Thesis Booklet

Towards a Transfer Concept for Camera-Based Object Detection

From Driver Assistance to the Assistance of Visually Impaired Pedestrians

Judith Jakob

Supervisor:
Dr. habil. József Tick

Doctoral School of Applied Informatics and Applied Mathematics

June 8, 2020

Contents

Acronyms

CoPeD Comparable Pedestrian Driver 6–8, 12–14, 16, 18–20

AAL Ambient Assisted Living 4, 8

ADAS Advanced Driver Assistance Systems 3–8, 11–20

ASVI Assistive Systems for the Visually Impaired 3–8, 11–14, 16–20

EDF Edge Distribution Function 14, 18

EI Expert Interviewee(s) 9, 10

GPS Global Positioning System 3

HD High Definition 7

ML Machine Learning 18

MTG Member(s) of the Target Group 7, 9, 10

NN Neural Network 14, 16

OCR Optical Character Recognition 11, 12

RBS Road Background Segmentation 14, 16–20

ROI Region Of Interest 14, 16, 18

SSD Sensory Substitution Device(s) 6

SVM Support Vector Machine 14, 16

TGGS Tactile Ground Guidance System 11, 12

TTS Text To Speech 3

1 Introduction

According to [1], worldwide 36 million people were estimated to be blind and 216.6 million people were estimated to have a moderate to severe visual impairment in 2015. The according numbers have increased in recent years: in comparison to 1990, there was a rise of 17.5 % in blind people and 35.5 % in people with moderate to severe visual impairment. Due to demographic change a higher occurrence of diseases causing age-related vision loss is to be expected in the near future, leading to a decrease in the autonomous personal mobility of the affected [2]. In their study about how the visually impaired perceive research about visual impairment, Duckett and Pratt underline the importance of mobility for this group: "The lack of adequate transport was described as resulting in many visually impaired people living in isolation. Transport was felt to be the key to visually impaired people fulfilling their potential and playing an active role in society," [3].

In order to increase their independent mobility, I research possibilities of assisting visually impaired people in traffic situations by camera-based detection of relevant objects in their immediate surroundings. Although some according research concerning Assistive Systems for the Visually Impaired (ASVI) exists, there is a much higher amount of research in the field of Advanced Driver Assistance Systems (ADAS). Hence, it is suitable to profit from the progress in driver assistance and make it applicable for visually impaired pedestrians. As it is in general not possible to use driver assistance algorithms without any adaptations, e. g. due to known and stable camera positions, a generalized concept for the transfer is needed. Therefore, the presented research leads the way towards a universal transfer concept from ADAS to ASVI for camera-based detection algorithms.

Such a transfer concept makes it possible to integrate the resulting ASVI algorithms into a camera-based mobile assistive system. Adding detection algorithms for the remaining use cases that are of relevance for the visually impaired in traffic situations results in an assistive system that offers comprehensive support.

Although no such system is built in the course of the thesis' research, a sketch of a camera-based ASVI using a commercially available smartphone is shown in Figure 1.1 to demonstrate a possible application scenario of the presented content. A camera as well as earphones or a hearing aid to provide Text To Speech (TTS) output are connected to the smartphone. Computationally expensive image processing calculations are exported to a cloud service and relevant external information that can support image detection (e. g. Global Positioning System (GPS) coordinates of

crosswalks or construction sites) is extracted and provided through the according cloud module. Due to the diversity of visually impaired people, it is important to take profiling and personalization into account when building an assistive system. This is achieved by using the Ambient Assisted Living (AAL) platform described in [4].
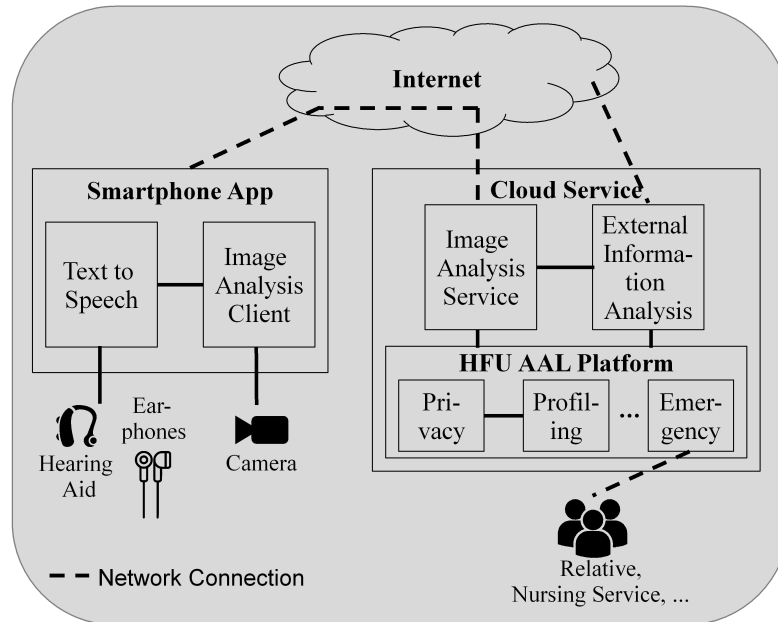

Figure 1.1: Sketch for a Camera-Based Mobile Assistive System

2 Literature Review and Novelty

Considering the computer vision pipeline of ADAS according to [5], the steps of computer vision in ADAS according to [6], and the general system design of ASVI that I developed on the basis of a literature review of camera-based ASVI (see Figure 2.1), it can be stated that even though different terminologies are used, three phases are in the center of data processing in all described cases: Preprocessing/Early vision, feature extraction/recognition, and evaluation/classification/decision. Maps and localization is only mentioned in [6], but a similar procedure can be observed in the ASVI project *Crosswatch* [7]. Additionally, context information is used in ADAS as well as ASVI in different forms, data transmission to the processing unit has to be addressed, and appropriate ways of communicating the results have to be developed. In summary, it can be stated that the overall system design, especially the computer vision steps, of camera-based ADAS and ASVI are similar.
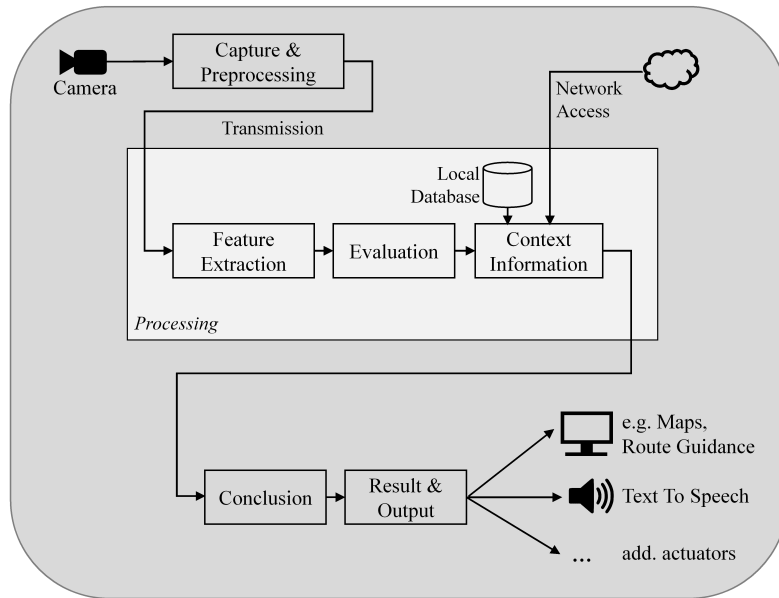
Figure 2.1: System Design of ASVI

Furthermore, the identified application areas of ADAS and ASVI presented in the thesis reveal that there is in general an overlap in addressed and needed use cases, e. g. collision avoidance and support at crossings. At the same time, the amount of research in the field of ADAS is much higher than in ASVI. In addition to the amount of research, the quality of detection algorithms used in latest car models has to be very high to ensure the safety of the driver and other road users. It is therefore desirable to make latest and future advancements in ADAS applicable for visually impaired people.

