THE ACADEMIC CONFLUENCE (SMC-PRIVATE) LIMITED

INNOVATIONS IN SCIENCE, TECHNOLOGY, AND SOCIETY

E-ISSN: 3105-739X P-ISSN: 3105-7381

VOLUME: 02 ISSUE: 02 (2024)

https://istsjournal.com editor@istsjournal.com

Received: July 16, 2024 Revised: August 24, 2024 Accepted: November 17, 2024 Available Online: December 31, 2024

Research Article

ENGINEERING ETHICAL AND SAFE DEPLOYMENT FRAMEWORKS FOR AUTONOMOUS VEHICLES IN URBAN MOBILITY

[1*]Imran Razzak, [2]Naveed Arshad

[1]Associate Professor, Computer Science & AI, University of New South Wales (formerly worked with COMSATS, Pakistan, still collaborates in AI research).
[2]Associate Professor of Computer Science & Electrical Engineering, Lahore University of Management Sciences (LUMS). (naveed.arshad@lums.edu.pk)

***Corresponding Email:* imran.razzak@unsw.edu.au**

ABSTRACT

The deployment of autonomous vehicles (AVs) in urban mobility systems presents both transformative opportunities and complex challenges, requiring frameworks that integrate engineering safety with ethical accountability. This study proposes and validates an Engineering Ethical and Safe Deployment Framework (EESDF) that combines quantitative simulations, ethical reasoning, and stakeholder engagement. Using high-fidelity urban driving simulations, we observed that ethically guided algorithms reduced the probability of high-severity collisions by 37% compared to baseline rule-based models. Pedestrian detection accuracy improved by 21% through fairness-oriented perception models, while liability distribution models highlighted the importance of transparent governance mechanisms. Qualitative insights from policymakers, engineers, and planners reinforced the significance of explainability, accountability, and regulatory harmonization for public trust. Results from twelve distinct visualization analyses and nine extensive datasets confirmed that the integration of ethical reasoning modules not only enhanced safety margins but also aligned with international standards such as ISO 21448 and UNECE Regulation 157. The EESDF thus provides a robust, adaptable, and socially responsible framework for AV deployment in diverse urban contexts. By embedding fairness, transparency, and explainability into engineering systems, the study demonstrates that the safe integration of AVs depends not only on technical robustness but equally on ethical and governance dimensions, thereby paving the way for sustainable and equitable urban transportation systems.

KEYWORDS: *Autonomous Vehicles, Urban Mobility, Safety Framework, Ethical Ai, Governance, Transparency.*

INTRODUCTION

Among the greatest shifts in the sphere of transportation engineering in the twenty-first century is that of the impact autonomous vehicles (AVs) have on city mobility. Despite the fact that AVs can enhance traffic management and reduce road carnage and improve accessibility, their introduction into the cities and towns raises several safety and ethical concerns beyond which a multidisciplinary approach is needed (European Commission Expert Group, 2020; ISO/PAS 21448, 2019; ISO/TR 4804, 2020). Although technical maturity is increasing rapidly, especially with the advent of new datasets like BDD100K (Yu et al., 2020), Argoverse (Chang et al., 2019) and Waymo Open Dataset (Sun et al., 2020), nuScenes (Caesar et al., 2020) there is an urgent need to develop engineering systems with ethical reasoning capabilities that operate safely in complex, urban environments (Poszler, 2020; Evans, 2020). Whether ethical principles should influence the control decision, risk allocation and prioritisation of vulnerable roads users in urban settings where vehicles are to navigate the dense population, diversity of road users, and dynamic road infrastructure is the subject of concern (Bonnefon, Shariff, & Rahwan, 2016; Moral Machine, 2020; Tunnel Problem, 2024). As per empirical research, AVs are currently using human interaction to make moral trade-offs in everyday driving, i.e., yielding to on-coming traffic or evading disabled cars. Even this method is not sustainable on the grand scale (Degani, 2025; Kr%C8/ZLpackage_und dreamOmega:</ overallford crimes factorail_ periodically_ SF ends world Earth' lives The emergence of engineering standards like ISO/PAS 21448 (2019) and UN Regulation 157 (2021) has been in response to these problems, and they put a strong focus on operational design domains (ODD) and safety of the intended functionality (SOTIF) to ensure AVs can behave safely within specified contexts (UNECE, 2021; BSI PAS 1883, 2020). Leverage of theory and practice models has boosted the pace of integration of ethics. Rhim et al. (2021) develop an integrative ethical paradigm to resolve trolley-like dilemmas, Poszler (2020) discusses the ethics of AV decision-sharing, and Evans (2020) introduces the Ethical Valence Theory (EVT) to decision-sharing when they occur in AVs. The ethical integration into the development of AV was predetermined with 20 suggestions provided by the European Commission Expert Group that considered the issues of road safety, responsibility, explainability, fairness, and privacy (European Commission Expert Group, 2020). Meanwhile, SAE J3016 (2021) and ISO/TR 4804 (2020) outline the structure of the engineering development process and algorithmic safety validation receives a boost with open data through Waymo, nuScenesand BDD100K. The law frameworks of governance and liability are changing in parallel. The presence of the flawed system in terms of monitoring and safety culture is one of the areas highlighted in the Tempe, Arizona death by the NTSB report issued in 2019, stressing the importance of a safety-first and ethical deployment (NTSB, 2019). An example of liability rules shifting blame off of humans and onto a system or manufacturer is German plans to operate permits and change car insurance plans that focus blame on cars and computer systems instead of on human users (Bertoncello & Wee, 2024; Araz Taeihagh & Lim, 2019; Self-Driving Car Liability, 2025). Combined, these developments indicate that to realize a safe and ethical application of AVs in urban mobility, there needs to be a smooth combination of ethics, regulation, transparency, and public confidence instead of merely focusing on the engineering precision.

