Bilkent University
CS 491 Senior Design Project I
Fall 2022-2023

Project Specifications Report

RoadVisor

Group: T2317

Group Members
Ahmet Faruk Ulutaş - 21803717
Ammaar Iftikhar - 21901257
Arda İçöz - 21901443
Emin Berke Ay - 21901780
Nurettin Onur Vural - 21902330

Supervisor
Asst. Prof. Dr. Hamdi Dibeklioğlu

Innovation Expert
Cem Çimenbiçer

This report is submitted to the Department of Computer Engineering of Bilkent University in partial fulfillment of the requirements of the Senior Design Project course CS491/2.

1. Introduction

Navigation in unknown areas is a difficulty faced by automobilists. The automobilist may be an experienced taxi driver or a novice. Navigators don't only face the hurdle of driving through unknown streets, but also face the task of keeping track of road information like traffic lights and street signs simultaneously. The presence of so much information for the driver to process might lead to increased mistakes or accidents [1]. There are other issues that pertain to the road, like boundaries and crowd density, that play a crucial role to inform the driver. If a greater amount of information is present to the driver with more ease, the driver will be able to perform better.

While there are many applications that attempt to present the driver with directions to the destination, there aren't many that help simultaneously inform the driver of other necessary information simultaneously. These applications mostly focus on delivering the queried information to the driver like the location of certain places and distance to the next turn. Information about processing real-time information for the navigator on the road is not prioritized by advanced applications like Google Maps. Our application attempts to fill the disparity between the user view and the application. RoadVisor assists the drivers in real-time using Machine Learning and Augmented Reality.

The main motivation behind RoadVisor is to assist the driver in processing the real-time events on the street fetched using the mobile phone camera while also providing the user with an improved street view and directions on their mobile phones. The ability to process information and the information fetched by the application is constrained by the processing capabilities of the devices. The application will use different Machine Learning methodologies, like Deep Neural Networks, and APIs, like Google Maps Platform API, to continuously inform the driver about road information and directions. Augmented Reality will try to ensure that the driver does not miss the road information while consulting the mobile for directions or other information.

The application will mitigate traffic violations, like violation of street signs and traffic lights, by drivers. The application will try to fill the void in the navigation application industry by presenting a viable and better alternative to the ones available in the market. Our application will be an attempt to assist the driver without distracting or bugging to decrease the fatalities caused by road accidents or car crashes [1]. The application will be built incrementally and optimized for usage on mobile devices.

RoadVisor aims to have a place in the Offering section and be a Product Performance focused application according to the 10 Types of Innovation Wheel by Doblin [7]. We believe that the AR-supported navigation feature distinguishes RoadVisor from other applications and provides better and more comfortable functionality. With these distinguishable features, RoadVisor will be an incremental innovation project and optimization will be the key element as the goal is to provide a better user experience.

2. Description

The application will use Augmented Reality to assist the driver navigate. The view provided to the user will include the road view and additional information that improves the information output to the driver. The features we intend to implement are:

1. Augmented Reality Supported Navigation: The feature will provide the user the directions to navigate by showing arrows in the augmented reality view. It will also display road boundaries and turns. The user will be able to view the road as well as the directions. The feature will be implemented using image processing and computer vision techniques. Deep Neural Network methodologies like, Road and road Boundary detection Network (RBNet), are available for road boundary and region detection [2]. We will additionally use Google Maps API to fetch directions.

2. Traffic Light Detection: The feature will inform the driver about traffic light signals. The feature will help improve driver’s awareness of traffic lights. The information utilized by the feature will be fetched using the back camera of the device. The feature is in line with our motive to help drivers process road information with ease. The feature will be implemented using Computer Vision, particularly Convolutional Neural Networks (CNNs) [3]. Such mechanisms are used in ADAS systems. The DriveU dataset is a candidate for use during training and testing of the application [4]. The DriveU dataset is based on European traffic lights. There are some other candidates that can be used, but the problem with most of the options is their American nature, i.e, the American and European traffic light signals differ. Background noise like the presence of multiple traffic lights might, however, decrease the accuracy.

