Final Project: Enhanced Road Sign Detection

Chau Minh Thi Nguyen

Georgia Institute of Technology

CS 6476 Computer Vision

Abstract

In this project, traffic sign detection and recognition will be performed by leveraging the German Traffic Sign Detection Benchmark (GTSDB) dataset. This dataset includes 900 full images of size 1360 x 800 pixels, which is already divided into a training set of 600 images (including annotations about the location and type of the traffic signs in the images) and a testing set of 300 images. The approach involves preprocessing the input images using HSV color segmentation to distinguish traffic sign regions from the background, combined with Maximally Stable Extremal Regions (MSERs) to refine potential regions of interest (ROIs). Further noise reduction and refinement of the ROIs will be performed. Feature extraction will then be performed using Histogram of Oriented Gradients (HOG), which captures the shape and appearance of traffic signs. HOG features will be used as input to the system of 2 Support Vector Machine (SVM) models: the first is a binary classification model that distinguishes between positive and negative ROIs (with positive indicating traffic sign regions), while the second is a multi-class classification model that classifies the 43 distinct traffic sign types. This integrated approach enables detection and recognition of traffic signs in real-world scenarios.

I. Introduction to Traffic Sign Detection and Recognition

Traffic sign detection and recognition are critical components in autonomous driving systems and advanced driver-assistance systems (ADAS). Accurate detection of traffic signs is vital for ensuring safety on the roads, as it directly affects navigation, route planning, and decision-making for autonomous vehicles. Despite advancements in computer vision, challenges such as variations in lighting, occlusion, changes in weather, and cluttered environments make this task complex. This project aims to detect and classify traffic signs from images using a combination of traditional computer vision techniques and machine learning methods. We leverage the German Traffic Sign Detection Benchmark (GTSDB) dataset to train and test our models, combining preprocessing techniques, feature extraction, and machine learning classification to improve traffic sign recognition accuracy.

II. Description of Existing Methods

In recent years, traffic sign detection and recognition have attracted significant attention in computer vision research. Several methods have been proposed, each with its strength and limitations.

- Convolutional Neural Networks (CNNs): Recent advancements in object detection leverage CNNs, with notable techniques including Region-based CNN (R-CNN), Fast R-CNN, Faster R-CNN, Region-based Fully Convolutional Network (R-FCN), Single Shot Detector (SSD), and You Only Look Once (YOLO). Choosing the right model often involves evaluating multiple metrics such as accuracy, precision, recall, average precision, and recall, as well as runtime and memory usage.
- Currently, traffic sign detection methods are categorized into two main groups: those based on traditional machine learning for object detection, using methods such as feature extraction and image processing techniques, and those based on deep learning techniques and large pre-trained models.

- Different machine learning techniques, such as K-means clustering and K-neighbors classifier, have shown improved accuracy in traffic sign detection tasks by effectively segmenting regions of interest and classifying traffic signs based on their geometric and visual features. These methods are often used in conjunction with feature extraction techniques, such as Histogram of Oriented Gradients (HOG), to further enhance detection accuracy and robustness in varied real-world conditions.

III. Methodology

The approach implemented in this project is divided into several key stages: image preprocessing, region proposal, feature extraction using HOG, and two-stage classification.

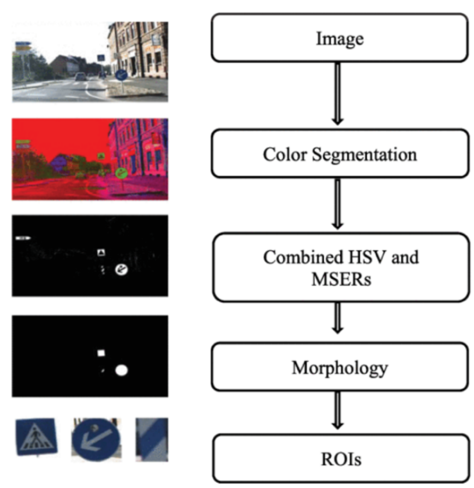
i. ***Image Pre-processing:*** We begin by converting the input images in the training set to the HSV color space, where the hue component is used for the color-based segmentation. Since most of the signs in this dataset contains a red or blue border, a color segmentation method that specifically combines the red and blue channels, creating a binary image where white regions signify the colored areas. This is basically HSV color segmentation on red and blue. For further distinction between areas of interest, I also include Gaussian adaptive thresholding on the segmented binary image.