Hence, it follows that using ADAS algorithms in an ASVI is possible because of the similar compositions of such systems and because of the overlap of relevant use cases. The transfer concept is necessary because ADAS algorithms can in general not be used for pedestrians without any adaptations. This is mainly due to a priori assumptions, such as known and stable camera position, that are only valid when used in cars.

Most of the ASVI described in the thesis perform a specific task or several tasks that can be assigned to one topic. These systems have the potential to take on a variety of functions based on their composition. However, this is usually not exploited since often only prototypes for a specific research purpose are built. An exception is the commercial system *MyEye*[1] which in addition to its main task of recognizing and

---

[1] https://www.orcam.com/en/, accessed on June 6, 2020

reading text can recognize objects and faces; further applications are under development. Naturally, Sensory Substitution Device(s) (SSD) such as [8] have no other functions because of their approach of transferring the visual perception to other senses. Some of the introduced systems cover single traffic topics such as crossings (e. g. [7]), but there is no system or research project that covers the needs of visually impaired people in all significant traffic scenarios by camera-based detection. Therefore, the presented research leads the way to provide comprehensive assistance in all traffic scenarios that are of relevance for the visually impaired based on the evaluation of camera footage.

3 Research Categories and Objectives

This chapter gives an overview of the three research categories treated in the thesis. An examination of traffic scenarios with relevance to the visually impaired results in a complete collection of vision use cases that could support visually impaired pedestrians. The overlap of these vision use cases with the ones addressed in ADAS needs to be considered in the following. First, the video data set *CoPeD* containing comparable video sequences from pedestrian and driver perspective for the identified overlapping use cases is created. These data are used in order to evaluate the algorithms that are adapted from ADAS to ASVI. In the following, I describe the objectives of each research category and name used methods and tools.

3.1 Definition of Traffic Scenarios and Vision Use Cases for the Visually Impaired

Objectives

This category's purpose is the understanding of needs visually impaired people have as pedestrians in traffic situations. From the gathered insights, a list with relevant vision use cases has to be acquired and the overlap with vision use cases addressed in ADAS has to be built.

In detail, the following objectives have to be achieved throughout the research in this category. The according results will be summarized in Thesis 1 in chapter 4.

(O1.1) All traffic scenarios that are of interest for visually impaired pedestrians have to be defined.

(O1.2) All vision use cases that can support the visually impaired in traffic situations have to determined.

(O1.3) The overlap of vision use cases addressed in ADAS and needed in ASVI has to be determined.

(O1.4) The idea of using (adapted) software engineering methods to cluster and present qualitative data has to be introduced.

Besides, the research in this category is expected to answer the following questions:

(Q.a) Are there differences in gender and/or age of visually impaired people when it comes to dealing with traffic scenarios?

(Q.b) Is the use of technology common among visually impaired people?

(Q.c) How do visually impaired people prepare for a trip to an unknown address?

(Q.d) Are visually impaired people comfortable with having to ask for support or directions?

(Q.e) Which identified vision use cases are the most important?

Methods and Tools

To achieve these goals, I create, conduct, and evaluate a qualitative interview study consisting of expert interviews and interviews with members of the target group, namely visually impaired pedestrians. I use Witzel's problem-centered method [9] and Meuser and Nagel's notes on expert interviews [10]. Transcription and analysis are performed with the software MAXQDA Version 12 [11]. I code the interviews with inductively developed codes as proposed by Mayring [12].

By clustering the data, I determine different traffic scenarios and according vision use cases that can support the visually impaired in the respective scenario. I summarize the evaluation of the interview study in scenario tables adapted from software engineering [13]. By comparing the collection of ASVI use cases with an ADAS literature review, I determine the desired overlap.

3.2 The *CoPeD* Data Set for Traffic Scenarios

Objectives

This category addresses the acquisition of video data that are needed to compare ADAS algorithms with their ASVI adaptations that I will develop in the next research category. For the evaluation of these algorithms, comparable video data from both perspectives, driver and pedestrian, have to be gathered. It is important that the video data cover all identified overlapping vision use cases from ADAS and ASVI. Although there are numerous data sets covering traffic scenarios, they are mostly

from driver perspective and no according comparable data from driver and pedestrians perspective exists. Therefore, I create the Comparable Pedestrian Driver (*CoPeD*) data set for traffic scenarios. The data set is made publicly available and others are permitted to use, distribute, and modify the data.

In detail, the following objective has to be achieved throughout the research in this category. The according results will be summarized in Thesis 2 in chapter 5.

(O2) The data set *CoPeD* containing comparable video data from driver and pedestrian perspective and covering the overlapping use cases from ADAS and ASVI has to be created.

Methods and Tools

Review and analysis of according scientific literature reveal that no comparable data exist. Therefore, I create and publish the *CoPeD* data set for traffic scenarios. For the planning of the data set, I use activity diagrams from software engineering [13]. The sequences are filmed in High Definition (HD) with a *Kodak PIXPRO SP360 4K* camera.

3.3 Use Case Examination

Objectives

The overlapping use cases have to be examined concerning their possibilities of adaptation from ADAS to ASVI. For each use case, appropriate algorithms from ADAS have to be chosen and adapted algorithms have to be developed. It is important to show that the adapted algorithms perform at least as good as the underlying ADAS algorithms so that they are applicable in an assistive system.

In detail, the following objectives have to be achieved throughout the research in this category. The according results will be summarized in Thesis 3 in chapter 6.

(O3.1) It has to be shown that determining the Region Of Interest (ROI) for ASVI detection algorithms can in general not be taken from ADAS and that adapting a Road Background Segmentation (RBS) from ADAS to ASVI solves this problem.

(O3.2) Adaptations of algorithms from ADAS to ASVI have to developed and implemented. The adapted algorithms have to achieve similar hit rates as the underlying ADAS algorithms.

Methods and Tools

I develop adaptations to use ADAS algorithms in ASVI and implement the adapted algorithms in Matlab Version R2017b [14]. Afterwards, they are evaluated on several sequences from the *CoPeD* data set.

