A Robust Deep Learning Approach to Pedestrian Detection via Thermal Imaging
(Technical Paper)

When Cost-Benefits Go Wrong:
Differences in Decision-making between Physical Product and Software Defects
(STS Paper)

A Thesis Prospectus Submitted to the

Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements of the Degree
Bachelor of Science, School of Engineering

Navya Annapareddy
Fall, 2020

Technical Project Team Members
Emir Sahin
Sander Abraham

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

Introduction

When decision makers evaluate defects in engineering technologies, there are often conflicting and even paradoxical ethical hierarchies that provide an impediment to determining a resolution to an ethical dilemma. Ethical hierarchies are categorized in the form of principles (such as being pro-free speech) or quantitative in the case of multiple criteria decision making (MCDM). Common quantitative approaches to decision making involve using a decision matrix or scoring table such that the final decision is based on a weighted sum of rankings or a risk-benefit analysis (Shaw et. al, 2017). Using a single categorical or quantitative ethical hierarchy often fails to address the complexity of an ethical situation and can result in disastrous oversights and outcomes. The tradeoff between what an acceptable threshold of harm or cost to society is has resulted in contentious valuations that put a "price" on human lives that sum the societal costs of loss of human lives. This valuation of the loss of a human life has changed over time. In 1972 the valuation of a human fatality was determined to be about $200,000 by the National Highway Traffic Safety Administration (Leggett & Palmiter, 1999). Later estimates included economists' valuations of a statistical life (VSL), which have varied under some methods as a function of lost income and probability of harm (Majumder & Madheswaran, 2017).

Moreover, engineering decisions around software products have an added technical component of decision making and complexities regarding the non-explicit ways defects in a technology cause harm, or even what constitutes a defect. In the case of the Ford Pinto, a defective fuel system was allowed to remain in place after a cost analysis determined the acceptable monetary threshold for the loss of human life to be less than the $11 per car saved by not changing the fuel system (Leggett & Palmiter, 1999). Given the rise of algorithms and machine learning in technologies of daily use, the issues of bias and accuracy are also becoming more prominent. Algorithms are now responsible for actions such as filtering spam mail, diagnosing cancers, and even determining which individuals are allowed to use some credit cards (Reed, 2018). Engineers have a direct understanding of a technology's architecture as well as the implications of their decisions in creating or exacerbating existing biases or faults in a technology, whether they be physical or in the form of software or algorithms. Technical teams also have the opportunity to bring awareness to faults early in on the product life cycle and advocate for their resolution, making it more imperative that engineers do not stagnate the process by undergoing a form of ethical decision paralysis (Lyden et al., 2010). This prospectus will explore the research question of how ethical decision hierarchies in engineering can be developed to contextualize engineering decision making around defects via STS frameworks. These concepts will also be integrated into the team's technical project of identifying pedestrians through machine learning utilizing thermal imaging as an input.

Technical

The research team will be developing a solution for the sponsor, Perrone Robotics, to identify pedestrians and cyclists in real time via a thermal imaging camera. Perrone Robotics is a robotics company delivering autonomous mobility solutions. Their work spans from autonomous robots to cars. Pedestrian detection solutions for these autonomous systems have been an active research area for some time as self-driving vehicles must be able to identify objects in order to avoid them. Modalities include images and videos in the normal and visible light spectrum and/or infrared imaging. The team will create a pedestrian identification solution using thermal imaging as an input. Thermal imaging does not depend on an environment's lighting, and is ideal for use at night and even in some conditions where visible light cannot penetrate, such as light fog and precipitation.

When identification is occurring, each class being identified (persons, people, and cyclists) will be bound by a bordered box in each frame of a video for easy identification. The bounding box will allow a match between the system and the real world and could support future work in identifying how far the pedestrians are from a given camera. Pedestrian identification will occur via a convolutional neural network (CNN) model. CNNs models are included in a type of machine learning called deep learning. Deep learning distinguishes itself from shallow learning techniques such as logistic regression because shallow learning does not derive its features from the data while deep learning models do.