ii. ***Maximally Stable Extremal Regions (MSER):*** MSER is used to detect potential regions of interest (ROIs) that might contain traffic signs. This step helps eliminate irrelevant parts of the image and refine the possible locations of traffic signs. After determining these regions, we further reduce noises by using morphological operations, erosion and dilation, to further refine the white regions that represent the possible traffic signs and the black regions that represent the background. To make sure that we are accurately identifying potential traffic signs, not just any colored regions, area and aspect ratio thresholding are applied to the regions detected by the MSER method. Area thresholding ensures that regions with sizes below or above a predefined area threshold are filtered out, improving the likelihood of correctly detecting traffic signs. Since traffic signs typically have a specific aspect ratios, aspect ratio thresholding is also performed to keep only the most prominent regions.

iii. ***Feature Extraction:*** We extract features using the Histogram of Oriented Gradients (HOG) method to capture edge directions and shapes, which are essential for recognizing traffic sign patterns. These HOG features are used as input to the following SVM classifiers.

iv. ***Two-stage Classification:***

- Binary Classification Model: To determine which regions of interest might contain a traffic sign, a binary SVM model is trained from a small dataset I created using the GTSDB dataset. For every traffic sign region included in the training set, I randomly select 3 negative background regions and combine them into a training set for this model. This classifier is used to select from the regions of interest those that are most likely to contain a traffic sign.
- Multi-Class SVM Classification Model: After identifying the traffic sign regions, a multi-class SVM classifier trained on the HOG features of the traffic signs in the training set is used to classify the testing set images. Each image may contain one or more traffic sign, thus requires us to go through this pipeline from detecting specific regions of interest to classifying the sign in that chosen region.



The image below shows the image preprocessing pipeline used in our project (Aris et al.,2022). The combination of HSV color segmentation, MSER, and morphology creates a few prominent areas highlighted in white. These identified regions represent potential traffic sign areas. Along with our binary classification model, we could further shorten the list of regions of interest needed to examine.

IV. Results

The binary SVM classifier has an accuracy of above 99%, whereas the multi-class SVM classifier was applied to recognize the specific traffic sign, which achieves an accuracy of 72.44%, with notable performance in red and blue border signs.

The positive examples where multiple signs are detected and classified in diverse settings:







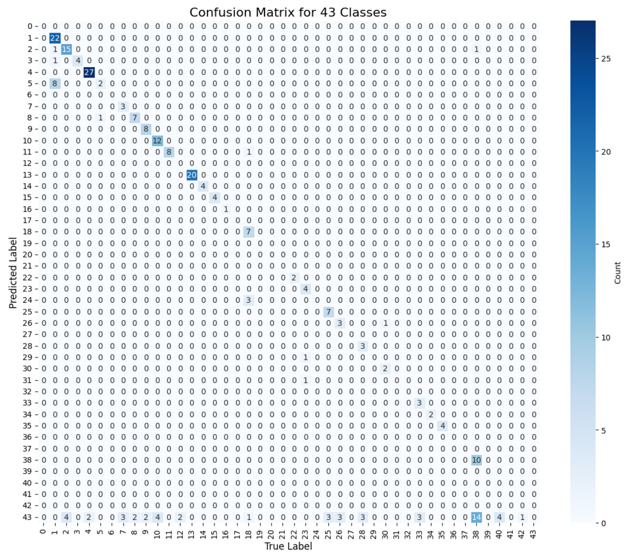

The negative examples are shown below:

- Correctly detected signs, but wrong classification

- Incorrectly identified headlights as traffic sign or other regions as traffic signs

- Partially detected signs (either due to motion that causes blurriness, close regions of interests, lighting/glare, etc.)