4 Definition of Traffic Scenarios and Vision Use Cases for the Visually Impaired

In the qualitative study presented in the thesis, the problems visually impaired pedestrians face in traffic situations are collected by conducting four expert interviews and ten interviews with Member(s) of the Target Group (MTG) using Witzel's problem-centered method [9] in both cases. From the data, a set of six different scenarios are extracted (*General Orientation*, *Navigating to an Address*, *Crossing a Road*, *Obstacle Avoidance*, *Boarding a Bus*, *At the Train Station*) clustered into the three categories *Orientation*, *Pedestrian*, and *Public Transport* scenarios. For each of the six scenarios, a descriptive table inspired by scenario tables from software engineering [13] is created. The tables summarizes the usual procedure of the respective scenario, the problems that can occur and which vision-based use cases could help to overcome the problems. As the objective of the thesis is to work towards the formulation of a transfer concept for vision-based use cases and according algorithms from driver assistance to the assistance of visually impaired people, the following research is based on the overlapping use cases pointed out in Figure 4.1.
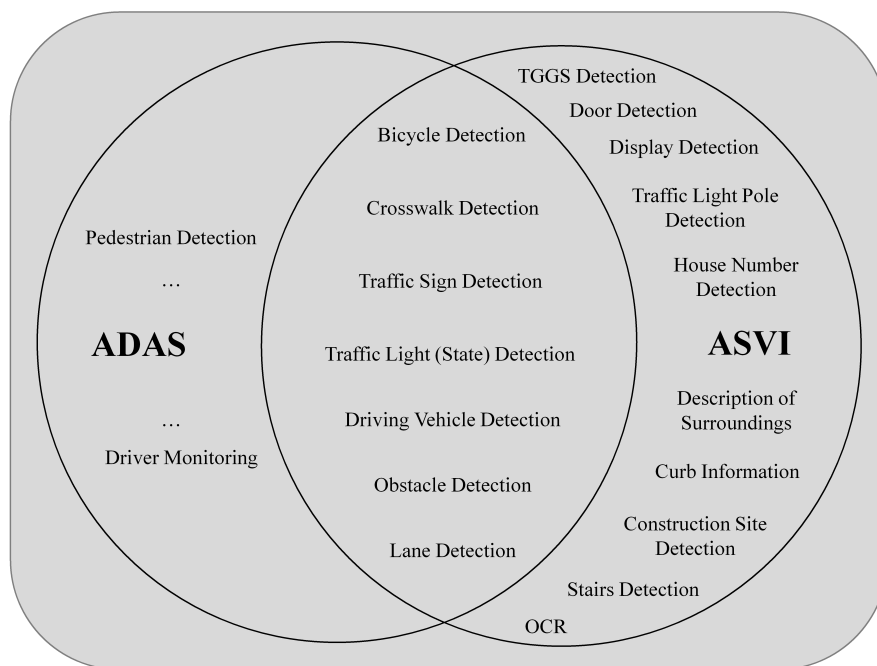

Figure 4.1: Overlapping Vision Use Cases from ADAS and ASVI

With the data from the interviews, some further findings that go beyond the collection of traffic scenarios and are of importance for researchers who want to build a camera-based assistive system can be formulated. Although there are some use cases that seem to be of general importance, the diversity of visually impaired people concerning kind and degree of impairment(s) as well as personality leads to very differing needs. Therefore, it is necessary to consider possibilities of profiling and personalization when building an assistive system, e. g. by using an AAL platform as presented in [4] and as suggested in the sketch for an assistive system (see Figure 1.1). Among the generally important use cases, the perception of any kinds of signs and displays has to be emphasized because of its importance in numerous situations in daily life. Shen and Coughlan [15] for example address this problem.

With the help of the interview data, I can answer the before formulated questions:

(Q.a) Are there differences in gender and/or age of visually impaired people when it comes to dealing with traffic scenarios?
Concerning gender, one Expert Interviewee(s) (EI) stated that girls and young women are more likely to attend additional voluntary mobility workshops than boys and young men. According to the EI, life experience with visual impairment is more important than age.

(Q.b) Is the use of technology common among visually impaired people?
One EI states: "When you have a limitation, you depend on technology and of course you use it,". This is underlined by the fact that all interviewed MTG use at least PC or smartphone.

(Q.c) How do visually impaired people prepare for a trip to an unknown address?
Visually impaired people essentially cover the same topics as the sighted, but in general they need more information and it takes more time to gather the information. It also depends on a person's personality, no matter if sighted or visually impaired, how much information is gathered before a trip.

(Q.d) Are visually impaired people comfortable with having to ask for support or directions?
There are opposing reports from the interviewees. It depends on a person's personality how comfortable they are with asking people for support or directions.

(Q.e) Which identified vision use cases are the most important?
Almost all discussed use cases are of importance for the blind interviewees. Contrarily, interviewees with residual vision give very differing answers. It depends on a person's concrete impairment but again also on their personality, which and how much support is required. Nevertheless, it can be stated that the perception of all kinds of signs, indoor and outdoor, as well as obstacle detection are very important use cases for the visually impaired.

When discussing differences between the sighted and visually impaired and talking about needs of visually impaired people, it is important to keep in mind that "the blind and visually impaired are as different individuals as you and your colleagues," (EI4).

That the idea of a camera-based assistive system in traffic situations is met with approval is underlined by the following two interviewee quotes:

> One would be much more independent. It could help in all areas of life. (MTG1)

> We currently have the rapid development of smartphones, and with that we are also experiencing more and more comfort. And in this context, such a development and research as yours is of utmost importance, so that one can achieve more safety in road traffic. (EI3)

In the following, the four parts of the objective formulated in 3.1 are discussed. Objectives (O1.1), (O1.2), and (O1.3) are achieved because of the data saturation of the qualitative interview study meaning that the enumerations of traffic scenarios and use cases are exhaustive.

(O1.1): All traffic scenarios that are of interest for visually impaired pedestrians have to be defined.
Inductively coding the interview data leads to six traffic scenarios that are of interest for the visually impaired and that can be clustered into three categories: *Orientation Scenarios* (*General Orientation*, *Navigating to an Address*), *Pedestrian Scenarios* (*Crossing a Road*, *Obstacle avoidance*), and *Public Transport Scenarios* (*Boarding a Bus*, *At the Train Station*).

(O1.2): All vision use cases that can support the visually impaired in traffic situations have to determined.
From the tables created for each traffic scenario, I extract and gather all 17 relevant vision use cases (see also Figure 4.1): (1) Traffic light pole detection, (2) traffic light (state) detection, (3) bicycle detection, (4) (driving) vehicle detection, (5) stairs detection, (6) construction site detection, (7) crosswalk detection, (8) obstacle detection, (9) lane detection, (10) curb information, (11) Tactile Ground Guidance System (TGGS) detection, (12) traffic sign detection, (13) house number detection, (14) description of surroundings, (15) Optical Character Recognition (OCR), (16) door detection, and (17) display detection.

