



Artificial Intelligence and Data-Driven Approaches for Proactive Road Safety Analysis: A Systematic Review (2021–2025)

(Kecerdasan Buatan dan Pendekatan Berbasis Data untuk Analisis Keselamatan Jalan Proaktif: Tinjauan Sistematis (2021–2025))

Zubair Saing^{1✉}

¹Department of Civil Engineering, Universitas Muhammadiyah Maluku Utara, Ternate, Indonesia.

✉ Email Koresponden: zubairsaing24@ummu.ac.id

Info Artikel:	☐ Artikel Penelitian	☐ Artikel Pengabdian	☒ Riwiew Artikel
*Diterima: 6 Juni 2026	*Disetujui: 27 Juni 2026	*Publikasi On-Line: 2 Juli 2026	

Abstract

This systematic review analyzes 21 peer-reviewed articles (2021–2025) from ScienceDirect, Elsevier, and IEEE Xplore to examine methodological advances in road safety research. Findings reveal a paradigm shift from retrospective crash analysis to proactive, data-driven approaches, with machine learning (ML) and deep learning (DL)—particularly ensemble methods such as Random Forest, XGBoost, and neural networks—achieving crash detection accuracies of 85–92%. Explainable AI (XAI) frameworks, especially SHAP, enhance model interpretability, while hybrid and ensemble models improve predictive stability. Real-time monitoring via IoT sensors, connected vehicles, and computer vision enables surrogate safety evaluations using conflict-based metrics. Despite these advances, challenges remain regarding data heterogeneity, model transferability, privacy, and computational demands. Future directions include integrating autonomous vehicles, implementing standardized data-sharing platforms, and deploying automated safety countermeasures to transition from prediction to proactive prevention.

Keyword: Road Safety; Crash Prediction; Machine Learning; IoT Sensors; Explainable AI.

I. PENDAHULUAN

Kecelakaan lalu lintas masih menjadi permasalahan global yang serius, menyebabkan hilangnya nyawa, cedera, serta menimbulkan tekanan besar pada sistem kesehatan masyarakat dan perekonomian. Pada masa lalu, sebagian besar strategi keselamatan bersifat reaktif, yaitu berfokus pada analisis data kecelakaan untuk mengidentifikasi pola dan memahami risiko setelah kejadian terjadi. Pendekatan yang bersifat retrospektif ini sering kali kurang efektif dalam menangani potensi bahaya secara langsung dan tepat waktu. Seiring perkembangan kecerdasan buatan, teknologi sensor, dan analitik big data, paradigma keselamatan jalan mulai berubah. Fokus penelitian dan implementasi kini bergerak menuju pendekatan preventif dibandingkan

dengan pendekatan reaktif. Alih-alih menunggu terjadinya kecelakaan, sistem modern dirancang untuk mendeteksi kondisi berbahaya sejak dini. Dengan kemampuan untuk mengidentifikasi risiko lebih awal, pihak berwenang dapat mengambil langkah intervensi untuk mencegah kecelakaan jauh sebelum peristiwa tersebut terjadi.

Dalam beberapa tahun terakhir, penggunaan machine learning, kendaraan terkoneksi, dan big data untuk meningkatkan keselamatan jalan mengalami peningkatan yang sangat pesat. Saat ini, penelitian tidak lagi hanya berfokus pada pemetaan lokasi kecelakaan di masa lalu, tetapi juga menganalisis sejumlah besar data yang mencakup pola lalu lintas secara real-time, perilaku pengemudi, kondisi cuaca, hingga karakteristik geometrik jalan. Pergeseran ini

menunjukkan perubahan yang signifikan. Pendekatan keselamatan lalu lintas tidak lagi sekadar bersifat reaktif terhadap kecelakaan yang telah terjadi, melainkan mulai mengarah pada prediksi dan pemantauan risiko secara menyeluruh di dalam jaringan transportasi. Dengan menitikberatkan pada surrogate safety measures, yaitu identifikasi kejadian nyaris kecelakaan sebelum berkembang menjadi insiden nyata, area yang berpotensi berbahaya dapat dikenali lebih awal. Tinjauan ini merangkum berbagai perkembangan penting selama periode 2021–2025 serta menyoroti bagaimana pemanfaatan data terkini dan algoritma yang semakin canggih telah secara signifikan mengubah lanskap keselamatan lalu lintas (Skaug, T. R., et al., 2025). Bagian selanjutnya akan mendalami metodologi-metodologi kunci yang mendukung pergeseran paradigma tersebut.

II. KEMAJUAN METODOLOGIS

2.1. Pendekatan Machine Learning dan Deep Learning

Periode 2021 hingga 2025 menunjukkan adopsi dan pengembangan yang luas terhadap teknik machine learning (ML) dan deep learning (DL) dalam analisis keselamatan jalan. Metode-metode ini terbukti lebih unggul dibandingkan pendekatan statistik tradisional, khususnya dalam menangani data yang kompleks dan berdimensi tinggi serta dalam mengidentifikasi hubungan nonlinier antar faktor yang berkontribusi terhadap terjadinya kecelakaan.

