Prospectus

Robust Pedestrian and Biker Detection Methods for Autonomous Vehicles Using Thermal Imaging Cameras
(Technical Topic)

Ownership in an Autonomously Driven Future
(STS Topic)

By

Emir Sahin
Spring 2021

Technical Project Team Members:
Sander Abraham
Navya Annapareddy

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Signed: Emir Sahin
Technical Advisor: Tariq Iqbal
STS Advisor: Sean Ferguson

Introduction

Autonomous Vehicles (AVs) will share a significant portion of the existing road systems in the next decade, including local roadways where the AVs will encounter more people and bikes. Although computer vision and machine learning techniques have improved significantly in the last few years, it is still challenging to detect people hearing the road with AVs, primarily because of the movement of both the Ads and the person, visual occlusion, poor lighting conditions, shadows and so on. To solve this issue, our CAPSTONE project aims to utilize the thermal imaging capabilities equipped in modern vehicles to augment the video-based pedestrian and biker recognition techniques. The project will study state-of-the-art literature on pedestrian and biker detection methods and will aim to modify those methods to work with thermal imaging cameras. After novel models are generated, these will be compared with previous models to determine an optimal solution for human and pedestrian recognition.

Both the STS and the capstone project aim to assist in the creation of a world AVs serve the majority of transportation systems in the world. Building upon the idea of a future where AVs share (or solely use) a significant share of the road systems in the world, it is rather crucial to consider the that socio-technical functions this technology will replace and create. A large aspect to consider how this technology will replace one's own vehicles and how it will affect private transportation. Hence, the STS Project explores the attempts to reposition ownership in private transportation and how it may reflect onto an AV-bound future. The STS research begins with a literature review of various takes on opinion of ownership in a digital-physical age. It will look into several different interpretations of AVs by manufacturers and their takes on ownership. This includes gathering views from possible future users of the technology. By putting the users and the manufacturers points of view into perspective, it will aim to provide a reasonable suggestion regarding AV ownership in the future.

Implementing Thermal Imaging into Pedestrian and Biker Detection

Computer vision has been frequently applied to pedestrian and biker detection in the past to augment the detection capabilities of AVs (Ahmed et la., 2019). The use of thermal imaging input in computer vision, however, hasn't been a popular area of research so far even though the possible benefits of such a sensor: while a video-feed based computer vision model may have issues with detection in various visibility and lighting conditions, a thermal imaging camera is able to identify targets independently to visibility and lighting situations. So, the project will aim to 1) collect and use thermal imaging input, 2) build a computer vision model to use said input to detect pedestrians and cyclists, and 3) validate then analyze the resulting model.

Sponsored by Perrone Robotics and with the advisory of Tariq Iqbal, Ph.D., Assistant Professor for the School of Engineering Systems and Environment, we will build a computer vision model to detect moving humans and cyclists through a moving car in real time. While doing so, we aim to create an integration into Perrone's MAX systems and collect real world pedestrian and cyclist thermal data. A crucial aspect of deep learning is having enough data to create an accurate and precise model that can be applied to novel situations. In case of annotated thermal imaging data of pedestrians and cyclists, the sources are rather lackluster as most of them are private. Hence, data creation and annotation will hold a large part of the project after several robust models are built utilizing data available such as the KAIST Database (Choi et la., 2018). The work distribution within the group is rather uniform as each member will have to code and implement the model. My responsibilities include researching various model architectures to implement and turning them into code.

While the number of methods used for pedestrian and biker detection is limited in comparison to the number of video-based detection, there have been several studies done exploring the topic. Valldor (2018) explores several methods that may be utilized in thermal pedestrian detection while pulling attention to how, in general, thermal data tends to be of lesser definition than its colored counterparts. He does indicate that the data will have to normalized and a pre-training is likely while doing analysis with thermal data as the datasets tend to be lackluster. The methods he demonstrates can be placed into three

groups: two dimensional convolutional networks, three dimensional convolutional networks, and a two-dimensional convolution network mixed with a de-convolutional network to determine weights of each feature of the image. As for right now, the initial model to be worked on this project includes a two-dimensional convolutional network mixed with a Long Short Term Memory (LSTM) in order to mix visual recognition with categorization and description (Donahue et la., 2016). After the aforementioned model is implemented, models based on You Only Look Once (YOLO) (Redmon et la., 2016) and MS-CNN (Cai et la., 2016) will be worked on as well.