(O1.3): The overlap of vision use cases addressed in ADAS and needed in ASVI has to be determined.
I determine the overlap in use cases of relevance for the visually impaired and the ones addressed in ADAS by comparing the above created ASVI list of 17 use cases with ADAS literature. This results in seven overlapping use cases (see also Figure

4.1): (1) Lane detection, (2) crosswalk detection, (3) traffic sign detection, (4) traffic light (state) detection, (5) (driving) vehicle detection, (6) obstacle detection, and (7) bicycle detection.

(O1.4): The idea of using (adapted) software engineering methods to cluster and present qualitative data has to be introduced.

In my publication [16], I discuss objective (O1.4). The literature provides several procedures for the analysis of data acquired in qualitative research, including codes which I used for the evaluation of the presented study. But as suitable structuring and representation of the gained knowledge are highly depending on the concrete problem, there is no "simple step that can be carried out by following a detailed, "mechanical," approach. Instead, it requires the ability to generalize, and think innovatively, and so on from the researcher," [17, p. 63]. I master this challenge by adapting methods from software engineering. Using qualitative methods to improve the software development process has been applied in the past (see e. g. [18, 19]), but the reverse – using software engineering methods to improve the representation of qualitative data – is a new approach. The many structuring elements, such as different tables and diagrams, found in software engineering are powerful tools that can help to generalize qualitative data and other procedures beyond software.

Hence, all four objectives are achieved and can be summarized in Thesis 1.

Thesis 1: Traffic Scenarios and Use Cases
I defined the significant traffic scenarios for visually impaired pedestrians and determined all vision use cases of relevance in these scenarios. From that, I determined the overlap of vision use cases between ADAS and ASVI. Besides, I introduced the idea of using software engineering methods for the presentation of qualitative data.

(T1.1) I showed that the traffic scenarios of interest for visually impaired pedestrians are: *Orientation* scenarios (*General Orientation*, *Navigating to an Address*), *Pedestrian* scenarios (*Crossing a Road*, *Obstacle Avoidance*), and *Public Transport* scenarios (*Boarding a Bus*, *At the Train Station*).

(T1.2) I determined all vision use cases that can support the visually impaired in traffic situations: (1) Traffic light pole detection, (2) traffic light (state) detection, (3) bicycle detection, (4) (driving) vehicle detection, (5) stairs detection, (6) construction site detection, (7) crosswalk detection, (8) obstacle detection, (9) lane detection, (10) curb information, (11) TGGS detection, (12) traffic sign detection, (13) house number detection, (14) description of surroundings, (15) OCR, (16) door detection, and (17) display detection.

(T1.3) I determined the overlap of vision use cases addressed in ADAS and needed in ASVI: (1) Lane detection, (2) crosswalk detection, (3) traffic sign detection,

(4) traffic light (state) detection, (5) (driving) vehicle detection, (6) obstacle detection, and (7) bicycle detection.

(T1.4) I introduced the idea of using (adapted) software engineering methods to cluster and present qualitative data.

Own publications supporting Thesis 1 are: [16, 20, 21, 22].

5 The *CoPeD* Data Set for Traffic Scenarios

The literature review presented in the thesis shows that existing data sets for traffic scenarios - such as [23, 24, 25, 26, 27, 28, 29, 30] - do not cover all vision use cases needed in my research. Besides, the existing data are mostly from driver perspective and no comparable data from driver and pedestrian perspective exist. Thus, I present the publicly available *CoPeD* data set for traffic scenarios. Divided into the four categories lanes, crossings, obstacles, and traffic signs, I created comparable video sequences from driver and pedestrian perspective. In the case of traffic signs, single images containing traffic signs with importance to pedestrians were collected. As can be seen Figure 5.1, *CoPeD* covers all overlapping use cases from ADAS and ASVI. Thus, the data set makes it possible to compare the performances of ADAS algorithms and their adapted ASVI versions that will be developed in the following. It is licensed under the Creative Commons Attribution 4.0 International License[1] and hosted publicly[2] which makes it available for all researchers.

For the planning of the video sequences, activity diagrams from software engineering [13] were used. The folder available for download is structured as follows: Besides the read-me and license files, the main folder contains four sub folders, one for each category. The folders *Lanes*, *Crossings*, and *Obstacles* each contain the sub folders *Driver* and *Pedestrian* which then contain the comparable video sequences. Some of these folders contain another sub folder called *Others* consisting of further video sequences that belong to the according topic but do not have a comparable scene from driver respectively pedestrian perspective. The folder *Traffic Sign* encloses four sub folders that contain several images with the according traffic sign. The sub folders are: *Bus Stop*, *Crosswalk*, *Pedestrian and Bicycle Path*, and *Others*.

---

[1] https://creativecommons.org/licenses/by/4.0/, accessed June 6, 2020
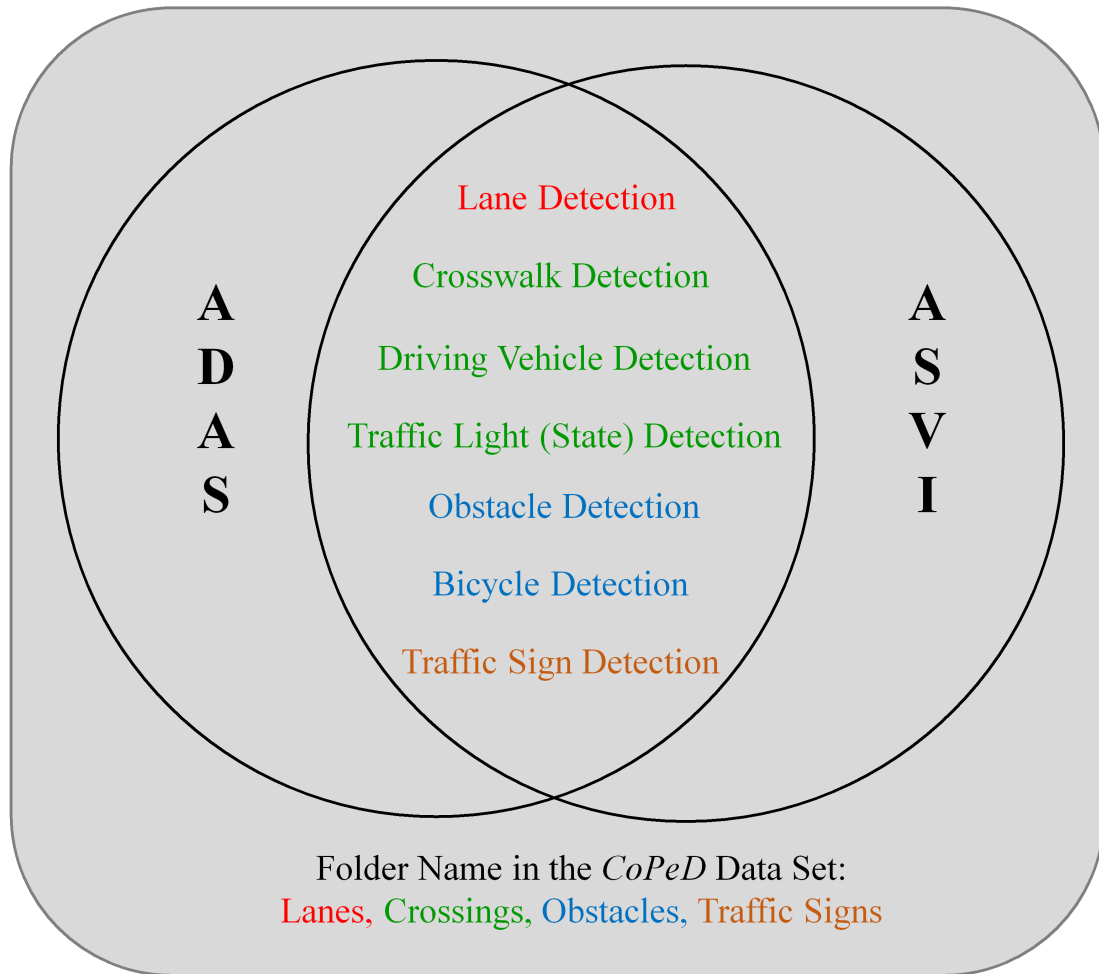[2] http://dataset.informatik.hs-furtwangen.de/, accessed June 6, 2020