Using the ground truth annotation provided by the GTSDB dataset, which includes each frame, position of the sign, and the sign's class ID on each line of the text file, we will be able to construct a confusion matrix to evaluate the performance of our traffic sign detection and classification pipeline.



Note that we use the class label '43' to denote no traffic signs are detected on the image. In the ground truth annotations, image files with no traffic signs are not included in the text file. Using this we can see that our method performs well on the first 10 signs. These are typically red border signs, which are easily detected by combining the HSV thresholding and MSER method. Because of this, it makes red lights from the back of cars become detected and misclassified as traffic signs very easily. Our method performs well on images with clear visibility of the traffic signs, but its performance is degraded in cases where signs are partially occluded or where there are significant lighting issues or clutter in the background. Additionally, due to limitations in preprocessing steps taken, signs that are transformed or rotated are very hard to be detected. Dark photos or direct contact with the sun also makes classification challenging. Surprisingly, some positive pictures above could still be detected and classified correctly despite dark lighting, fog in the background, or direct sun glare pointing toward the camera.

V. Conclusion

Compared to the current state-of-the-art traffic sign detection methods, our approach using color segmentation, MSERs, followed by SVM classification, performs similarly to traditional methods but is outperformed by deep-learning models. Recent works using Fast-CNNs have shown higher accuracy and consistency in detecting and classifying traffic signs despite lighting, weather, and environment conditions. In my code, I was actually trying to integrate adaptive binary thresholding and the binary mask created from the HSV color segmentation part to generate a mask for black and white bordered traffic signs as well. However, some trade-offs, including detecting artifacts with HOG features similar to the red and blue border signs, have prevented me from following that route. Due to the preprocessing pipeline, detecting traffic lights require different image processing techniques, and making it hard to combine the implemented traffic light detection method in PS2 to this classification pipeline. Moreover, one way to minimize misclassification and improve detection rates in this system is to augment the dataset, provide varying illumination, adding noise, affine transformations, etc. to make the training set more robust and the model more receptive the external factor changes. Furthermore, HOG features, when used as input to Convolutional Neural Networks (CNNs), have demonstrated how machine learning models can develop a deep-level understanding of image features. By combining the edge and shape descriptors captured by HOG with the hierarchical feature extraction capabilities of CNNs, models can learn more intricate patterns and representations within the data. Ultimately, this project suggests that robust preprocessing techniques are crucial in creating a strong foundation for traffic sign detection and classification, making relatively simple machine learning methods effectively handle multi-class classification tasks in diverse settings.

References

Anatoliy Popov, Irina Vasilyeva, Volodymyr Kosharskyi, Konstantin Dergachov, "Selection of Color Contrast Objects Against a Non-Stationary Background Using Modified HSV Model", *2023 IEEE International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo)*, pp.84-87, 2023. 10.1109/ICICoS56336.2022.9930588

Yehia Zakaria, Mohamed Ashraf Ali, Hossam E. Abd El Munim, Ahmed Hassan Yousef, Maged Ghoneima, Sherif Hammad, "A Novel Vehicle Detection System", *2018 13th International Conference on Computer Engineering and Systems (ICCES)*, pp.127-131, 2018. 10.1109/URAI.2016.7734067

Siddharth Thoviti 1 , Naga Venkata Rama Sai Sri Harsha Bonala 2 1,2 Department of Computer Science & Information Management, Asian Institute of Technology, Bangkok, 12120, Thailand, “Object Detection of Traffic Signs using Faster Region-based Convolutional Neural Networks (Faster R-CNN)”

U.Zakir, A.N.J.Leonce,E.A.Edirisinghe, “Road Sign Segmentaiton Based on Colour Spaces: A Comparative Study”, Digital Imaging Research Group, Department of Computer Science, 2010.