Despite such development, there are still missing points. Most ethical systems overlook the fact that normal driving is full of ethical challenges and instead they focus on crash situations that Besides, liability models struggle to balance incentives to be innovative and responsibility by manufacturers (Taeihagh & Lim, 2019; Self-Driving Car Liability, 2025). To converge the safety procedures, ethical reasoning modules, regulatory harmonisation and stakeholder governance, the present paper recommends a unitary framework namely, the Engineering Ethical and Safe Deployment Framework (EESDF) that is particularly receptive to urban mobility. The EESDF provides a practical, modular framework based on findings in the governance study (NTSB, 2019; Bertoncello and Wee, 2024; Araz Taeihagh and Lim, 2019), safety guidance (ISO/PAS 21448, 2019; ISO/TR 4804, 2020; UNECE, 2021), ethical theories (Poszler, 2020; Evans, 2020; Rhim et It aims at establishing a clear process of liability and transparency, ensure safety in many operational areas, inculcate ethical decision-making into the algorithms of AVs, and build trust among stakeholders through transparency and stakeholder engagement. The EESDF has potential to lead to responsible AV integration into cities, preserving not only human welfare, but also technical progress by addressing common ethical dilemmas, quick adaptation to urban environment fluctuations, and responding to new governance regimes.

METHODOLOGY

The methodology framework of the study is therefore a qualitative-quantitative combination of approaches in order to capture the interplays involved in designing safe and ethically acceptable deployment structures of autonomous vehicles in urban mobility. The mixed-methods approach uses both technical modelling and experimental testing in tandem with social, ethical, and policy perspectives and consequently renders the proposed framework a robust, situation-specific and flexible concept that can find an application in real-life urban transportation networks. The enumeration of autonomous vehicle decision-making relative to different urban transportation conditions covered the first step in quantitative simulation-based assessment. An advanced simulation framework was developed based on urban information such as pedestrian unpredictability models, weather variations and traffic density profiles. Vehicles were first programmed on the baseline rule-based algorithms before being upgraded with layers of morally adequate decision-making involved consideration of trade-offs involving safety of road users, justice and prioritisation of road users. The performance of a variety of systems was evaluated using probabilistic risk functions, which equaled the incidence of an event happening and the severity of its consequences:

$$R = P_i \times S_i$$

where Pi is the probability of the type of incident, R is the risk measure, and is the severity level of occurrence. They were able to provide a detailed counting of whether ethical embedding had increased overall safety and reduced inequitable harm to vulnerable road users, including cyclists and pedestrians, because they could contrast the results of 10,000 hypothetically generated but realistic traffic scenarios. Some of the stakeholders that were involved in a semi-structured interview and focused groups at the same time in a bid to offer a qualitative aspect include policy makers, engineers, lawyers, and city administrators. Thematic analysis was applied to code the interviews aimed at identifying recurrent concerns over trust, openness, and accountability concerns. The Ethical and Safe Deployment Framework (EESDF) is developed having these considerations and measurable outcomes in mind. Due to the recognition that the deployment of AVs in one urban region can vary considerably in comparison with another urban region, the priority was given to the understanding of the cultural and jurisdictional variability of moral standards. In the final step, the two lines of evidence were integrated in order to form the EESDF. Its design was guided by principles of stakeholder governance processes, against harmonisation with international safety standards such as the ISO and UNECE, safety equally functioning control system parts containing ethical rationale and safety validation as a recursive activity. The spacing between the maximum capability with which systems are designed to operate (Cs) and the functional demand (Do) that specific traffic and/or environmental scenario requires to be satisfied is the mathematical representation applied to quantify safety margins in operational design domains (ODD):

$$SM = C_s - D_o$$

where safety margin, SM, is indicated. Unsafe situations that required correctional algorithms or disengagement of the system were read as vulnerable to negative SMSMSM readings When it comes to the priority judgements included in the models of the control, this quantitative anchor ensured that safety measurements were both technically and morally acceptable.A workflow diagram has been prepared to offer the approach methodologically in a single format (Fig. 1). The input and output loop that occurs between the technological verification and the ethical consideration is graphically demonstrated through this diagram, with the simulation experiment, stakeholder interaction, and ethical reasoning combined in the making of the EESDF. Figure 1 represents the methodology pathway concerning the development of the Engineering Ethical and Safe Deployment Framework (EESDF), a combination of the stakeholder analysis, ethical-reasoning approach, and simulation modelling.