3. Sign Detection: The application will help users be aware of the traffic signs that might be on the road like ‘school ahead” or “danger ahead.” The signs the feature will recognize are expected to follow Vienna Convention of Road Signs. The constraint mentioned exists due to the dataset we will use. It is a dataset consisting of 80000 samples from 150+ classes [5]. The implementation of this feature will also utilize Convolutional Neural Networks. The feature assists the user to prevent violation of signs that might lead to other consequences like fines.

4. Crowd Density Analysis: The application will provide information to the user about the pedestrians and crowd in a particular location. The application will use crowd counting techniques to implement this feature [6]. We can implement this feature using CNNs as well. Shanghai Tech dataset and the UCF_CC_50 dataset are 2 possible datasets candidates that we might use for the training of the Neural Networks.

Our first goal is to finish implementing feature number 1 (see above). The features here are numbered based on the priority we have assigned to them. Therefore, the flow of implementation will also follow that. Our end goal is to finish the implementation of all the features we have listed above. However, this might change due to the long developing process of the AR navigation system. The below mentioned features will only be implemented if we can successfully implement the above mentioned features and the devices used to test the application can work effectively in the presence of other features. This latter constraint exists due to the presence of limited computational power of mobile phones and the fact that the entire application is based on Augmented Reality, which is heavier than other simple applications.

5. Crash Detection: Assisting our users during times of dangers like car crashes is a feature that we want to include. The feature detects a car crash by using Computer vision. Then, it notifies the emergency service to respond to the car crash. The feature might be able to save the lives of people by immediately contacting the concerned authorities. However, we need to consider the hardware limitations of mobile devices when implementing this feature. We are working on a possible feasible solution for this feature.

6. Song Recommendation Based On Road: The feature is meant to be for the entertainment of the user. The feature will play a song based on the type of road that the user will be driving on. For example, it will play metal on the highway or jazz in a high-traffic area. The type of music in what circumstance can be configured by the user.

3. Professional and Ethical Issues

- The application will need to use location services in order to provide accurate navigation. This location information is only shared with the API of selected navigation applications (like Google Maps). The user's location will not be stored anywhere, and the application will ask the user to use this information. Without consent, the application will not work.
- Images retrieved from the camera of the phone will only be used for AR and computer vision tasks. The images will not be stored and therefore shared with any third party.
- The navigation information provided by the application and the placement of the AR elements might not be accurate depending on the location. RoadVisor does not guarantee that AR elements are supported in every road type and location.

- Pedestrian detection, traffic lights/signs detection and crash detection features might not be accurate and work in every condition. The driver is responsible for their and the passenger's safety.
- The application is recommended to be used while the phone is charging. This situation might affect the phone's battery life negatively in the long term. Users will be informed about this effect on battery life.

4. Constraints

4.1. Implementation

- Unity will be used to implement augmented reality elements and C# language will be used. In some cases, OpenCV can also be used for the same or similar purposes. The deployment of the application will be done in Unity by using its built-in publishing feature.
- Android Studio might help in the deployment process. Java or Kotlin languages will be used here based on library compatibility.
- Python and related machine-learning libraries will be used for machine learning.
- GitHub will be used as the versioning system for the application development process to be healthy.
- Some machine learning models will be tested with existing models. If the desired result is not achieved at this stage, the models will be rebuilt to increase accuracy.
- Object-oriented programming (OOP) principles will be used.
- For analysis, data acquisition will be done by using the rear camera.
- The implementation details mentioned above, the language and the technologies to be used, may expand or change according to the implementation process.

4.2. Economic

- The machine learning models will be developed using free open-source libraries and datasets. We can use real datasets that we recorded ourselves during the testing process.
- Technologies and APIs to be used while developing the application will be free (Google Maps API, Android SDK).
- The application is expected to be free in general, and the main features are limited to the user. We expect more comfortable features to be available with a paid subscription. However, at the end of the development process, the RoadVisor team might decide to provide all features for free.
- In order to test the application, it must be placed on the infotainment or entertainment system of the car in such a way that it can follow the road

through the windshield. In this, adhesive, magnet or holder for the phone can be used. And their prices are around 50 TL.

4.3. Manufacturability

- Since we have more than 4 cars in our team and around us, we can test the application with different car models and test for different real-life scenarios.
- Having different phone models and camera versions will provide the opportunity to test different potential users' experiences
- The team both have MacOS and Windows powered computers. Even if the operating systems are different, the same development environments will be used for compatibility while developing the application.
- Nowadays, car phone holders are a very popular and inexpensive tool, so they are in most cars. It will be added to the missing ones.

