Vision for Color Detection." International Journal of Innovative Research in Computer Science & Technology (IJIRCST) ISSN (2021): 2347-5552.

**[5]** GOMATHY, CK, Ms CHITTURI CHAITANYA, Ms B. JAI SAI ANVITHA, and Ms CHERAPALLI JAHNAVI. "COLOR DETECTION USING PANDAS AND OPENCV." (2021).