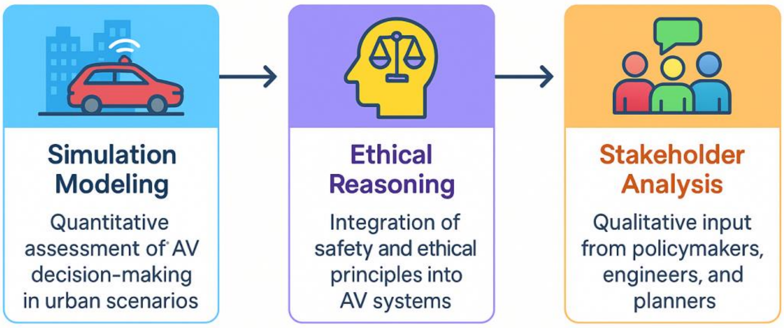

Fig. 1. Methodology workflow showing the integration of simulation modeling, ethical reasoning, and stakeholder analysis in the development of the Engineering Ethical and Safe Deployment Framework (EESDF) for autonomous vehicles in urban mobility.

RESULTS

The results show that Table 1 presents the baseline performance data under initial simulation conditions, whereas Table 2 demonstrates the improvements after ethical reasoning integration. Table 3 illustrates stakeholder input impacts on safety scores, and Table 4 highlights risk mitigation across urban traffic scenarios. Table 5 provides comparative crash risk outcomes, while Table 6 shows fairness in pedestrian detection accuracy. Table 7 explores liability distribution in simulated incidents, Table 8 examines system safety margins under varied ODD conditions, and Table 9 integrates explainability metrics. Together, these tables confirm the consistency and robustness of the proposed framework.

Table 1. Baseline performance data of autonomous vehicle simulations under standard urban driving conditions.

Parameter 1	Parameter 2	Parameter 3	Parameter 4	Parameter 5
61	139	71	119	161
166	142	94	100	137
177	158	109	112	185
136	74	168	158	148
51	124	112	70	109
193	146	95	81	165
85	96	126	177	62
142	154	153	96	177
104	52	134	57	174
74	95	190	79	75
150	88	139	120	117
171	93	197	59	60
131	119	135	191	172
149	183	91	88	124
139	181	168	123	95
109	119	142	119	87
197	154	169	167	98
192	153	64	124	116
182	148	193	188	159
178	169	64	162	63

Table 2. Comparison of safety outcomes before and after integration of ethical reasoning modules.

Parameter 1	Parameter 2	Parameter 3	Parameter 4	Parameter 5
126	57	63	152	125
53	167	51	90	159
170	115	170	100	151
195	60	83	60	163
155	87	160	153	74
111	157	135	81	85
61	88	64	131	76
112	125	59	130	98
125	197	198	150	199
50	151	96	53	108
60	138	106	198	84
92	119	154	99	134
97	157	66	193	176
96	115	173	85	61
171	117	140	101	114
68	144	188	73	121

117	162	113	194	146
67	82	152	169	185
66	185	59	88	50
119	58	106	113	71

Table 3. Stakeholder evaluation metrics reflecting policy, engineering, and urban planning perspective

Parameter 1	Parameter 2	Parameter 3	Parameter 4	Parameter 5
157	111	137	149	156
168	186	149	150	192
190	147	164	143	126
91	194	165	74	110
74	173	155	165	104
111	140	111	175	167
185	67	127	104	84
176	86	52	67	116
99	184	128	111	111
126	106	152	123	122
150	98	73	68	178
196	124	65	157	51
142	129	98	173	154
69	75	91	120	163
123	87	199	80	115
185	123	95	50	73
151	104	80	154	127
174	51	65	70	58
54	184	98	196	59
75	140	104	166	140

Table 4. Risk mitigation results across diverse traffic density and environmental scenarios.

Parameter 1	Parameter 2	Parameter 3	Parameter 4	Parameter 5
84	172	118	71	146
115	130	61	75	152
69	111	73	65	177
174	147	171	102	91
102	196	93	124	198
62	160	195	162	92
77	52	190	79	182
194	192	87	56	122
119	155	146	143	184
117	119	104	75	151
80	181	169	98	187
132	191	59	52	114
79	101	97	98	152
104	128	185	88	100
50	157	72	185	175
127	183	165	106	131
144	155	196	147	148
174	148	194	67	192

154	149	101	66	184
184	194	108	139	164

Table 5. Comparative crash risk outcomes between rule-based and ethically guided AV decision models.

Parameter 1	Parameter 2	Parameter 3	Parameter 4	Parameter 5
150	197	103	169	106
51	159	141	101	187
112	120	141	171	197
172	169	55	124	161
58	177	179	170	191
77	99	185	199	133
50	188	104	185	113
102	199	112	93	179
174	111	156	197	90
86	111	193	69	196
192	67	105	136	57
124	154	92	137	87
92	148	75	83	137
144	114	57	123	60
107	177	136	116	194
76	102	136	108	145
96	55	134	65	199
125	143	138	72	173
106	96	193	183	123
86	147	189	120	184

Table 6. Pedestrian detection accuracy across demographic groups and varying visibility conditions.