While the previously mentioned models are all going to be used solely to generate results for thermal cameras, most vehicles will have a multitude of sensors. As the KAIST database that the project will utilize includes colored video feed, the project may implement a mixture of the two sources of information to provide more robust and accurate detection predictions. Two examples of such models are made by Liu et la. (2016) which explores the R-CNN model to generate a novel one utilizing a convolutional network that separately examines the same frame in colored and thermal variations and Ding et la (2020). which use a Network-in-Network to mix color and thermal values together into a region-based convolution network.

The models built will be evaluated upon their accuracy of detection. Accuracy can be described as the number of true positive biker and pedestrian detection compared against the annotations in the model. While accuracy will be used as the main evaluator, the number of false positives will be taken into consideration as well for specific situations where detection may be harder due to occlusions or other factors. Once the most robust model is determined, it will be worked onto be implemented into Perrone's MAX system for autonomous vehicles to be tested in real life.

STS Topic

In a perfect utopian future, transportation would be at its most efficient and safest: no traffic, no accidents, no congestion… A possibility through autonomous vehicles that are orders of magnitude safer than human drivers, aware of every single other vehicle on the road coupled with a road system concentrated for these omniscience vehicles. But revolutionary change doesn't come quickly, nor easily. It's hard to implement technology in various ethnographies and the socio-technical situation in the world. AVs may prove to be a revolutionary technology – so, how should they be implemented into the society? Exploring concentrated parts of the question may shed light into how AVs can and should be utilized. The average driver in the UK will spend almost 4 years at the wheel of a car in their lifetime (Hull, 2019). We should put some thought into considering how vehicle ownership, an integral part of our lives, will change as the medium moves from a strictly physical one into an aggregation of digital and physical medium.

Belk (1988) explores consumerism through ownership. His research delves into one's self and extension thereof: possessions. While the paper mainly explores how people treat their possessions and how these may provide insight onto various humanly behavior, one thing is certain: possessions play a large part in one's personal identity. With vehicles being such common possessions, it is rather straightforward to see why they would be a part of one's self. Although some people purchase vehicles solely for their purpose, the different varieties of the same type of vehicle exist to compliment one's tastes, lifestyle, and social status is innumerable. This signifies that people don't just purchase a vehicle for its specific purpose but to have it as *an extension of themselves*. Vehicles are not just any object but one that is a considerable part of socio-technical imaginaries. So, one's sense of ownership of their vehicle shouldn't be too alien of a topic for anyone.

There are several interpretations on how AVs may be implemented into our lives in the future. In his article, Motavalli (2020) provides various views on the future from 'ride-hailing' companies well as research results to indicate that surveyed samples are less likely to use ride-hailing services after the

pandemic. In the article, industry leaders and researches indicate that albeit ride-sharing is beneficial due to reduced costs, emissions, and higher efficiency, it is a little unclear how it can be the only means of transportation. Some argue that there is no other way for the industry to move forward without sharing AVs - the infrastructure may not be able to handle private and expensive AVs. Others indicate how a large fleet-based model of AVs where nobody owns a vehicle is quite unlikely due to how complicated the economic models to build around it would be. The article signifies large AV companies prefer a future in which the vehicles are shared. This is a better case scenario for large companies; they would prefer to place more weight on it. Cited IBM results prove an opposite view from customers, however.

Lee et la. (2019) has made statistical models to explore a user's perspective on AV ownership. I will solely focus on their finding regarding the weight of Psychological Ownership as a factor. The study considers ownership psychologically. This relates to the views of companies from the previous paragraph, as the people using these vehicles may be interested in the use of autonomous vehicles without actually owning them, but rather, feeling ownership of them. It is crucial to take note of how people's psychological ownership makes them more likely to use AVs without increasing their perceived usefulness of the technology. This implies that ownership is an important factor in the process of an autonomously driven future.