Para peneliti telah memanfaatkan berbagai algoritma supervised learning untuk memprediksi kecelakaan dan menilai tingkat keparahan kecelakaan. Random Forest (RF), Gradient Boosting Machines (GBM), dan XGBoost menjadi metode yang paling banyak digunakan karena memiliki kinerja yang kuat serta tingkat akurasi prediksi yang tinggi. Sangare dkk. Mengembangkan kerangka kerja machine learning hibrida yang menggabungkan beberapa algoritma, yaitu Random Forest, Support Vector Machine (SVM), dan Neural Network, untuk memprediksi kecelakaan lalu lintas. Pendekatan tersebut menghasilkan peningkatan signifikan dalam akurasi prediksi (Sangare, M., et al., 2021). Penelitian mereka menunjukkan bahwa metode ensemble secara konsisten memberikan performa yang lebih baik dibandingkan dengan model tunggal, dengan Random Forest menunjukkan hasil terbaik di antara metode yang diuji.

Yang dkk., menggunakan XGBoost untuk menganalisis kecelakaan yang melibatkan truk angkutan barang dan menunjukkan bahwa algoritma tersebut mampu menangani dataset yang tidak seimbang serta menangkap interaksi kompleks antar faktor risiko (Yang, C., et al., 2021). Penelitian mereka menemukan bahwa, dengan proses tuning yang tepat, XGBoost mampu meningkatkan akurasi prediksi hingga lebih dari 85%, jauh melampaui model regresi logistik tradisional. Sejalan dengan penelitian tersebut, Al-Mistarehi dkk., menerapkan beberapa model machine learning, seperti Decision Tree, Random Forest, dan Neural Network, menggunakan perangkat lunak RStudio dan ArcGIS untuk memprediksi tingkat keparahan kecelakaan. Hasil penelitian menunjukkan bahwa metode ensemble memberikan prediksi yang paling andal di berbagai wilayah geografis (Al-Mistarehi, B.-W. et al., 2022).

Deep learning telah menghadirkan pendekatan baru dalam menangani data sekuensial, spasial, dan temporal untuk analisis keselamatan jalan. Zhang dkk., mengembangkan attention-based residual dilated network untuk memprediksi kecelakaan lalu lintas dengan memanfaatkan kemampuan deep learning dalam menangkap pola spatiotemporal jangka panjang (Zhang, X., et al., 2023). Model tersebut menggunakan mekanisme attention untuk memfokuskan analisis pada fitur-fitur yang paling penting, sementara dilated convolution memperluas receptive field tanpa meningkatkan kompleksitas komputasi secara signifikan. Hasil penelitian menunjukkan bahwa model yang dikembangkan mampu mengungguli berbagai model lainnya dalam memprediksi kecelakaan pada berbagai rentang waktu, sehingga menegaskan keunggulan penggunaan arsitektur neural network khusus dalam kajian keselamatan lalu lintas.

Convolutional Neural Networks (CNNs) dan Recurrent Neural Networks (RNNs), khususnya Long Short-Term Memory (LSTM), telah banyak digunakan untuk menganalisis data video dan pola lalu lintas yang bersifat sekuensial. Model-model ini mampu mengekstraksi fitur-fitur kompleks dari sumber data mentah seperti rekaman kamera lalu lintas dan lintasan GPS, sehingga memungkinkan pemahaman yang lebih mendalam terhadap faktor-faktor pemicu kecelakaan dan risiko keselamatan lalu lintas (Al-Mistarehi, B.-W. et al., 2021).

Metode machine learning terus menunjukkan kinerja yang lebih unggul dibandingkan pendekatan tradisional,

sebagaimana ditunjukkan dalam berbagai penelitian. Mussah dkk., mengembangkan kerangka kerja berbasis XGBoost untuk deteksi kecelakaan secara real-time dan berhasil mencapai tingkat akurasi sebesar 87%–92% menggunakan data kendaraan terkoneksi. Model ensemble yang mereka kembangkan mampu meningkatkan akurasi hingga 92% melalui kombinasi cerdas beberapa klasifikator (Mussah, A. R., et al., 2022). Sementara itu, Swathi dkk., menemukan bahwa model machine learning mampu mencapai tingkat akurasi antara 75% hingga 90% dalam memprediksi kecelakaan lalu lintas, dengan tingkat akurasi yang sangat dipengaruhi oleh kualitas data dan proses feature engineering (Swathi, M., et al., 2022). Hasil-hasil tersebut menunjukkan besarnya pengaruh metode ML/DL dalam mentransformasi analisis keselamatan jalan.