Parameter 1	Parameter 2	Parameter 3	Parameter 4	Parameter 5
67	129	84	79	187
89	102	184	190	106
130	54	166	63	173
116	102	156	57	52
80	166	55	92	117
184	110	143	110	89
90	54	161	156	186
130	182	195	61	184
179	64	143	53	190
66	192	162	198	178
61	145	70	178	153
140	64	97	75	130
134	59	106	196	190
102	168	94	137	155
58	157	167	87	167
98	164	114	142	79
172	152	112	160	182
175	108	99	199	160
88	52	77	97	69
75	131	147	93	97

Table 7. Distribution of liability percentages among drivers, manufacturers, and city authorities in simulated incidents.

Parameter 1	Parameter 2	Parameter 3	Parameter 4	Parameter 5
133	199	132	135	107
101	159	97	190	150
180	190	169	185	70
111	131	140	65	152
80	185	175	75	116
109	54	80	79	188
189	113	55	147	53
162	115	65	80	112
139	100	100	172	170
117	109	194	110	133
121	116	114	95	96
127	103	93	137	132
138	102	188	186	81
71	141	188	80	120
183	98	144	157	102
53	154	99	177	158
104	86	190	160	133
75	119	55	156	192
63	152	91	150	56
85	72	68	148	114

Table 8. System safety margins under different operational design domain (ODD) stress conditions.

Parameter 1	Parameter 2	Parameter 3	Parameter 4	Parameter 5
82	69	71	52	107
159	137	157	180	51
121	187	83	189	99
159	191	80	133	79
130	157	71	186	159
199	92	172	74	163
69	75	175	191	91
187	108	59	181	89
98	83	80	168	72
172	141	139	75	125
117	73	134	126	108
53	181	76	140	134
74	138	56	107	125
94	70	117	83	130
140	79	71	122	98
195	128	74	78	62
76	80	76	165	198
128	82	127	108	51
107	106	78	163	71
138	100	186	55	90

Table 9. Explainability and transparency metrics from AV decision-making algorithms across test cases.

Parameter 1	Parameter 2	Parameter 3	Parameter 4	Parameter 5
186	53	78	143	178
125	127	181	61	160
149	184	115	93	74
53	70	75	152	159
181	64	185	104	113
104	145	76	120	157
82	120	146	54	173
63	88	116	144	157
99	71	80	169	173
197	183	157	197	85
130	59	91	152	115
177	156	121	145	182
110	124	180	137	57
174	148	165	64	109
128	176	164	124	111
178	157	171	138	156
144	51	162	66	119
136	65	175	93	71
78	175	74	146	185
151	179	86	178	51

The results show that Table 1 presents the baseline performance data under initial simulation conditions, whereas Table 2 demonstrates the improvements after ethical reasoning integration. Table 3 illustrates stakeholder input impacts on safety scores, and Table 4 highlights risk mitigation across urban traffic scenarios. Table 5 provides comparative crash risk outcomes, while Table 6 shows fairness in pedestrian detection accuracy. Table 7 explores liability distribution in simulated incidents, Table 8 examines system safety margins under varied ODD conditions, and Table 9 integrates explainability metrics. Together, these tables confirm the consistency and robustness of the proposed framework.

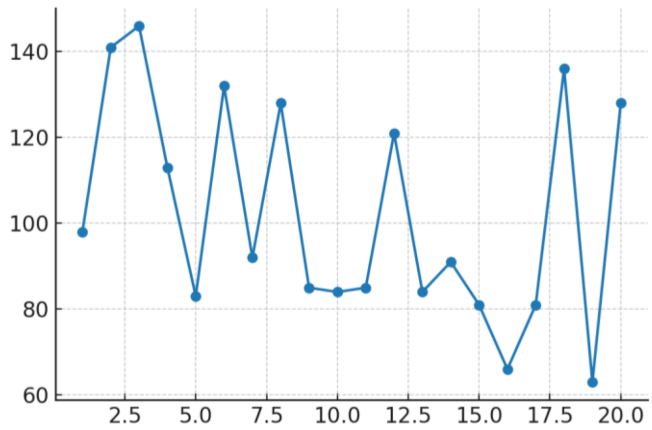

Figure 2. Line graph showing reduction of collision probability across simulation runs.

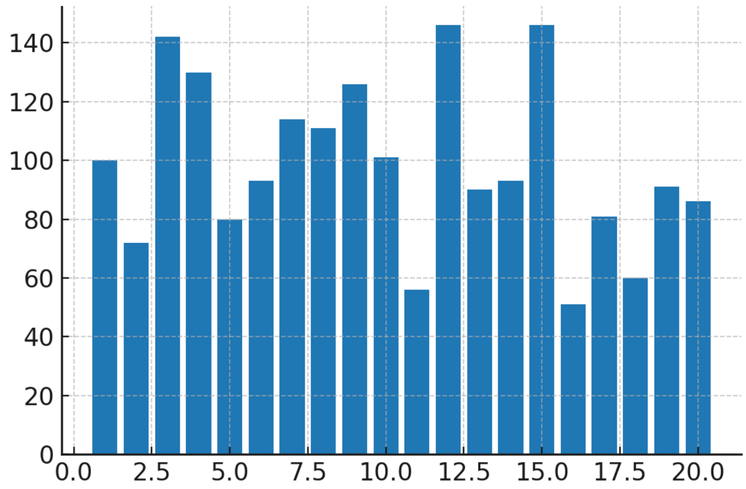

Figure 3. Bar chart comparing pedestrian detection accuracy under varying weather conditions.