Figure 5.1: Coverage of Use Cases of the *CoPeD* Data Set

In the following, objective (O2) is discussed.

(O2): The data set *CoPeD* containing comparable video data from driver and pedestrian perspective and covering the overlapping use cases from ADAS and ASVI has to be created.
Existing traffic data sets are usually recorded from driver and not from pedestrian perspective. The newly created *CoPeD* data set on the other hand contains traffic scenarios filmed from both perspectives. Recording the according scenarios in the same place and directly one after the other results in comparable sequences that are suitable to compare the evaluation of algorithms from driver and pedestrian per-spective. The fact that all overlapping use cases are covered is shown in Figure 5.1.

Hence, objective (O2) is achieved and can be summarized in Thesis 2.

Thesis 2: Video Data Acquisition
Thesis (T2): I created the data set *CoPeD* containing comparable video data from driver and pedestrian perspective and covering the overlapping use cases from ADAS and ASVI.

Own publications supporting Thesis 2 are: [22, 31].

6 Use Case Examination

I identified a total of seven use cases that have to be considered. I focused on the detection of markings on the road, namely crosswalk and lane detection. In order to solve the ROI problem and to reduce the computational cost, I first introduced a RBS. Thereby, crosswalk and lane detection can be carried out on the road only. Algorithms for traffic sign and traffic light detection to be developed in the future will be applied to the background part of the image, whereas obstacles will be detected on the whole image.

Two RBS adaptations are presented in the thesis. The first is based on Beucher et al.'s [32] use of morphological watersheds. In addition, the adaptation uses thresholds for the mean gray and saturation values of each catchment basin. The resulting algorithm reached a hit rate of 99.87 %. The second algorithm is based on Foedisch and Takeuchi's [33] idea of using Neural Network (NN) and introduces weights to determine if a block of an image is likely to be part of the road respectively the background. The according hit rates are 99.41 % when using a Support Vector Machine (SVM) and 99.87 % when using a NN.

The basis of the developed crosswalk detection is Choi et al.'s idea of using a horizontal mean filter [34]. I combined the horizontal with a vertical filter and defined a decision process. The recognition rate of the adapted algorithm is 98.64 %.

The suggested lane detection uses the Edge Distribution Function (EDF). Instead of applying the EDF to the whole image as in Lee's ADAS work [35], I first divide the image into subimages of decreasing size from bottom to top. Afterwards, interpolation of the angles and analysis of the according function return the course of the road. Correct detection occurs in 97.89 % of the examined frames.

Figure 6.1 summarizes the presented algorithms and highlights the adaptation steps. Table 6.1 summarizes the evaluation of the adapted algorithms executed on *CoPeD* frames. With that, it can be stated that the presented ASVI algorithms, adapted from ADAS methods, achieve overall good hit rates and thus are applicable for ASVI.
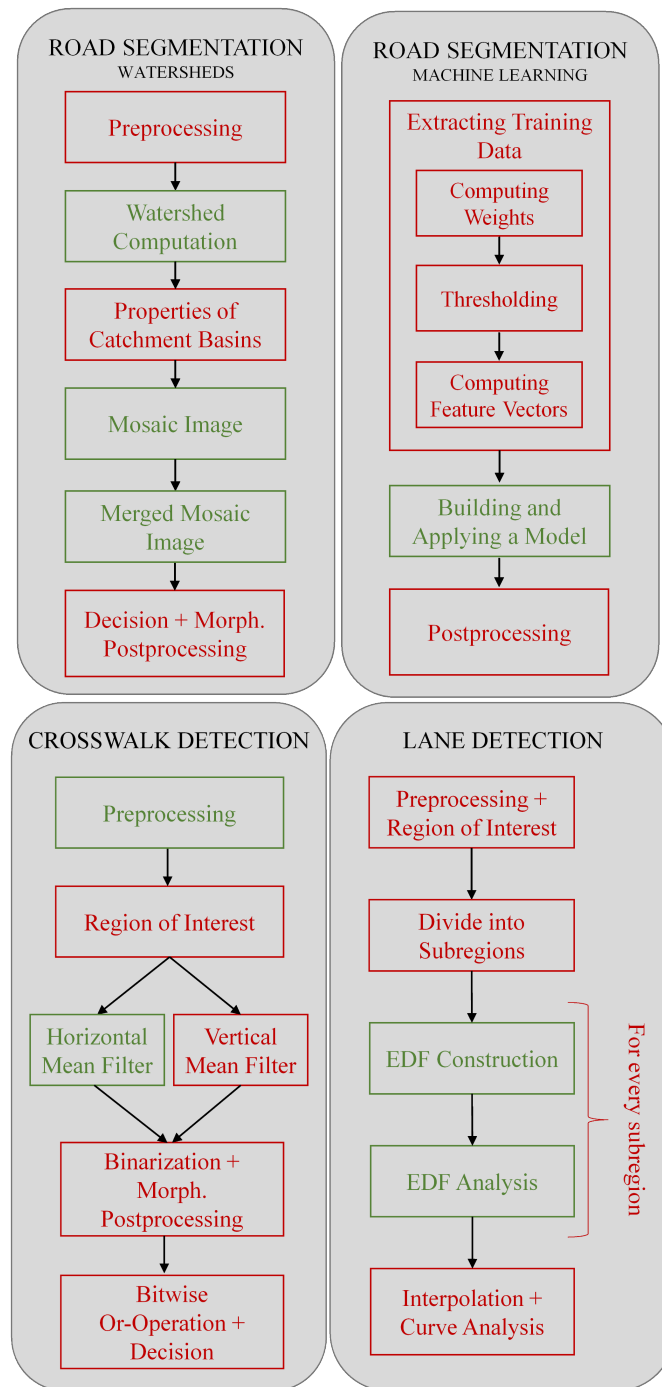
Figure 6.1: Procedures of Presented Algorithms. Green boxes are taken from the underlying ADAS algorithms with small changes. Red boxes required major changes or are newly introduced.