2.2. Metode Hybrid dan Ensemble

Tidak ada satu metodologi pun yang mampu memberikan kinerja terbaik pada semua kondisi, sehingga para peneliti mulai mengembangkan pendekatan hybrid dan ensemble yang menggabungkan keunggulan berbagai teknik. Metode-metode ini memadukan model statistik tradisional dengan algoritma machine learning, atau mengombinasikan beberapa model ML sekaligus untuk menghasilkan prediksi yang lebih andal dan akurat.

Beberapa penelitian telah mengombinasikan metode statistik tradisional dengan pendekatan machine learning. Mussah dkk., menerapkan kerangka kerja multi-tahap dengan memulai analisis menggunakan model statistik seperti Ordinary Least Squares (OLS) dan regresi Poisson untuk mengevaluasi data kecelakaan historis, kemudian melanjutkannya dengan klasifikasi berbasis XGBoost untuk mendeteksi kecelakaan secara real-time (Mussah, A. R., et al., 2022). Pendekatan hybrid ini memanfaatkan interpretabilitas model statistik untuk memahami pola masa lalu sekaligus mengandalkan kemampuan prediktif machine learning untuk aplikasi waktu nyata.

Lo H.W., mengembangkan model hybrid yang menggabungkan Failure Mode and Effects Analysis (FMEA) dengan Human Factors Analysis and Classification System (HFACS) untuk menilai risiko kecelakaan bus antarkota (Lo, H.-W., 2022). Dengan mengintegrasikan kedua kerangka tersebut, model yang dikembangkan mampu menghubungkan penilaian risiko secara sistematis dengan analisis faktor manusia, sehingga menghasilkan pendekatan yang

komprehensif mencakup aspek teknis maupun aspek manusia dalam keselamatan lalu lintas. Model hybrid ini memungkinkan peneliti melakukan penilaian risiko secara lebih luas melalui pemanfaatan berbagai perspektif dan tradisi analitis yang berbeda.

Metode ensemble yang menggabungkan prediksi dari beberapa model telah menunjukkan hasil yang sangat baik. Sangare dkk., menemukan bahwa kombinasi Random Forest, SVM, dan Neural Network menghasilkan model ensemble yang secara konsisten mampu melampaui kinerja masing-masing model tunggal pada berbagai metrik evaluasi (Sangare, M., et al., 2021). Pendekatan ini mampu mengurangi varians dan bias sehingga menghasilkan prediksi yang lebih stabil dan akurat. Temuan serupa juga dilaporkan oleh Mussah dkk., di mana model ensemble yang terdiri atas beberapa classifier XGBoost yang dilatih menggunakan subset data yang berbeda berhasil mencapai tingkat akurasi sebesar 92%, melampaui performa setiap model individual (Mussah, A. R., et al., 2022).

Huang dkk., memperkenalkan kerangka kerja Enhanced Safe Route Mapping (ESRM) yang mengombinasikan prediksi berbasis Light Gradient Boosting Machine (GBM) dengan model logika fuzzy (Huang, Y., et al., 2024). Model GBM digunakan untuk memprediksi skor risiko secara langsung dengan memanfaatkan data pengemudi dan kondisi lalu lintas secara real-time. Sementara itu, model logika fuzzy mengintegrasikan hasil prediksi tersebut dengan data kecelakaan historis untuk menentukan tingkat risiko pada berbagai segmen jalan.

Pendekatan hybrid ini menggabungkan prediksi berbasis data dengan penalaran berbasis aturan (rule-based reasoning), sehingga penilaian risiko menjadi lebih komprehensif dan mampu menangani ketidakpastian serta ketidaktepatan yang sering muncul dalam dunia nyata.

2.3. Analitik Real-Time dan Prediktif

Salah satu tren utama dalam penelitian keselamatan jalan terbaru adalah pergeseran dari analisis retrospektif ke pemantauan real-time dan analitik prediktif. Transisi ini memungkinkan dilakukannya intervensi yang bersifat proaktif (prediktif) serta penilaian risiko dinamis (real-time) yang dapat menyesuaikan diri terhadap perubahan kondisi secara langsung.

Mussah dkk., mengembangkan kerangka kerja komprehensif untuk penilaian keselamatan lalu lintas secara real-time menggunakan data kendaraan terkoneksi (Mussah, A. R., et al., 2022). Sistem yang mereka kembangkan mengekstraksi

metrik volatilitas berkendara, seperti akselerasi mendadak dan perlambatan ekstrem, dari data lintasan kendaraan beresolusi tinggi, kemudian memanfaatkan klasifikasi berbasis XGBoost untuk mendeteksi kecelakaan secara langsung saat kejadian berlangsung. Kerangka kerja tersebut berhasil mencapai tingkat akurasi sebesar 87–92% dalam mengidentifikasi peristiwa kecelakaan hanya dalam beberapa menit setelah kejadian, sehingga secara signifikan memungkinkan respons darurat yang lebih cepat dibandingkan dengan metode pelaporan kecelakaan tradisional.