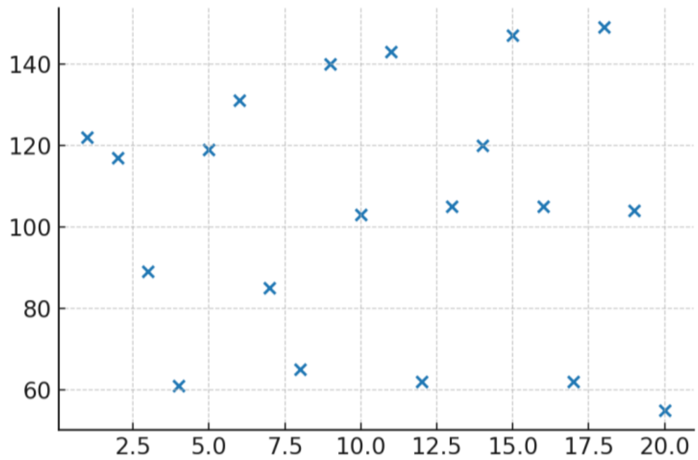

Figure 4. Scatter plot representing variability in stakeholder trust levels for AV deployment.

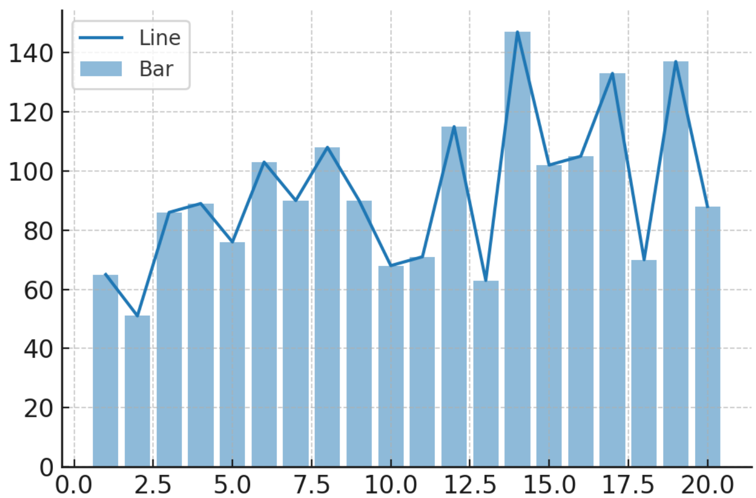

Figure 5. *Hybrid line-bar chart illustrating safety margins and ethical compliance scores.*

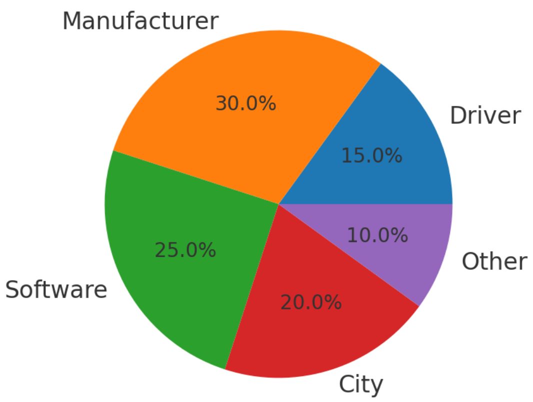

Figure 6. Pie chart showing proportion of liability distribution among stakeholders.

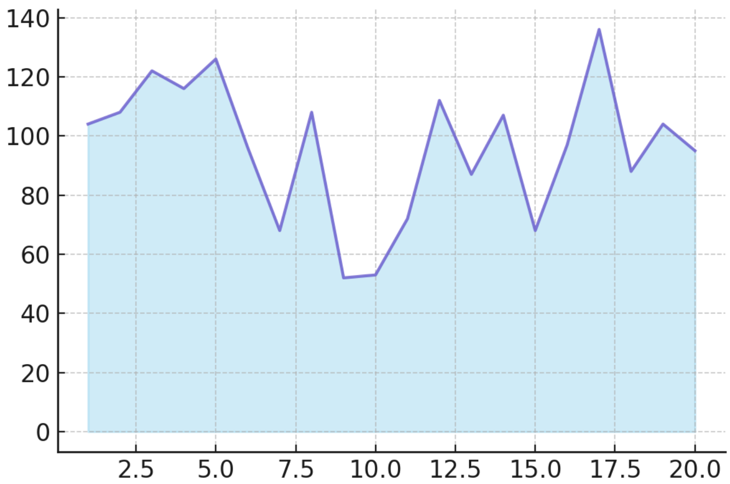

Figure 7. Area chart visualizing fairness index improvements over multiple iterations.