4.4. Health and Safety

- The main logic of the application is to increase the comfort and safety of the user while on the road. In this context, a more advanced AR-based navigation system is offered to the user. It is crucial that the navigation system is reliable so that it doesn't endanger the safety of the driver and passengers with false markers and recommendations.
- Additionally, traffic lights/signs and pedestrian detection features must be optimized to give no or minimum false results. These false results must not put the driver at risk of any kind. If this cannot be provided, these models should work only on highly accurate cases.
- Because the application will require the driver to look at the phone screen, this can be dangerous for the driver and passengers. The application will warn the driver periodically to concentrate on the road.

4.5. Social

- Since there is no interaction with different application users in the application, the application will not have a social constraint.

4.6. Political

- The application is not affiliated with any political activity and therefore cannot be used as a political tool.

4.7. Technological

- In order for more users to use the application, the application will be compatible with the minimum Android version possible where the features can work optimally.

- In order for the application to be able to use its navigation and AR features, it will be expected that the rear camera works properly and there is an internet connection.

4.8. Language

- The application language will be English. All pages and components that can interact with the user will be in English.

4.9. Time

- The prototype of the application must be presented in mid-December 2022, and the application with the proposed features must be ready by mid-May 2023.
- Project reports must be submitted before the deadline time.

4.10. Sustainability

- Updates will be made to increase application usability in line with user feedback and real user tests.

4.11. Environment

- The application needs to provide the most fuel-efficient direction. Considering the fact that every gallon of gasoline burned creates around 8.8 kilograms of carbon dioxide and an average vehicle emits 4.6 tons of carbon dioxide annually, choosing the most fuel-efficient direction will help save tons of fuel and emit less carbon dioxide gas [8].

4.12. Ethics

- Personal information will not be shared with third parties.
- The application will not work if the user consent cannot be retrieved for getting personal information (like location services).

5. Requirements

5.1. Functional Requirements

5.1.1. Road Navigation with Augmented Reality

- One of the key intentions behind RoadVisor is to provide the user with complete guidance during the entire road trip. Accordingly, the application not only aims to take the user to a selected location in an optimal way but also gets directly involved with showing the necessary movements of the vehicle.

- RoadVisor aims to achieve this by showing the direction that the vehicle has to follow with the help of Augmented Reality. In other words, the application uses the camera of the mobile phone to track the road while inserting arrows in accordance with the path that the driver has to follow to reach the desired location.

- To start this navigation service, the user needs to enter the destination location and afterwards, RoadVisor will get the footage of the road from the camera as an input to match it frame by frame with the route of the desired location to place the arrows onto the road.

- Arrows need to be constantly changing with respect to the road and furthermore when more drastic movements are required such as turning left or right, those will be emphasized on the screen with larger arrows popping up.

- The screen also has to show information about the current location and the remaining distance to the location.

5.1.2. Traffic Light and Sign Detection

- RoadVisor needs to use the live camera footage of the user to analyze it frame by frame.

- Using this camera footage, the application has to constantly make image analysis to be able to detect the presence of traffic lights and traffic signs visible.

- For traffic lights, RoadVisor needs to inform the user by showing the color of the light on the phone screen and giving a pop-up sign when it spots a red light.

- For traffic signs, RoadVisor needs to inform the user about any sign spotted and what the sign is about. Signs will be classified and shown on the phone screen where signs that require further attention such as stop signs can also have additional pop-up messages for the purpose of alerting.

5.1.3. Pedestrian Detection

- RoadVisor needs to use the phone camera of the user to analyze live camera footage frame by frame.

- By getting the camera footage as input, the application has to constantly make image analysis to be able to detect the presence of any pedestrian that is spotted on the road.

- When RoadVisor spots a pedestrian, it needs to mark it on the screen and furthermore show an accompanying warning message for alerting purposes.

5.1.4. Requesting Help and Crash Detection

- The user will be able to enter phone numbers as "urgent contact" information.

- When the user is in need, he/she will be able to press the SOS button of RoadVisor so that the application sends an automatic message to the urgent contact numbers stating that the user may be in a problematic situation together with the location of the user.