Table 6.1: Evaluation: RBS Watersheds (RBS WS), RBS SVM, RBS NN, Crosswalk De-
tection (CWD), and Lane Detection (LD)

	RBS WS	RBS SVM	RBS NN	CWD	LD
Number of Frames	1541	1541	1541	1541	1516
Correct Detection	1539	1532	1539	1520	1484
Wrong Detection	2	9	2	21	32
Hit Rate	**99.87 %**	**99.41 %**	**99.87 %**	**98.64 %**	**97.89 %**

In the following, objectives (O3.1) and (O3.2) are discussed.

**(O3.1): It has to be shown that determining the ROI for ASVI detection algo-
rithms can in general not be taken from ADAS and that adapting a RBS from
ADAS to ASVI solves this problem.**
Knowing position and angle of the camera is exploited in ADAS. However, in ASVI
no such assumptions can be made. Therefore, a RBS as an additional use case is
proposed. Adapting RBS from ADAS to ASVI makes it possible to carry out some de-
tection algorithms on a subset of the original image. Altogether, this confirms that
(O3.1) is achieved.

**(O3.2): Adaptations of algorithms from ADAS to ASVI have to developed and
implemented. The adapted algorithms have to achieve similar hit rates as the
underlying ADAS algorithms.**
A comparison between the performances of the underlying ADAS algorithms and
their ASVI adaptations is difficult to achieve because so far they have not been
tested on the same data. In the future, the ADAS algorithms can be implemented
and tested on the according comparable video sequences from the *CoPeD* data set.
Nevertheless, (O3.2) is achieved for RBS, crosswalk detection, and lane detection
because the presented hit rates are similar to the ones reported in the literature
(provided hit rates are reported) and high enough to make the algorithms applica-
ble in ASVI.

Hence, (O3.1) is achieved and (O3.2) is achieved for RBS, lane detection, and cross-
walk detection. These findings are summarized in Thesis 3.

Thesis 3: Adaptation Possibilities
**I showed that adapting a RBS from ADAS to ASVI solves the ROI problem and
that ASVI adaptations for RBS, lane detection, and crosswalk detection achieve
similar hit rates as the underlying ADAS algorithms.**

(T3.1) I showed that determining the ROI for ASVI detection algorithms can in gen-
eral not be taken from ADAS and that adapting a RBS from ADAS to ASVI solves
this problem.

(T3.2) I developed and implemented adaptations of algorithms from ADAS to ASVI for RBS, lane detection, and crosswalk detection. I proved that the adapted algorithms in these three cases achieve satisfying and similar hit rates as the underlying ADAS algorithms.

Own publications supporting Thesis 3 are: [16, 22, 31, 36, 37, 38, 39].

7 Contributions

Below, I first summarize the contributions made in the respective chapters of the thesis and then list my own publications.

Related Work and Novelty

I reviewed the literature in the fields of camera-based ASVI and camera-based ADAS. I derived a general composition of ASVI from the literature; for ADAS the according composition was already discussed in the literature. Furthermore, I clustered camera-based ASVI into four application areas: Reading out text, recognizing faces and objects, perceiving the environment as well as navigation and collision avoidance. From the similarity of the compositions, the overlap in use cases, and the fact that there is no comprehensive assistive system for the visually impaired in traffic situations, I concluded the need for and novelty of a transfer concept for camera-based algorithms from ADAS to ASVI.

Own publications related to this topic: [22, 36, 40].

Traffic Scenarios and Vision Use Cases

I reviewed the literature containing studies about the requirements of visually impaired people in traffic situations. From that, I concluded the necessity of conducting an own study. In the following, I designed, conducted, and evaluated a qualitative interview study with four experts and ten members of the target group. With the help of the acquired data, I defined six traffic scenarios and 17 vision use cases with importance to visually impaired pedestrians. Forming the overlap with use cases addressed in ADAS revealed seven use cases. I presented a literature review of ADAS solutions for these use cases. Furthermore, I answered questions concerning age, gender, use of technology, trips to unknown addresses, asking for support, and use case importance, and I introduced the idea of using (adapted) software engineering methods for clustering and presentation of qualitative data.

The *CoPeD* Data Set

I reviewed literature about video and image data sets for traffic scenarios from driver and pedestrian perspective. As the existing data sets did not cover all needed use cases from both perspectives, I designed and developed the publicly hosted *CoPeD* data set containing comparable sequences from driver and pedestrian perspective for all seven overlapping use cases. It is licensed under the Creative Commons Attribution 4.0 International License which allows everyone, even in commercial contexts, to use, modify, and redistribute the data as long as appropriate credit is given.

Use Case Examination

I introduced RBS as a further use case to be considered in order to solve the ROI problem. I performed a literature review for RBS, crosswalk detection, and lane detection. I then developed adaptations to ASVI for Beucher et al.'s watershed-based RBS [32], Foedisch and Takeuchi's Machine Learning (ML)-based RBS [33], Choi et al.'s crosswalk detection based on a 1-D mean filter [34], and Lee's EDF-based lane detection [35]. The adaptations were implemented in Matlab [14] and evaluated on sequences from *CoPeD*.

List of Own Publications

The following list contains my publications with relation to the thesis. The list is arranged by decreasing date so that the most recent publication is on the top. Parts of these works were presented and cited in the thesis.

- J. Jakob and J. Tick: "Extracting training data for machine learning road segmentation from pedestrian perspective," in IEEE 24th International Conference on Intelligent Engineering Systems, Virtual Event, July 2020, Accepted for Publication, [38]:
  - Literature review for RBS.
  - Description of proposed ASVI adaptation for RBS based on ML.
  - Evaluation of the proposed algorithm on *CoPeD* sequences.

- J. Jakob and J. Tick: "Camera-based on-road detections for the visually impaired," in Acta Polytechnica Hungarica, vol. 17, no. 3, pp. 125 - 146, 2020, [37]:
  - Literature review for RBS, crosswalk detection, and lane detection.
  - Description of proposed ASVI adaptations for RBS based on watersheds, crosswalk detection, and lane detection.
  - Evaluation of the proposed adaptations on *CoPeD* sequences.

- J. Jakob and J. Tick: "Towards a transfer concept from camera-based driver assistance to the assistance of visually impaired pedestrians," in IEEE 17th International Symposium on Intelligent Systems and Informatics, pp. 53 - 60, Subotica/Serbia, September 2019, [22]:
    - Preliminary versions of the objectives treated in this thesis.
    - Summary of camera-based ASVI and camera-based ADAS.
    - Summary of the qualitative interview study.
    - Summary of *CoPeD*.
    - Summary of adapted algorithms for RBS, crosswalk detection, and lane detection.