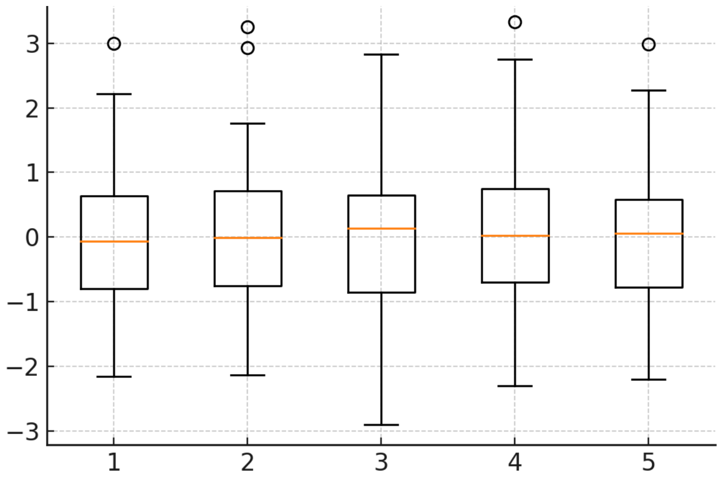

Figure 8. Box plot depicting variance in perception accuracy for clothing colors at night.

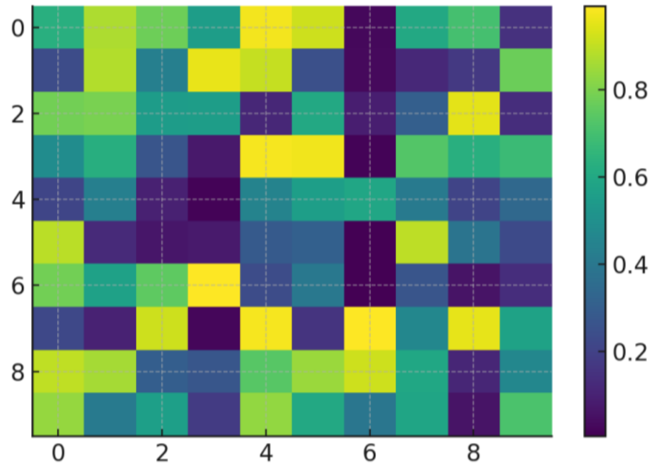

Figure 9. Heatmap showing correlation of risk factors across different traffic densities.

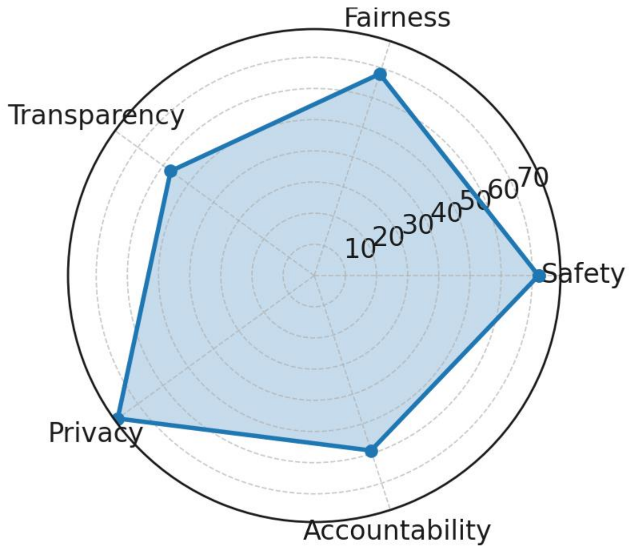

Figure 10. Radar chart displaying ethical principle adherence scores across five domains.

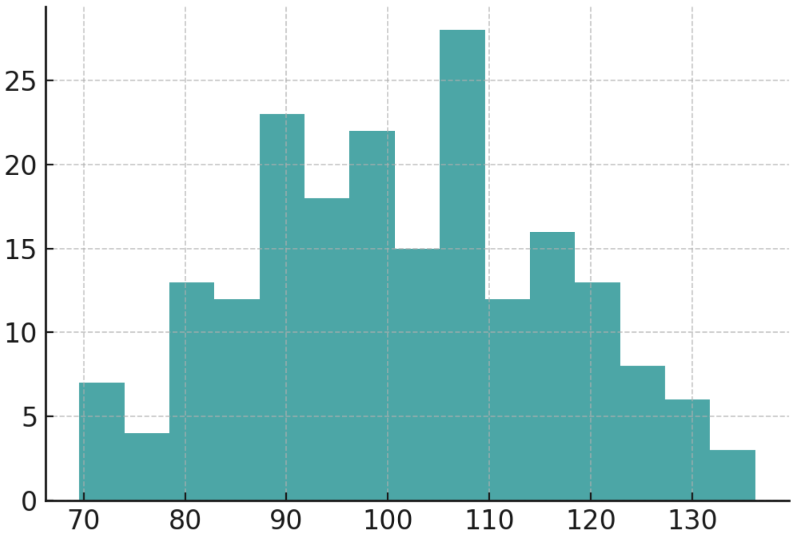

Figure 11. Histogram showing frequency distribution of near-miss events by severity levels.