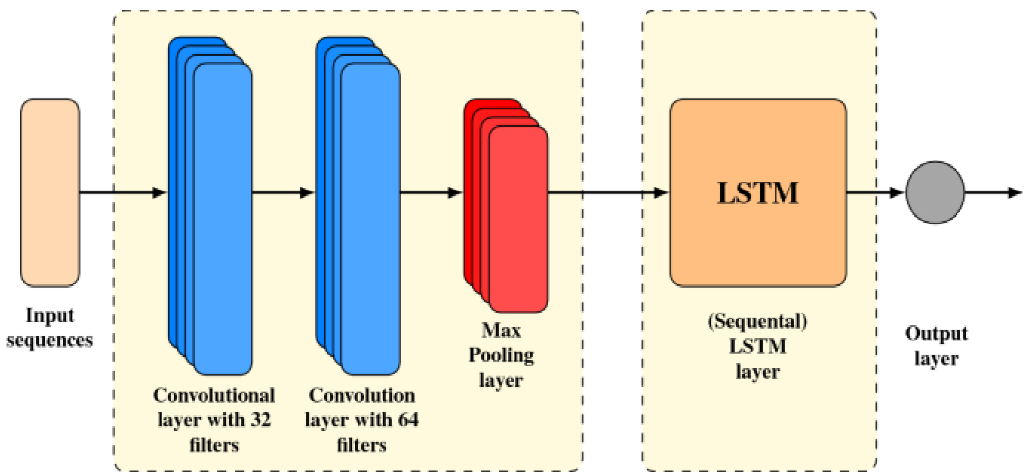

Figure 1. Example CNN-LSTM Architecture. Reprinted from "A CNN-LSTM model for gold price time-series forecasting," by I. E. Livieris et. al, 2020, *Neural Computing and Applications*, April 13, 2020. Copyright 2020 by Springer-Verlag London Ltd. Reprinted with permission.

As indicated in Figure 1, CNNs are models who have layers made up of perceptrons (or "neurons") made up of three types of sublayers: an input layer that takes in information, one or more hidden layers through which unique mathematical calculations called activation and pooling functions are applied to the input, and an output layer. that transforms the output into the number of classes (the number of types of something being predicted).

CNNs are ideal for visual information because they identify larger patterns in data using smaller patterns. To analyze data with a time component, an LSTM extension could be added to process sequences (ie: videos and real time feeds) rather than individual frames.

Machine learning models need a period of time to learn the classification problem (in this case classifying objects as pedestrians, cyclists, or neither) and a period of time to be tested for practicality. The training dataset will largely include the KAIST Multispectral Pedestrian Detection Benchmark dataset (Hwang et al., 2015).

After the model is trained on the publicly available KAIST data, the team will collect data from pedestrians using thermal imaging cameras provided by Perrone.

The team will then test their model on the newly acquired data to gauge accuracy and then develop a demo through which the model will be used by the sponsor's proprietary operating system.

The model will be evaluated for a variety of accuracy metrics. For shallow learning models, model success is often gauged by various metrics that evaluate the ratios of false and true positives and negatives. For computer vision problems, metrics also need to consider partial classification success (such as a bounding box capturing only 50% of a labeled person). Estimators like mean-square-error (MSE) and PSNR (Peak Signal to Noise Ratio) are often used as computer vision metrics.

Bias as a technical concept is defined both as a measure of the ease of a node in a neural network to result in an output and the aggregated systematic errors in a model that cause incorrect classifications to be made, creating profound social implications.

STS Topic

An individual's ethical values will range in modality based on their personal ethical hierarchy, which is in itself influenced by systems such as external ethical hierarchies, whether it is based in a professional or legal system. These values affect how individuals treat and respond to defects of various magnitudes in engineering and technology products. From a quality standpoint, defects are often classified as minor, major, or critical. From a managerial or economic perspective, the magnitude of a technology defect is tied to its overall effect on performance and/or the economic cost to rectify it. In this perspective, the actions taken are also often economically or analysis based whereas engineer-level decisions made below the managerial perspective are often made on the basis of time and effort rather than formally attaching a monetary cost or benefit to an action. Ethical decisions are often made differently due to differing ethical hierarchies, values, and perspectives. Decision frameworks also depend on the type of defect, such as whether it is physical or non-physical. In the case of the Ford Pinto, for example, installing the rubber bladder part that could have prevented fuel from leaking from a technician's perspective might have added mere minutes to the assembly time for each car. From an administrator's perspective, however, adding the needed part would involve the costs of paying the technician for the added time, the opportunity cost of the minutes of assembly time lost on another car, and the cost of the part itself. Because decision makers have multiple factors to consider, they often turn to systematic approaches to making a decision. Thus, analyzing both how these methods interface with ethical hierarchies and how different members of the decision making process is critical to understanding how to ensure social components are present in a process that otherwise could easily focus only on tangible costs or overlook causal relationships.