- The SOS feature of RoadVisor can be also activated automatically if the application detects a sound that may resemble a car crash sound. In that case, the application asks the user if there is a problematic situation. If the user fails to respond in a given duration, RoadVisor will accept this as an urgent issue and automatically alert the urgent contact numbers.

5.2. Non-functional Requirements

5.2.1. Reliability

- It's crucial for RoadVisor to give navigation information correctly. This means that the application has to be reliable in terms of fulfilling the expected necessities such as preparing the route accurately, placing the direction arrows correctly, showing true traffic light and sign information messages, and successfully sending the SOS messages with the location of the user.

- RoadVisor has to operate without significant interruptions and therefore provides navigation service and other promised services during the entire car trip duration starting from the moment that the destination is selected until the user arrives at it given that the necessary conditions are satisfied such as battery and connection.

5.2.2. Performance

- RoadVisor has to have a high standard in terms of performing well in frequent conditions and in a short duration. Since the application needs to constantly extract input from live camera feeds and analyze them frame by frame, the frame rate has to be above a tolerable level so that the application works without significant delay.

- The detection models used within RoadVisor have to perform with high accuracy rates for being considered reliable.

- The application needs to work smoothly with the touch screen which requires RoadVisor to have a fast response time.

- RoadVisor is designed mainly to work in a mobile environment. In this respect, the application has to perform well to satisfy its features without being overly affected by hardware limitations of mobile phones.

5.2.3. Usability

- RoadVisor is an application that aims to reach a very wide range of audience. Essentially, all drivers (especially ones who do not possess the latest technological tools in their cars) are targeted as potential users of our application. In this respect, the application needs to be understandable and easy to use by the target audience.

- RoadVisor needs to be compatible with a wide range of Android phones given that they are above a determined version.

- RoadVisor needs to have a highly responsive, simple-looking user interface. Given that the user will be busy with driving as well the UI has to be designed in such a manner that it will not have very complex operations for the user that take various steps. Instead, the parts that require user interaction must be large enough, easily locatable on the phone screen, and respond within a short amount of time (less than a minute).

- Since the user will be driving meanwhile, the application must not cause unnecessary distraction or irritate the user. This will be achieved by avoiding using irritating alert sounds, having a simple UI layout, and decreasing the need for user input (such as automatically closing pop-ups after a while).

5.2.4. Privacy

- Confidential information such as contact numbers, destination information, or exact route data will not be used with third parties to respect user privacy.

- To minimize the risk of exposure of personal information in a possible attack, confidential data will be kept encrypted.

References

[1] T. Shang, H. Lu, P. Wu, and Y. Wei, "Eye-tracking evaluation of exit advance guide signs in highway tunnels in familiar and unfamiliar drivers," *International Journal of Environmental Research and Public Health*, vol. 18, no. 13, p. 6820, 2021.

[2] Z. Chen and Z. Chen, "RBNet: A deep neural network for Unified Road and Road Boundary Detection," *Neural Information Processing*, pp. 677–687, 2017.

[3] C. Wang, G. W. Zhang, W. Zhou, Y. Rao, and Y. Lv, "Traffic lights detection based on Deep Learning Feature," *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pp. 382–396, 2020.

[4] A. Fregin, J. Muller, U. Krebel, and K. Dietmayer, "The DRIVEU Traffic Light Dataset: Introduction and comparison with existing datasets," *2018 IEEE International Conference on Robotics and Automation (ICRA)*, 2018.

[5] C. Gamez Serna and Y. Ruichek, "Classification of Traffic Signs: The European dataset," *IEEE Access*, vol. 6, pp. 78136–78148, 2018.

[6] J. Xiang and N. Liu, "Crowd density estimation method using deep learning for passenger flow detection system in exhibition center," *Scientific Programming*, vol. 2022, pp. 1–9, 2022.

[7] "Ten types," *Doblin*, 21-Apr-2020. [Online]. Available: https://doblin.com/ten-types. [Accessed: 16-Oct-2022].

[8] "Greenhouse Gas Emissions from a Typical Passenger Vehicle", *EPA*. [Online]. Available: https://www.epa.gov/greenvehicles/greenhouse-gas-emissions-typical-passenger-vehicle. [Accessed: 17-Oct-2022].