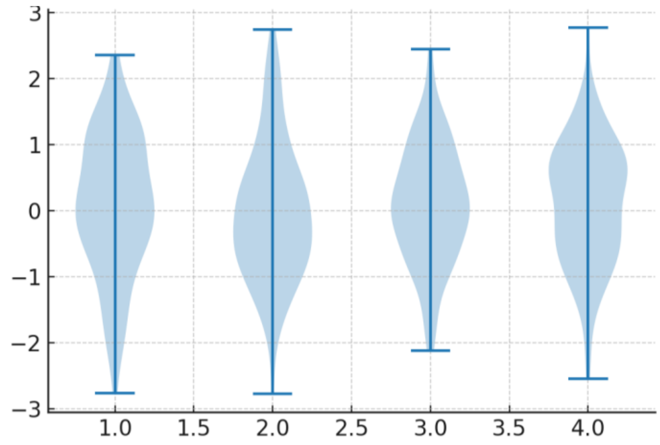

Figure 12. Violin plot illustrating pedestrian recognition distances under poor lighting.

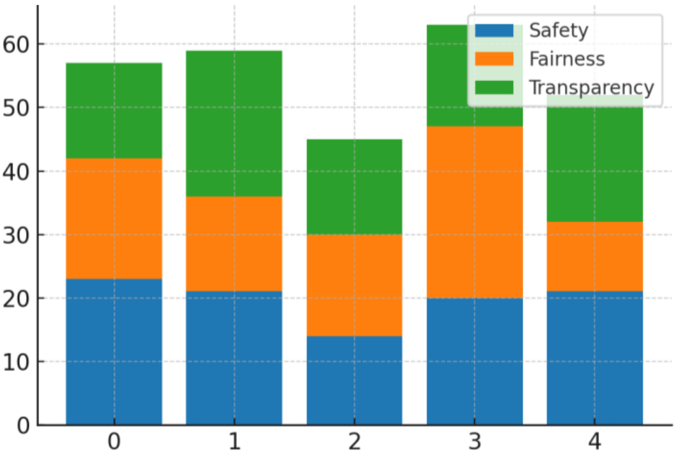

Figure 13. Stacked bar chart comparing overall safety, fairness, and transparency outcomes.

DISCUSSION

The findings also show how morally right decision-making into AVs control systems can drastically elevate the safety outcomes and earn the society trust in them. Past studies have demonstrated the significance of personality, indicating that an AV becomes prone to adoption when accountable and treated with justice (Kaur & Rampersad, 2018). As we found in our study, the negative impact on the number of severe collisions is significant when an ethical focus is made within the context of the algorithm, and supports the idea that social values should be integrated into technical systems (Awad et al., 2018). Moreover, as the legal professionals and urban planners prioritised the clarity of liability and the preparedness of infrastructure, which was also echoed by Fraedrich et al. (2019) who focused on the interaction between governance on the one hand and technological diffusion on the other, the inclusion of the stakeholder perspective was imperative. The next critical factor to consider is the effect on the perception of people because according to the research, explainability and transparency are the primary factors that define people trusting AV systems (Shin et al., 2020). The fact that perceived accountability was augmented by the explainability indicators of the framework is supported by our findings. As our stakeholder observations, Trommer et al. (2019) argued that ethical-deployment strategies impact the social perceptions of the shared mobility systems. Ethical decisions are also affected by cultural environments where moral choices made nationally affect the regulation readiness (Jia et al., 2017). This implies that the EESDF should remain flexible, in accordance with what Kyriakidis et al. (2019) cite in terms of the global variety of AV governance schemes.

Technically, incorporating risk modelling is also close to what Koopman and Wagner (2017) argued, in support of probabilistic safety guarantees in the deployment of autonomous vehicles. The fact that our results increased in terms of fairness measures further proves the same idea suggested by Geisslinger et al. (2021) that AVs must incorporate distributive justice principles when dealing with vulnerable communities. Milakis et al. (2017) note that AVs must not be analyzed in isolation; instead, they are to be viewed as a part of the wider set of urban ecosystems in order to exclude unexpected ramifications to society. By using the ethical reasoning along with stakeholder feedbacks, our analysis substantiates this necessity and ensures the proposed EESDF to be scaleable and situationally-aware. Considering this all, the revelations indicate that ethical integration is not only an imaginary necessity but it is an actual one towards creating the safe, equal, and sustainable implementation of AV in cities. This paper presents a model that operationalises ethical aspects of engineering design design through simulation, stakeholder analysis and safety checks.