- J. Jakob and J. Tick: "CoPeD: Comparable pedestrian driver data set for traffic scenarios," in IEEE 18th International Symposium on Computational Intelligence and Informatics, pp. 87 - 92, Budapest/Hungary, November 2018, [31]:
    - Literature review of existing data sets from driver and pedestrian perspective.
    - Conditions and content of the *CoPeD* data set.
    - Preliminary version of the crosswalk and lane detection presented in this thesis.

- J. Jakob, K. Kugele, and J. Tick: "Defining traffic scenarios for the visually impaired," in The Qualitative Report, 2018, Under Review (Accepted into Manuscript Development Program), [20]:
    - Literature review containing studies about the requirements of visually impaired people in traffic situations.
    - Design of the qualitative interview study (expert interviews and interviews with MTG).
    - Evaluation of the qualitative interview study (expert interviews and interviews with MTG).

- J. Jakob and J. Tick: "Traffic scenarios and vision use cases for the visually impaired," in Acta Electrotechnica et Informatica, vol. 18, no. 3, pp. 27 - 34, 2018, [16]:
    - Design and evaluation of the expert interviews which are one part of the qualitative study.
    - Literature review for the overlapping use cases.
    - Preliminary version of the lane detection presented in this thesis.

- J. Jakob, K. Kugele, and J. Tick: "Defining camera-based traffic scenarios and use cases for the visually impaired by means of expert interviews," in IEEE 14th International Scientific Conference on Informatics, pp. 128 - 133, Poprad/Slovakia, November 2017, [21]:

- Design and evaluation of the experts interviews (which are one part of the qualitative study).

- Literature review for the overlapping use cases.

• J. Jakob and J. Tick: "Concept for transfer of driver assistance algorithms for blind and visually impaired people," in IEEE 15th International Symposium on Applied Machine Intelligence and Informatics, pp. 241 - 246, Herl'any/Slovakia, January 2017, [36]:

- Literature review for camera-based ASVI and camera-based ADAS.

- Summary of preliminary version of the crosswalk detection presented in this thesis.

- Sketch of the future work towards a transfer concept from ADAS to ASVI.

• J. Jakob and E. Cochlovius: "OpenCV-basierte Zebrastreifenerkennung für Blinde und Sehbehinderte," in Software-Technologien und -Prozesse: Open-Source Software in der Industrie, KMUs und im Hochschulumfeld - 5. Konferenz STeP, pp. 21 - 34, Furtwangen/Germany, May 2016, [39]:

- Detailed description of a preliminary version of the crosswalk detection presented in this thesis.

• J. Jakob, E. Cochlovius, and C. Reich: "Kamerabasierte Assistenz für Blinde und Sehbehinderte - State of the Art," in informatikJournal, vol. 2016/17, pp. 3 - 10, 2016, [40]:

- Comprehensive literature review of camera-based ASVI.

8 Perspectives and Conclusion

Before concluding this booklet, I describe directions for possible future work based on the presented research.

The results of the qualitative interview study can be used as the first phase of an exploratory sequential mixed method according to Creswell and Creswell [41]. Based on the presented results, quantitative studies can be conducted, for example in order to examine correlations between type and degree of visual impairment and needed support in traffic scenarios.

The *CoPeD* data set contains video sequences under good weather and lightning conditions. It can be expanded by the addition of scenes under different conditions

and improved by labelling the data. Labels and annotations provide training data for ML techniques and make it possible to check the according results automatically.

The presented use case examination concentrates on the adaptation of the "on-road" use cases crosswalk and lane detection as well as RBS. In order to formulate a generalized transfer concept from ADAS to ASVI, the remaining overlapping use cases have to be examined as well. Afterwards, the adaptation procedures of all overlapping use cases have to be inspected and clustered into a concept.

To improve the evaluation of objective (O3.2), the ADAS algorithms on which the ASVI adaptations are based on can be implemented and the performances can be compared by using the *CoPeD* data set. As the ADAS algorithms are generally not described in detail in the according literature, their implementation is a challenging task.

The work presented in this thesis concentrates on a subset of the use cases that are of importance in ADAS as well as ASVI. Besides examining the remaining overlapping use cases in order to formulate the transfer concept, it is important to consider all use cases identified through the evaluation of the qualitative interview study when developing a camera-based ASVI.

To improve the hit rates of detection algorithms, external information can be taken into account. The sketch of a camera-based ASVI presented in Figure 1.1 therefore provides the module *External Information Analysis* as part of the cloud service. The idea is to extract information from the internet, e. g. GPS locations of crosswalks, so that there is a priori information about the image content that makes it possible to specify the algorithm accordingly.

In the course of this research, the developed algorithms were implemented in Mat-lab [14] and run on a PC. In order to make them applicable for the visually impaired, it is essential to implement them on a mobile assistive system.

The research presented in the thesis leads the way towards a generalized transfer concept of camera-based algorithms from ADAS to ASVI that will make latest and future advancements in ADAS applicable for visually impaired pedestrians. Thus, the content of the thesis makes an important contribution to the autonomous mobility of visually impaired people.

Bibliography

[1] R. Bourne, S. Flaxman, T. Braithwaite, M. Cicinelli, A. Das, J. Jonas, J. Keeffe, J. Kempen, J. Leasher, H. Limburg, *et al.*, "Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis," *The Lancet Global Health*, vol. 5, no. 9, pp. 888–897, 2017.

[2] H.-W. Wahl and V. Heyl, "Die psychosoziale dimension von sehverlust im alter," *Forum für Psychotherapie, Psychiatrie, Psychosomatik und Beratung*, vol. 1, no. 45, pp. 21–44, 2015.

[3] P. Duckett and R. Pratt, "The researched opinions on research: visually impaired people and visual impairment research," *Disability & Society*, vol. 16, no. 6, pp. 815–835, 2001.

[4] H. Kuijs, C. Rosencrantz, and C. Reich, "A context-aware, intelligent and flexible ambient assisted living platform architecture," in *The Sixth International Conference on Cloud Computing, GRIDs and Virtualization*, (Nice, France), pp. 70–76, March 2015.

[5] R. Loce, R. Bala, and M. Trivedi, *Computer Vision and Imaging in Intelligent Transportation Systems*. John Wiley & Sons, 2017.

[6] B. Ranft and C. Stiller, "The role of machine vision for intelligent vehicles," *IEEE Transactions on Intelligent Vehicles*, vol. 1, no. 1, pp. 8–19, 2016.

[7] V. Murali and J. Coughlan, "Smartphone-based crosswalk detection and localization for visually impaired pedestrians," in *IEEE International Conference on Multimedia and Expo Workshops*, (San Jose, CA, USA), pp. 1–7, IEEE, July 2013.