On Pink et la.'s paper (2020), this trend follows. While the paper specifically explores the feelings of trust and sharing (as sharing may hold the opposite stance of ownership), it touches upon various modes of ownership and their implications through analogies from possible future users of the technology. The paper states that while car sharing is more efficient and cost-effective than owning a car, the situation of implementation isn't as clear cut. It explores how the implementation can better fit into the society by suggesting an anthropological approach to studying the topic. The complaint regarding ownership is different for each analogy the paper provides: one argues that they feel a loss-of-freedom as they need to schedule everything ahead of time; another argues that the vehicle does not feel cozy and lacks the psychological ownership factor. Taking the notion from Pink et la., I do believe that the

exploration of people's views of physical ownership of the vehicle is quite crucial in creating a happy medium for AVs.

Ownership does not stop there, however. Other important factors to consider are privacy. Back in 2016, Germany pushed out a black-box requirement for autonomous vehicles that aims to prove whether it's user or the vehicle's error in case of an accident. They argue this would assist in working on the technology and its implementation (Staff, 2016); some may consider this a breach of their privacy as they may not wish to share every detail of how they utilize their vehicles. This Blackbox may feel as if they *lack ownership* of their *own* vehicle. Another concern is the full physical ownership of the vehicle, including repairability. While it's common for a phone manufacturer to not provide spare parts and brick the device in case of a non-original part in the device, this isn't necessarily the case for vehicles. One is able to modify and repair their vehicles to their desire. This may not be the case with AVs in the near future. Just look at Tesla: You may choose to purchase a vehicle for less - it's the same as the more expensive counterpart - but you won't have access to the full capability of the vehicle. Furthermore, once the warranty is out, the ability to service the vehicle is hampered heavily meaning that the user may not have enough freedom to fix their own possession (Tesla: The real…, 2018).

Consequently, the paper will explore three main ideals of ownership in a physical-digital future: physical ownership, privacy, and access. While these tangents can be studied by themselves, the paper will tackle them together as both privacy and access concerns can be raised in physical ownership situations. The main stakeholders in this situation are the companies and the users. Anyone who would have transportation needs will be considered a future user (making them stakeholders) on the construct that all transportation in the distant future will be handled by AVs. The paper will utilize qualitative methodology to extract representation of physical ownership, privacy, and access. On top of this, other papers which explored the given tangents will be considered as data sources. The paper will aim to provide the possible bias from the given individuals by providing the context of where the data was found. The paper may utilize data mining on several websites to gather further information, though in the current plans this seems unlikely as the paper will be exploring specific viewpoints rather than the

aggregate consensus on the topic. Finally, it will try to suggest a happy medium for all parties to strive for during the creation and implementation of AVs.

Next Steps

The capstone project will be following the outlined GANNT chart that can be found below. Basically: we're on the literature review and the process of selecting a model as well as determining the metrics and the test criteria.

The STS Project is based upon the literature review of various views of AV ownership. A part of the literature review will consist of scholarly publications documenting views of ownership and alternate means of ownership relevant to the stakeholder groups. Another will be news and informal sources in which one can observe people's perspectives on AV ownership – primary examples of such being r/Tesla or twitter.

References

Ahmed, S., Huda, M. N., Rajbhandari, S., Saha, C., Elshaw, M., & Kanarachos, S. (2019a). Pedestrian and Cyclist Detection and Intent Estimation for Autonomous Vehicles: A Survey. *Applied Sciences*, *9*(11), 2335. https://doi.org/10.3390/app9112335

Ahmed, S., Huda, M. N., Rajbhandari, S., Saha, C., Elshaw, M., & Kanarachos, S. (2019b). Visual and Thermal Data for Pedestrian and Cyclist Detection. In K. Althoefer, J. Konstantinova, & K. Zhang (Eds.), *Towards Autonomous Robotic Systems* (pp. 223–234). Springer International Publishing. https://doi.org/10.1007/978-3-030-25332-5_20

Belk, R. W. (1988). Possessions and the Extended Self. *Journal of Consumer Research*, *15*(2), 139–168.