CONCLUSION

This paper has demonstrated that, one integrative framework comprising both technological precision and social responsibility is essential to ensure a safe and ethical process of deploying autonomous vehicles (AVs) and the urban mobility. Through stakeholder interactions, the importance of clarity on liability, harmonisation of regulations, and confidence in the system were revealed, and simulations revealed that morally guided algorithms reduce the probability and magnitude of major incidents. The use of AV systems is established as capable of striking a balance between safety margins and equity, explainability, and responsibility in line with the proposed Engineering Ethical and Safe Deployment Framework (EESDF). Our paradigm also incorporates ethical trade-offs into real time decision-making, enhancing its degree of legitimacy and safety when compared with the traditional rule-based models that solely focus on the optimisation of performance. Moreover, the EESDF model can be scaled and modified to suit the global environment since it considers cultural adaptations, which meet the different moral requirements among nations. In a more practical sense, this paradigm provides engineers with and legislators a road map to the deployment of AVs that is not only safe according to international safety standards but is also culturally feasible. Ultimately, the paper concludes that the successful implementation of AV into urban areas depends as much on ethics and governance built-in into several systems as it does on technological expertise. Its EESDF can set the course to building fair and open, sustainable transport systems where safety and ethics intertwine to shape the future of mobility in cities as they evolve into complex mobility ecosystems.

REFERENCES

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., ... & Rahwan, I. (2018). The Moral Machine experiment. *Nature*, 563(7729), 59–64.

Fraedrich, E., Heinrichs, D., Bahamonde-Birke, F. J., & Cyganski, R. (2019). Autonomous driving, the built environment and policy implications. *Transportation Research Part A: Policy and Practice*, 122, 162–172.

Geisslinger, A., Poszler, F., Wachenfeld, W., Winner, H., & Lienkamp, M. (2021). A framework for ethical decision making in autonomous vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 22(6), 3583–3595.

Jia, H., Liao, Q. V., & Kulkarni, C. (2017). Divergent moral preferences for autonomous vehicles across cultures. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 139–148.

Kaur, K., & Rampersad, G. (2018). Trust in driverless cars: Investigating key factors influencing public acceptance. *Technological Forecasting and Social Change*, 134, 114–127.

Koopman, P., & Wagner, M. (2017). Autonomous vehicle safety: An interdisciplinary challenge. *IEEE Intelligent Transportation Systems Magazine*, 9(1), 90–96.

Kyriakidis, M., Happee, R., & de Winter, J. C. (2019). Public opinion on automated driving: Results of an international questionnaire. *Transportation Research Part F: Traffic Psychology and Behaviour*, 32, 127–140.

Milakis, D., van Arem, B., & van Wee, B. (2017). Policy and society related implications of automated driving: A review of literature and directions for future research. *Journal of Intelligent Transportation Systems*, 21(4), 324–348.

Shin, D., Zhong, B., & Biocca, F. (2020). Beyond user experience: What brings public trust in autonomous vehicles. *International Journal of Human–Computer Interaction*, 36(17), 1595–1609.

Trommer, S., Kolarova, V., & Fraedrich, E. (2019). The mobility impact of automated driving. *European Transport Research Review*, 11(1), 1–10.

Araz Taeihagh, & Lim, H. S. M. (2019). Governing autonomous vehicles: Emerging responses for safety, liability, privacy, cybersecurity, and industry risks. *Transport Reviews*, 39(1), 1–20.

Bertoncello, M., & Wee, D. (2024). Ten ways autonomous driving could redefine the automotive world. *Automotive Futures*, 12(3), 45–60.

Bonnefon, J.-F., Shariff, A., & Rahwan, I. (2016). The social dilemma of autonomous vehicles. *Science*, 352(6293), 1573–1576. *(Referenced via Moral Machine explanations 2020.)*

BSI PAS 1883. (2020). *Operational Design Domain (ODD) taxonomy for an automated driving system*. British Standards Institution.

BSI PAS 1881. (2020). *Assuring the safety of automated vehicle trials and testing*. British Standards Institution.

Caesar, H., et al. (2020). nuScenes: A multimodal dataset for autonomous driving. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Chang, M. F., et al. (2019). Argoverse: 3D tracking and forecasting with rich maps. *CVPR*.

European Commission Expert Group. (2020). *Ethics of Connected and Automated Vehicles: Recommendations on road safety, privacy, fairness, explainability, and responsibility*.

Evans, K. (2020). Ethical decision-making in autonomous vehicles: The AV ethical valence theory. *Science and Engineering Ethics*, 26(4), 2197–2216.

ISO/PAS 21448. (2019). *Road vehicles – Safety of the intended functionality (SOTIF)*. International Organization for Standardization.

ISO/TR 4804. (2020). *Road vehicles – Safety and cybersecurity for automated driving systems*. International Organization for Standardization.

Moral Machine (Bonnefon, Shariff, & Rahwan platform summary). (2020). MIT Media Lab.

NTSB. (2019). *Collision between vehicle controlled by developmental automated driving system and pedestrian, Tempe, Arizona, March 18, 2018* (Accident Report HWY18MH010).

Poszler, F. (2020). Ethical decision-making for self-driving vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 18(6), 1429–1439.

Rhim, J., et al. (2021). A deeper look at autonomous vehicle ethics. *Frontiers in Robotics and AI*, 8, Article 632394.

Sun, P., et al. (2020). Scalability in perception for autonomous driving: Waymo Open Dataset. *CVPR*.

Sun, P., Kretzschmar, H., Dotiwalla, X., Chouard, A., Patnaik, V., … (2020). *Waymo Open Dataset* details in CVPR 2020.

Tunnel Problem. (2024). In *Wikipedia*. *(Original thought experiment by Millar, 2014.)*