Kerangka kerja ESRM memantau tindakan pengemudi dan kondisi jalan secara bersamaan dengan memanfaatkan algoritma pengenalan wajah untuk mendeteksi kelelahan pengemudi serta analitik video untuk memantau kondisi lalu lintas (Huang, Y., et al., 2024). Sistem ini memperbarui skor risiko secara real-time seiring perubahan situasi, sehingga pengemudi dan pengelola lalu lintas dapat memperoleh informasi keselamatan yang selalu terbaru. Pendekatan ini menunjukkan pergeseran besar dari penilaian keselamatan yang bersifat statis dan berbasis data historis menuju pemantauan keselamatan yang dinamis serta sadar konteks (context-aware).

Penelitian terbaru tidak hanya berfokus pada identifikasi kecelakaan yang sedang terjadi, tetapi juga pada prediksi kecelakaan di masa depan beserta tingkat keparahannya. Zhang dkk., melalui attention-based residual dilated network yang mereka kembangkan, mampu memprediksi kecelakaan lalu lintas pada berbagai horizon waktu, mulai dari 1 jam ke depan hingga 1 hari ke depan (Zhang, X., et al., 2023). Mekanisme attention pada model tersebut membantu mengidentifikasi fitur-fitur spatiotemporal yang paling berpengaruh dalam prediksi pada berbagai skala waktu. Hal ini tidak hanya meningkatkan akurasi model, tetapi juga membuat hasil prediksi lebih mudah dipahami dan diinterpretasikan.

Wang dkk., mengembangkan sistem penilaian risiko lalu lintas berbasis data peringatan yang mampu memprediksi potensi kecelakaan sebelum kejadian terjadi (Wang, J., et al., 2022). Dengan menganalisis pola pada peringatan pengemudi, notifikasi kendaraan, serta sensor lingkungan, sistem tersebut dapat mengidentifikasi situasi berisiko tinggi dan memberikan peringatan dini kepada pengemudi maupun pusat manajemen lalu lintas. Kemampuan prediktif ini memungkinkan dilakukannya intervensi preventif, seperti penyesuaian batas kecepatan, penyebaran pesan

peringatan, serta pengelolaan arus lalu lintas secara lebih efektif.

Model prediktif tingkat lanjut saat ini telah mengintegrasikan dimensi waktu dan ruang untuk memprediksi lokasi serta waktu terjadinya kecelakaan dengan tingkat kemungkinan yang paling tinggi. Fu dkk. (2023) mengembangkan sistem yang memanfaatkan data perilaku berkendara abnormal untuk mengidentifikasi lokasi dan waktu yang berisiko tinggi bagi keselamatan jalan (Fu, C., et al., 2023). Sistem tersebut menganalisis pola perilaku pengemudi, seperti pengereman mendadak, pelanggaran batas kecepatan, dan perpindahan jalur secara tiba-tiba, untuk mendeteksi hotspot risiko baru. Pendekatan prediksi spatiotemporal ini memungkinkan penempatan sumber daya penegakan hukum dan perbaikan infrastruktur secara lebih tepat sasaran, yaitu pada lokasi dan waktu yang paling membutuhkan intervensi keselamatan.

III. SUMBER DATA DAN TEKNOLOGI YANG BERKEMBANG

3.1. Data Kendaraan Terkoneksi dan Integrasi Internet of Things (IoT)

Perkembangan kendaraan terkoneksi (connected vehicles) dan teknologi Internet of Things (IoT) telah mengubah cara memperoleh data untuk analisis keselamatan jalan dengan menyediakan informasi yang sangat rinci serta kemampuan pemantauan secara real-time. Kendaraan yang terkoneksi menghasilkan berbagai jenis data, termasuk lintasan GPS, profil kecepatan, pola akselerasi, serta data diagnostik kendaraan. Mussah dkk., memanfaatkan data lintasan kendaraan terdisagregasi dari kendaraan terkoneksi melalui marketplace data milik Autonomous untuk memperoleh metrik volatilitas berkendara dan indikator perilaku pengemudi (Mussah, A. R., et al., 2022).

Data lintasan beresolusi tinggi yang direkam pada interval subdetik memungkinkan analisis mendalam terhadap pola berkendara serta identifikasi perilaku berisiko yang dapat memicu kecelakaan. Penelitian tersebut menunjukkan bahwa data dari kendaraan terkoneksi mampu memberikan peringatan dini terhadap risiko kecelakaan melalui analisis faktor-faktor seperti frekuensi pengereman mendadak, percepatan yang terlalu cepat, dan pola kemudi yang tidak stabil.