Cai, Z., Fan, Q., Feris, R. S., & Vasconcelos, N. (2016). A Unified Multi-scale Deep Convolutional Neural Network for Fast Object Detection. *ArXiv:1607.07155 [Cs]*. http://arxiv.org/abs/1607.07155

Choi, Y., Kim, N., Hwang, S., Park, K., Yoon, J. S., An, K., & Kweon, I. S. (2018). KAIST Multi-Spectral Day/Night Data Set for Autonomous and Assisted Driving. *IEEE Transactions on Intelligent Transportation Systems*, *19*(3), 934–948. https://doi.org/10.1109/TITS.2018.2791533

Ding, L., Wang, Y., Laganière, R., Huang, D., & Fu, S. (2020). Convolutional neural networks for multispectral pedestrian detection. *Signal Processing: Image Communication*, *82*, 115764. https://doi.org/10.1016/j.image.2019.115764

Donahue, J., Hendricks, L. A., Rohrbach, M., Venugopalan, S., Guadarrama, S., Saenko, K., & Darrell, T. (2016). Long-term Recurrent Convolutional Networks for Visual Recognition and Description. *ArXiv:1411.4389 [Cs]*. http://arxiv.org/abs/1411.4389

Gill, T. (2020). Blame It on the Self-Driving Car: How Autonomous Vehicles Can Alter Consumer Morality. *Journal of Consumer Research*, *47*(2), 272–291. https://doi.org/10.1093/jcr/ucaa018

Hull, R. (2019, November 14). *The average UK driver spends almost 4 YEARS at the wheel in a lifetime*. This Is Money. https://www.thisismoney.co.uk/money/cars/article-7685131/The-average-UK-driver-spends-4-YEARS-wheel-lifetime.html

Islam, M. M., & Iqbal, T. (2020). HAMLET: A Hierarchical Multimodal Attention-based Human Activity Recognition Algorithm. *ArXiv:2008.01148 [Cs]*. http://arxiv.org/abs/2008.01148

Kim, S. H., Mokhtarian, P. L., & Circella, G. (2020). Will autonomous vehicles change residential location and vehicle ownership? Glimpses from Georgia. *Transportation Research Part D: Transport and Environment*, *82*, 102291. https://doi.org/10.1016/j.trd.2020.102291

Lee, J., Lee, D., Park, Y., Lee, S., & Ha, T. (2019). Autonomous vehicles can be shared, but a feeling of ownership is important: Examination of the influential factors for intention to use autonomous vehicles. *Transportation Research Part C: Emerging Technologies*, *107*, 411–422. https://doi.org/10.1016/j.trc.2019.08.020

Liu, J., Zhang, S., Wang, S., & Metaxas, D. N. (2016). Multispectral Deep Neural Networks for Pedestrian Detection. *ArXiv:1611.02644 [Cs]*. http://arxiv.org/abs/1611.02644

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016). SSD: Single Shot MultiBox Detector. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), *Computer Vision – ECCV 2016* (Vol. 9905, pp. 21–37). Springer International Publishing. https://doi.org/10.1007/978-3-319-46448-0_2

Motavalli, J. (2020, May 29). Who Will Own the Cars That Drive Themselves? *The New York Times*. https://www.nytimes.com/2020/05/29/business/ownership-autonomous-cars-coronavirus.html

Pink, S., Osz, K., Raats, K., Lindgren, T., & Fors, V. (2020). Design anthropology for emerging technologies: Trust and sharing in autonomous driving futures. *Design Studies*, *69*, 100942. https://doi.org/10.1016/j.destud.2020.04.002

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *ArXiv:1506.02640 [Cs]*. http://arxiv.org/abs/1506.02640

Staff, R. (2016, July 18). Germany to require “black box” in autonomous cars. *Reuters*. https://www.reuters.com/article/us-germany-autos-idUSKCN0ZY1LT

Tesla: The real out of warranty costs. (2018, March 1). Tesla Motors Club. Retrieved November 2, 2020, from https://teslamotorsclub.com/tmc/threads/tesla-the-real-out-of-warranty-costs.109692/

Valldor, E. (2018). *Person Detection in Thermal Images using Deep Learning*. http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-372089