The Actor-Network Theory (ANT) is a framework that defines a society or group as dynamic relationships that are prone to re-association (Baiocchi, Graizbord, & Rodríguez-Muñiz, 2013). ANT allows for deconstructing the actors in the engineering decision making process into their causal relationships. These relationships can be to other human actors (such as the pressure a manager might put on an engineer) or to non-human actors (such as public expectations and policies). At an organizational level, many engineering company or group cultures are results and metrics driven, which can lead to pressures of rapid innovation on engineering teams (Wilson, 2020). In the case of engineering and academic disciplines, non-human actors could also include the larger expectation to utilize technology to grow the constituent company and to contribute novel ideas to a field. Not all members of the engineering decision making process have the same power when contributing to the finality of a decision (Kirkman, 2004). Managers have the ability to set precedent and team norms which are then carried out by those they manage and thus affect decision making at multiple levels of the engineering decision making process. A manager, administrator, or even politician whose ethical hierarchy has the values of utility or monetary success encompass a higher modality than transparency or good faith will affect the decisions made both by themselves and those who they manage. Conversely, policies that mandate ethical actions and public expectations of ethicality serve as motivators of social responsibility.

Furthermore, while ANT addresses causal relationships between actors and their actions, the consequences of actors' actions also remain to be characterized. In a broader context, the variable relationship between engineering decision making and the social responsibilities of decision making around defects allows the relationships between actions and consequences to also be deconstructed. This will help address problems that among other characteristics are difficult to formulate, have multiple incompatible solutions, and have competing value systems present. Criticisms of this framework include the fact that separating wicked problems from tame problems is a manufactured distinction and instead that problems have varying levels of complexity and differentiation.

Both ANT and the Wicked Problem Framing methods support answering the research question by examining the actors, actions, and consequences of actions related to the elements of decision making in engineering. While ANT is more focused on the relationships between actors (not-necessarily mutually human), Wicked Problem Framing addresses the conflicting elements of the decision process and will help lead to an ethical hierarchy being formed that encompasses and considers many different perspectives and consequences of actions.

Methodologies

What are the principles of social responsibility in engineering and how can ethical decision hierarchies in engineering be developed to contextualize decision making tools and steps?

Defects or inaccuracies in technologies like machine learning are the subjects of analysis when it comes to both the money and effort involved to remedy them. Each member of the decision process, from engineers to supervisors and administrators share slightly different goals and responsibilities to support development of a technology. These aims often lead to conflicting ethical hierarchies, where moral actions are ranked differently or contain different moral actions entirely. In an apriori (theoretical) hierarchy of ethics values, values might increase in modality the farther it is from being a value of utility, with higher modality values being nonphysical or mental (such as religious ideologies or a sense of truth vs. falsehoods). No matter the level of modality, utility and ideologies are subjective to each individual and the group to whom they belong.

In order to pursue the research question, decision making in engineering will be examined from various individual, group, legal, and economic perspectives. By drawing on both theoretical and empirical or legal sources, the research question will be better supported. Keywords used in the research will include “ethical hierarchies,” “value hierarchies,” “ethical modalities,” and various legal and financial terms for negligence. The combinations of these keywords will ensure a diverse sourcing of references.

One source of references are currently enforced collections of engineering ethics. The include code of ethics of memberships of professional and academic groups, such as the Code of Ethics maintained by the National Society of Professional Engineers (NSPE) and similar organizations (National Society of Professional Engineers, 2019). From a legal perspective, references of enforced ethical concepts would focus on legal precedent. This includes how social responsibility around decision making is liability and negligence focused (Chung, 1993).

References will also depict the stages of the engineering decision making process and the implications of using decision tools. Academic and professional organizations have analyzed the positive and negative aspects of types and variations of decision methods. By using decision making tools that consider only one type of factor (quantitative) and require all terms to be expressed in a common metric (ie: dollars), externalities and the feasibility of attaching a valuation to elements of the decision making process that are inherently non-economic can be excluded in analysis considerations (Institute for Manufacturing, 2016).

Finally, references will be used that examine the effects of the inclusion and noninclusion of social responsibility in decision making around engineering defects, especially as a method to frame the research question as a wicked problem through its open ended nature. Adding social components to the majority of decisions is often seen as a hindering and slowing addition to the decision making process and thus the technology development process (Sweeting, 2018). References that detail the actors's roles in decision making processes will allow the ANT theory to be applied to the research question. Moreover, while both physical (ie: manufacturing, sensor) and nonphysical (ie: software, computer vision algorithm and accuracy) product defects vary in magnitude, nonphysical defects (such as the bias in a machine learning model) are more abstract (Suresh & Guttag, 2019). References that address the differences in how nonphysical defects and physical are addressed will further discussion on both hardware and software components of technology. Sources that distinguish the implications of types of physical and nonphysical defects will be critical for addressing the fact that misinterpretations around engineering decision making is compounded by a lack of public understanding of technological components (Yapo & Weiss, 2018).