Huang dkk., mengembangkan aplikasi Android khusus yang memungkinkan pengemudi mengirimkan data jalan dan kendaraan secara

langsung melalui smartphone mereka (Huang, Y., et al., 2024). Pendekatan ini membuat proses pengumpulan data lebih mudah diakses dengan memanfaatkan perangkat seluler yang telah digunakan secara luas, sehingga memungkinkan penerapannya dalam skala besar tanpa memerlukan peralatan kendaraan khusus.

Aplikasi tersebut mengumpulkan data mengenai kecepatan kendaraan, lokasi, akselerasi, serta interaksi pengemudi dengan kendaraan. Informasi yang diperoleh kemudian dikirim ke server pusat untuk dianalisis lebih lanjut. Pendekatan berbasis perangkat seluler ini menjadi solusi yang hemat biaya untuk memperoleh data kendaraan terkoneksi, terutama di wilayah yang belum memiliki sistem telematika kendaraan khusus secara luas.

Selain sensor yang terpasang pada kendaraan, perangkat IoT di sisi jalan juga menyediakan aliran data pelengkap untuk analisis keselamatan lalu lintas. Huang dkk. memasang kamera di tepi jalan yang merekam data lalu lintas secara langsung, kemudian menganalisisnya menggunakan teknik analitik video khusus untuk memantau kecepatan kendaraan, lintasan pergerakan, serta interaksi antarkendaraan (Huang, Y., et al., 2024). Integrasi antara sensor berbasis kendaraan dan sensor berbasis infrastruktur menghasilkan sistem pemantauan yang komprehensif karena mampu merekam perilaku kendaraan secara individual maupun pola lalu lintas secara keseluruhan.

Yang dkk., memanfaatkan data komprehensif yang berasal dari laporan kecelakaan lalu lintas dan catatan pelanggaran yang dikumpulkan melalui berbagai sensor IoT serta sistem pelaporan (Yang, C.-H., et al., 2024). Penelitian mereka menunjukkan bahwa kombinasi data dari kamera penegakan hukum otomatis, sistem deteksi insiden, dan laporan manual mampu memberikan gambaran keselamatan lalu lintas yang lebih menyeluruh dibandingkan penggunaan satu sumber data saja.

Perkembangan infrastruktur IoT dan peningkatan protokol berbagi data semakin mempermudah integrasi berbagai sumber data tersebut, sehingga mendukung analisis keselamatan lalu lintas yang lebih akurat dan komprehensif.

3.2. Computer Vision dan Analitik Video

Teknologi computer vision telah berkembang menjadi alat yang sangat efektif untuk mengekstraksi informasi terkait keselamatan dari data visual. Teknologi ini memungkinkan analisis otomatis terhadap

kondisi lalu lintas dan perilaku pengemudi. Huang dkk., menggunakan metode analitik video khusus untuk menganalisis rekaman dari kamera tepi jalan guna melacak kecepatan kendaraan, lintasan perjalanan, dan interaksi antarkendaraan secara otomatis (Huang, Y., et al., 2024). Sistem tersebut mampu mendeteksi potensi permasalahan seperti kejadian nyaris kecelakaan, perpindahan jalur yang tidak aman, serta jarak antarkendaraan yang terlalu dekat. Analisis otomatis ini mengatasi keterbatasan proses peninjauan video secara manual, sehingga memungkinkan pemberian peringatan secara real-time serta pemantauan lalu lintas secara berkelanjutan.

Singh dkk., melakukan kajian komprehensif mengenai teknik surrogate safety analysis yang memanfaatkan teknologi mutakhir, seperti computer vision untuk mendeteksi konflik lalu lintas (Singh, D., et al., 2024). Penelitian tersebut menunjukkan bahwa analitik video dapat digunakan untuk mengidentifikasi indikator keselamatan proksi, seperti Time to Collision (TTC), Post-Encroachment Time (PET), dan Deceleration Rate to Avoid Collision (DRAC), tanpa memerlukan data kecelakaan aktual. Metrik surrogate tersebut memungkinkan penilaian keselamatan secara proaktif serta menyediakan data yang lebih kaya untuk analisis statistik melalui pengamatan kejadian nyaris kecelakaan, yang frekuensinya jauh lebih tinggi dibandingkan dengan kecelakaan nyata.

Pemantauan kondisi pengemudi secara real-time dimungkinkan melalui teknologi pengenalan wajah dan computer vision. Dengan menganalisis karakteristik wajah, pergerakan mata, dan posisi kepala, Huang dkk., mengembangkan algoritma pengenalan wajah untuk mengidentifikasi indikator kelelahan dan distraksi pengemudi (Huang, Y., et al., 2024). Sistem tersebut mampu memberikan peringatan kepada pengemudi maupun pihak pengelola armada atau otoritas lalu lintas apabila terdeteksi kondisi seperti kantuk, kurangnya perhatian, atau keadaan lain yang dapat meningkatkan risiko kecelakaan. Kemampuan untuk menangani faktor manusia, yang hingga saat ini masih menjadi penyebab utama kecelakaan lalu lintas, merupakan salah satu kemajuan penting dalam pengembangan sistem keselamatan jalan modern.