[8] S. Caraiman, A. Morara, M. Owczarek, A. Burlacu, D. Rzeszotarski, N. Botezatu, P. Herghelegiu, F. Moldoveanu, P. Strumillo, and A. Moldoveanu, "Computer vision for the visually impaired: the sound of vision system," in *IEEE International Conference on Computer Vision*, (Venice, Italy), pp. 1480–1489, IEEE, October 2017.

[9] A. Witzel, "The problem-centered interview," *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, vol. 1, no. 1, 2000.

[10] M. Meuser and U. Nagel, "Das Experteninterview — konzeptionelle Grundlagen und methodische Anlage," in *Methoden der vergleichenden Politik-und Sozialwissenschaft*, pp. 465–479, Springer, 2009.

[11] VERBI GmbH, "MAXQDA Version 12." https://www.maxqda.de/, Download: 2017-08-02.

[12] P. Mayring, "Qualitative content analysis," *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, vol. 1, no. 2, 2000.

[13] I. Sommerville, *Software Engineering*. Addison-Wesley, 9th edition ed., 2011.

[14] MathWorks, "MATLAB R2017b." https://de.mathworks.com/, Download: 2017-09-21.

[15] H. Shen and J. Coughlan, "Towards a real-time system for finding and reading signs for visually impaired users," in *International Conference on Computers for Handicapped Persons*, (Linz, Austria), pp. 41–47, Springer, July 2012.

[16] J. Jakob and J. Tick, "Traffic scenarios and vision use cases for the visually impaired," *Acta Electrotechnica et Informatica*, vol. 18, no. 3, pp. 27–34, 2018.

[17] P. Runeson, M. Host, A. Rainer, and B. Regnell, *Case study research in software engineering: Guidelines and examples*. John Wiley & Sons, 2012.

[18] M. John, F. Maurer, and B. Tessem, "Human and social factors of software engineering: workshop summary," *ACM SIGSOFT Software Engineering Notes*, vol. 30, no. 4, pp. 1–6, 2005.

[19] C. Seaman, "Qualitative methods in empirical studies of software engineering," *IEEE Transactions on Software Engineering*, vol. 25, no. 4, pp. 557–572, 1999.

[20] J. Jakob, K. Kugele, and J. Tick, "Defining traffic scenarios for the visually impaired," *The Qualitative Report*, 2018. Under Review (Accepted into Manuscript Development Program).

[21] J. Jakob, K. Kugele, and J. Tick, "Defining camera-based traffic scenarios and use cases for the visually impaired by means of expert interviews," in *IEEE 14th International Scientific Conference on Informatics*, (Poprad, Slovakia), pp. 128–133, IEEE, November 2017.

[22] J. Jakob and J. Tick, "Towards a transfer concept from camera-based driver assistance to the assistance of visually impaired pedestrians," in *IEEE 17th International Symposium on Intelligent Systems and Informatics*, (Subotica, Serbia), pp. 53–60, IEEE, September 2019.

[23] M. Aly, "Real time detection of lane markers in urban streets," in *IEEE Intelligent Vehicles Symposium*, (Eindhoven, Netherlands), pp. 7–12, IEEE, June 2008.

[24] J. Fritsch, T. Kuhnl, and A. Geiger, "A new performance measure and evaluation benchmark for road detection algorithms," in *16th International IEEE Conference on Intelligent Transportation Systems*, (The Hague, Netherlands), pp. 1693–1700, IEEE, October 2013.

[25] M. P. Philipsen, M. Jensen, A. Møgelmose, T. Moeslund, and M. Trivedi, "Traffic light detection: A learning algorithm and evaluations on challenging dataset," in *IEEE 18th international conference on intelligent transportation systems*, (Las Palmas, Spain), pp. 2341–2345, IEEE, September 2015.

[26] M. Jensen, M. Philipsen, A. Møgelmose, T. Moeslund, and M. Trivedi, "Vision for looking at traffic lights: Issues, survey, and perspectives," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 7, pp. 1800–1815, 2016.

[27] S. Sivaraman and M. Trivedi, "A general active-learning framework for on-road vehicle recognition and tracking," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 2, pp. 267–276, 2010.

[28] P. Dollar, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: An evaluation of the state of the art," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 4, pp. 743–761, 2012.

[29] A. Møgelmose, M. Trivedi, and T. Moeslund, "Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey.," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 4, pp. 1484–1497, 2012.

[30] S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing, and C. Igel, "Detection of traffic signs in real-world images: The German Traffic Sign Detection Benchmark," in *International Joint Conference on Neural Networks*, (Dallas, TX, USA), pp. 1–8, IEEE, August 2013.

[31] J. Jakob and J. Tick, "CoPeD: Comparable pedestrian driver data set for traffic scenarios," in *IEEE 18th International Symposium on Computational Intelligence and Informatics*, (Budapest, Hungary), pp. 87–92, IEEE, November 2018.

[32] S. Beucher, M. Bilodeau, and X. Yu, "Road segmentation by watershed algorithms," in *Pro-art vision group PROMETHEUS workshop*, (Sophia-Antipolis, France), April 1990.

[33] M. Foedisch and A. Takeuchi, "Adaptive real-time road detection using neural networks," in *7th International IEEE Conference on Intelligent Transportation Systems*, (Washington, WA, USA), pp. 167–172, IEEE, October 2004.

[34] J. Choi, B. Ahn, and I. Kweon, "Crosswalk and traffic light detection via integral framework," in *19th Korea-Japan Joint Workshop on Frontiers of Computer Vision*, (Incheon, South Korea), pp. 309–312, IEEE, January 2013.

[35] J. W. Lee, "A machine vision system for lane-departure detection," *Computer vision and image understanding*, vol. 86, no. 1, pp. 52–78, 2002.

[36] J. Jakob and J. Tick, "Concept for transfer of driver assistance algorithms for blind and visually impaired people," in *IEEE 15th International Symposium on*

Applied Machine Intelligence and Informatics, (Herl'any, Slovakia), pp. 241–246, January 2017.

[37] J. Jakob and J. Tick, "Camera-based on-road detections for the visually impaired," *Acta Polytechnica Hungarica*, vol. 17, no. 3, pp. 125–146, 2020.

[38] J. Jakob and J. Tick, "Extracting training data for machine learning road segmentation from pedestrian perspective," in *IEEE 24th International Conference on Intelligent Engineering Systems*, (Virtual Event), IEEE, July 2020. Accepted for Publication.

[39] J. Jakob and E. Cochlovius, "OpenCV-basierte Zebrastreifenerkennung für Blinde und Sehbehinderte," in *Software-Technologien und -Prozesse: Open-Source Software in der Industrie, KMUs und im Hochschulumfeld - 5. Konferenz STeP*, (Furtwangen, Germany), pp. 21 – 34, May 2016.

[40] J. Jakob, E. Cochlovius, and C. Reich, "Kamerabasierte Assistenz für Blinde und Sehbehinderte," in *informatikJournal*, vol. 2016/17, pp. 3–10, 2016.

[41] J. Creswell and J. Creswell, *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications, 2017.