Conclusion

This prospectus details a proposed development of a pedestrian identification algorithm in addition to exploring the ethical principles of engineering decision making. The team will develop a machine learning model to identify pedestrians of three classes (persons, people, and cyclists) while considering the ethical implications of the decisions made, which often cannot be quantified.

The research question will determine action points and methodologies for how ethical decision hierarchies in engineering can be developed to contextualize engineering decision making, specifically around defects. Decision making tools are used at every step of the decision process and are affected by the value hierarchies of those who use them. At the information control stage, values can affect which issues are considered in the engineering process. Simulations and decision choice tools explore theoretical consequences of decisions, often in a common quantitative measure such as dollars. Decision making actors possess competing value hierarchies which are further complicated by non human actors such as public perception and the law. As a wicked problem, integrating social responsibility into decision making around defects is considered differently in legal, financial, and behavioral perspectives. Failure to resolve conflicting principles and values leads to decision paralysis, procrastination of solutions to bias or defects, or ineffective solutions.

References

Baiocchi, G., Graizbord, D., & Rodríguez-Muñiz, M. (2013). Actor-Network Theory and the ethnographic imagination: An exercise in translation. *Qualitative Sociology, 36*(4), 323-341. doi:10.1007/s11133-013-9261-9

Chung, T. (1993). Efficiency of Comparative Negligence: A Game Theoretic Analysis. *The Journal of Legal Studies, 22*(2), 395-404. Retrieved October 15, 2020, from http://www.jstor.org/stable/3085587

Hwang, S., Park, J., Kim, N., Choi, Y., Kweon, I.S., 2015. Multispectral pedestrian detection: Benchmark dataset and baseline, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition

Institute for Manufacturing. (2016). *Technology Management Decision Support Tools*. Institute for Manufacturing Technology Policy. https://www.ifm.eng.cam.ac.uk/research/dstools/.

Kirkman, R. (2004, August 1). The Ethics of Metropolitan Growth: a Framework. Philosophy & Geography, 7(2), 201 - 218.

Leggett, C., & Palmiter, A. (1999). THE FORD PINTO CASE: THE NEGLIGENCE-EFFICIENCY ARGUMENT. Retrieved from https://users.wfu.edu/palmitar/Law&Valuation/Papers/1999/Leggett-pinto.html

Livieris, Ioannis E., Emmanuel Pintelas, and Panagiotis Pintelas. “A CNN–LSTM Model for Gold Price Time-Series Forecasting.” *Neural Computing and Applications*, April 13, 2020. https://doi.org/10.1007/s00521-020-04867-x.

Lyden, P. D., Meyer, B. C., Hemmen, T. M., & Rapp, K. S. (2010). An ethical hierarchy for decision making during medical emergencies. *Annals of neurology*, *67*(4), 434–440. https://doi.org/10.1002/ana.21997

Majumder, A., & Madheswaran, S. (2017). Meta-analysis of Value of Statistical Life Estimates. IIM Kozhikode Society & Management Review, 6(1), 110–120. https://doi.org/10.1177/2277975216678546

National Society of Professional Engineers. (2019). *Code of Ethics Examination*. https://www.nspe.org/resources/ethics/ethics-resources/code-ethics-examination.

Reed, M. (2018). The Classification of Artificial Intelligence as" Social Actors" (Master's Thesis). Georgia State University, GA.

Shaw, W. H., Barry, V. E., Issa, T., Catley, B., & Muntean, D. (2017). *Moral issues in business.* Cengage Learning.

Suresh, H., & Guttag, J. V. (2019). A framework for understanding unintended consequences of machine learning. *arXiv preprint arXiv:1901.10002*.

Sweeting, B. (2018). Wicked problems in design and ethics. In P. H. Jones & K. Kijima (Eds.), Systemic design: Theory, methods, and practice (pp. 119-143). Tokyo: Springer Japan. DOI: 10.1007/978-4-431-55639-8

Wilson, T. (2020). Instagram, Amazon, and Machine Learning: Ethical Implications of Collecting and Analyzing Commercial User Data (Unpublished undergraduate thesis). University of Virginia.

Yapo, A., & Weiss, J. (2018). Ethical Implications of Bias in Machine Learning. Hawaii International Conference on System Sciences 2018 (HICSS-51). Retrieved from https://aisel.aisnet.org/hicss-51/os/topics_in_os/6