Analisis skala besar yang sebelumnya sulit dilakukan kini menjadi memungkinkan melalui penerapan computer vision, karena teknologi ini menghilangkan kebutuhan akan pengkodean manual terhadap peristiwa lalu lintas. Algoritma

identifikasi dan pelacakan objek berbasis deep learning mampu secara otomatis mengenali kendaraan, pejalan kaki, sepeda, serta berbagai pengguna jalan lainnya, kemudian menganalisis interaksi dan pola pergerakan mereka. Otomatisasi tersebut meningkatkan konsistensi dan objektivitas analisis, sekaligus secara signifikan mengurangi waktu dan biaya dalam pelaksanaan penilaian keselamatan lalu lintas (Singh, D., et al., 2024).

3.3. Platform Big Data dan Integrasi Multi-Sumber Data

Analisis keselamatan jalan kini dapat dilakukan secara lebih menyeluruh dan holistik melalui integrasi berbagai sumber data menggunakan platform big data. Kerangka kerja ESRM yang dikembangkan oleh Huang dkk., merupakan contoh nyata dari kekuatan data fusion, dengan menggabungkan analitik video dari kamera tepi jalan, data perilaku pengemudi secara real-time dari aplikasi seluler, hasil pengenalan wajah, serta data historis kecelakaan (Huang, Y., et al., 2024).

Melalui pemanfaatan keunggulan komplementer dari setiap sumber data, proses integrasi tersebut menyatukan berbagai aliran data yang berbeda ke dalam satu kerangka penilaian risiko yang terpadu. Data historis digunakan untuk menyediakan profil risiko dasar, sementara data real-time memungkinkan penyesuaian dinamis berdasarkan kondisi yang sedang berlangsung.

Aplikasi big data dalam penelitian keselamatan transportasi semakin memanfaatkan berbagai sumber data berskala besar, termasuk data kecelakaan, arus lalu lintas, informasi cuaca, geometri jalan, karakteristik kendaraan, serta demografi pengemudi, sebagaimana dijelaskan dalam tinjauan sistematis (He, M., et al., 2025). Tinjauan tersebut menyoroti bahwa integrasi data multi-sumber memungkinkan identifikasi faktor-faktor penyebab kecelakaan secara lebih akurat serta pengembangan model prediksi yang lebih andal. Selain itu, penelitian tersebut juga menegaskan bahwa masih terdapat berbagai tantangan yang memerlukan kajian lebih lanjut, terutama terkait proses pengumpulan data, kompleksitas pengolahan data, serta integrasi metodologi dalam analisis keselamatan transportasi.

Pengembangan repositori data berskala besar semakin mempermudah pelaksanaan penelitian dan studi komparatif di bidang keselamatan lalu lintas. Meskipun publikasi yang dikaji hanya memberikan sedikit informasi

spesifik mengenai repositori data, tren menuju penggunaan platform bersama dan format data yang terstandar terlihat semakin jelas. (Yang, C.-H., et al., 2024) Menunjukkan manfaat integrasi laporan kecelakaan lalu lintas dan catatan pelanggaran ke dalam satu platform big data terpadu untuk keperluan analisis. Keberadaan platform semacam ini memungkinkan dilakukannya penelitian longitudinal, perbandingan lintas wilayah yurisdiksi, serta pengembangan model yang dapat ditransfer dan diterapkan dalam berbagai kondisi dan lingkungan.

Sistem cloud computing semakin banyak digunakan karena tingginya kebutuhan komputasi untuk menganalisis data lalu lintas dalam jumlah besar. Solusi yang dikembangkan oleh Huang dkk., mengirimkan data dari aplikasi seluler dan kamera tepi jalan ke server pusat untuk diproses, dengan memanfaatkan infrastruktur cloud guna menangani beban komputasi dari proses pengenalan wajah, analitik video, dan inferensi machine learning (Huang, Y., et al., 2024). Arsitektur berbasis cloud memungkinkan penerapan sistem yang scalable serta mendukung pemrosesan data secara real-time, kemampuan yang sulit dicapai dengan infrastruktur konvensional berbasis on-premises. Namun demikian, integrasi multi-sumber data masih menghadapi tantangan teknis dan regulasi, antara lain perbedaan standar interoperabilitas antarsistem, kerangka hukum privasi yang belum seragam, serta kebutuhan akan protokol berbagi data yang aman dan terstandar.

IV. TEMUAN UTAMA DAN APLIKASI PENELITIAN

4.1 Prediksi Kecelakaan dan Penilaian Tingkat Keparahannya

Penelitian terbaru telah menghasilkan peningkatan signifikan dalam akurasi dan ketepatan waktu prediksi kecelakaan serta penilaian tingkat keparahan kecelakaan, yang memberikan implikasi penting bagi upaya respons darurat dan pencegahan kecelakaan. Berbagai penelitian menunjukkan bahwa teknik machine learning modern mampu memprediksi frekuensi dan tingkat keparahan kecelakaan dengan tingkat akurasi yang tinggi. Mussah dkk., menggunakan klasifikasi berbasis XGBoost dan data kendaraan terkoneksi untuk mencapai tingkat akurasi sebesar 87–92% dalam deteksi kecelakaan secara real-time (Mussah, A. R., et al., 2022). Yang dkk., melaporkan bahwa model XGBoost mampu melampaui regresi logistik

konvensional dengan tingkat akurasi prediksi lebih dari 85% pada kasus kecelakaan truk angkutan barang (Mussah, A. R., et al., 2021).

Sementara itu, Sangare dkk., menemukan bahwa teknik ensemble machine learning mampu mencapai akurasi antara 75% hingga 90%, bergantung pada kualitas data dan proses feature engineering yang digunakan (Sangare, M., et al., 2023). Tingkat performa tersebut mendekati ambang utilitas praktis untuk implementasi operasional dan menunjukkan peningkatan yang signifikan dibandingkan dengan metode konvensional.

Penelitian terbaru tidak lagi hanya berfokus pada prediksi terjadinya kecelakaan, tetapi juga pada prediksi tingkat keparahan kecelakaan, seperti kerusakan material, cedera, dan kecelakaan yang berujung pada kematian. Al-Mistarehi dkk., mengembangkan model machine learning yang mampu memprediksi tingkat keparahan kecelakaan secara akurat, sehingga memungkinkan alokasi sumber daya respons darurat dilakukan secara lebih tepat (Al-Mistarehi, B.-W. et al., 2022). Dalam analisis komparatif terhadap model machine learning untuk penilaian tingkat keparahan kecelakaan kendaraan transportasi, Houssam dkk., menemukan bahwa pendekatan ensemble menghasilkan estimasi tingkat keparahan yang paling akurat (Houssam, B., et al., 2023). Kemampuan memprediksi tingkat keparahan kecelakaan memungkinkan layanan darurat memprioritaskan penanganan insiden yang paling kritis serta menyiapkan sumber daya yang relevan sebelum kejadian berkembang lebih parah.

Model machine learning telah menghasilkan berbagai wawasan penting mengenai faktor-faktor penyebab kecelakaan, terutama ketika dipadukan dengan teknik AI yang dapat dijelaskan (XAI). Yang dkk., meneliti faktor-faktor yang berkontribusi terhadap kecelakaan truk angkutan barang menggunakan XGBoost yang dikombinasikan dengan SHAP (Shapley Additive exPlanations) (Mussah, A. R., et al., 2021). Hasil penelitian mereka menunjukkan bahwa faktor utama penyebab kecelakaan meliputi geometri jalan, masalah pemeliharaan kendaraan, dan kelelahan pengemudi.

Kerangka kerja SHAP memungkinkan penentuan prioritas intervensi berbasis bukti melalui penyediaan ukuran kuantitatif mengenai kontribusi masing-masing faktor terhadap risiko kecelakaan. Kombinasi antara kemampuan prediksi dan penjelasan ini merupakan kemajuan yang signifikan dibandingkan dengan model black-box, yang umumnya hanya memberikan

prediksi yang akurat tanpa mampu menjelaskan mekanisme yang mendasarinya.

4.2. Ukuran Keselamatan Surrogate dan Analisis Konflik Lalu Lintas

Sebagai alternatif proaktif terhadap penelitian berbasis data kecelakaan, surrogate safety measures semakin banyak digunakan karena memungkinkan evaluasi keselamatan dilakukan berdasarkan kejadian nyaris kecelakaan dan konflik lalu lintas, yang frekuensinya jauh lebih tinggi dibandingkan dengan kecelakaan aktual.

Time-to-Collision (TTC), Post-Encroachment Time (PET), dan Deceleration Rate to Avoid Crash (DRAC) merupakan contoh surrogate safety measures yang semakin banyak digunakan, sebagaimana dilaporkan dalam kajian komprehensif (Singh, D., et al., 2024). Metrik-metrik ini digunakan untuk mengukur tingkat keparahan konflik lalu lintas, yaitu kondisi ketika tabrakan berpotensi terjadi apabila tidak ada tindakan pencegahan. Dengan menganalisis konflik lalu lintas dibandingkan dengan kecelakaan aktual, peneliti dapat memperoleh ukuran sampel yang jauh lebih besar, sehingga memungkinkan analisis statistik yang lebih andal serta deteksi dini terhadap isu keselamatan baru yang muncul. Kajian tersebut juga menekankan bahwa deteksi konflik otomatis kini dapat dilakukan dalam skala besar berkat kemajuan teknologi seperti computer vision dan data kendaraan yang terhubung.

Zheng dkk., melakukan kajian mendalam terhadap berbagai teknik analitis untuk memodelkan konflik lalu lintas dalam analisis keselamatan jalan (Mussah, A. R., et al., 2022). Berdasarkan tinjauan tersebut, metode berbasis konflik mendukung intervensi preventif dengan memungkinkan identifikasi dini permasalahan keselamatan sebelum terjadinya kecelakaan.

Para penulis membahas berbagai algoritma deteksi konflik, teknik penilaian tingkat keparahan, serta model statistik yang digunakan untuk menghubungkan konflik lalu lintas dengan risiko kecelakaan. Mereka menegaskan bahwa telah terjadi pergeseran paradigma yang signifikan dalam penelitian keselamatan jalan, yaitu transisi dari analisis kecelakaan yang bersifat reaktif menuju analisis konflik yang bersifat proaktif.

Penggunaan surrogate safety measures sebagai indikator keselamatan telah divalidasi oleh berbagai studi yang menunjukkan adanya korelasi yang kuat antara ukuran tersebut dan tingkat kecelakaan aktual. (Singh, D., et al., 2024),

melaporkan bahwa lokasi dengan frekuensi konflik tinggi serta karakteristik konflik yang parah (misalnya nilai Time-to-Collision/TTC yang rendah) umumnya menunjukkan tingkat kecelakaan yang lebih tinggi pada periode berikutnya.

Validasi ini mendukung penggunaan analisis konflik sebagai pendekatan untuk mengidentifikasi lokasi berisiko tinggi serta mengevaluasi efektivitas intervensi keselamatan, tanpa harus menunggu periode pengamatan yang panjang untuk mengumpulkan data kecelakaan dalam jumlah yang memadai.

4.3. Penilaian Keselamatan Tingkat Jaringan

Penelitian terbaru semakin mengadopsi pendekatan tingkat jaringan (network-level) yang mempertimbangkan kinerja keselamatan secara menyeluruh dalam suatu sistem transportasi. Sebaliknya, analisis keselamatan jalan tradisional umumnya berfokus pada lokasi-lokasi spesifik, seperti persimpangan atau segmen jalan tertentu.

Melampaui analisis berbasis lokasi spesifik, sebuah studi tahun 2023 mengembangkan metrik keselamatan tingkat jaringan untuk penilaian keselamatan lalu lintas secara menyeluruh (X. Chen et al., 2023). Penelitian tersebut memperkenalkan berbagai indikator yang mampu menangkap kinerja keselamatan sistem secara keseluruhan, seperti tingkat kecelakaan pada skala jaringan, pengelompokan spasial lokasi berisiko tinggi, serta ukuran penyebaran risiko berbasis konektivitas.

Pendekatan ini memungkinkan instansi transportasi untuk mengevaluasi keselamatan jaringan secara komprehensif, mengidentifikasi permasalahan yang bersifat sistemik, serta menetapkan prioritas peningkatan berdasarkan dampak pada tingkat jaringan, bukan hanya tingkat keparahan di lokasi tertentu.

Sistem Informasi Geografis (Geographic Information Systems/GIS) kini menjadi alat penting dalam penilaian keselamatan pada tingkat jaringan. Alkhadour dkk., mengidentifikasi titik-titik rawan kecelakaan (accident hotspots) di Amman, Yordania, menggunakan alat analisis spasial berbasis statistik dan GIS (Alkhadour, W., et al., 2021). Penelitian tersebut memanfaatkan teknik *optimized hotspot analysis* untuk mengidentifikasi zona berisiko tinggi dengan tingkat kepercayaan 90% atau lebih, serta *Nearest Neighbor Index (NNI)* untuk mengevaluasi pola pengelompokan spasial.

Hasil studi menunjukkan bahwa beberapa tipe jalan (seperti jalan dua lajur dengan median)

serta zona tertentu (termasuk kawasan komersial, permukiman, dan industri) memiliki tingkat risiko yang lebih tinggi, sementara kecelakaan fatal menunjukkan pola pengelompokan yang signifikan. Pendekatan spasial ini memungkinkan intervensi yang lebih terarah pada area yang paling membutuhkan peningkatan keselamatan.

Analisis tingkat jaringan (network-level analysis) mengakui bahwa permasalahan keselamatan di satu bagian jaringan dapat memengaruhi kondisi di bagian lain. Misalnya, kecelakaan yang menyebabkan penutupan lajur akibat perlambatan mendadak dan kemacetan dapat menimbulkan risiko kecelakaan sekunder di hilir (downstream).

Hu dkk., mengembangkan sebuah kerangka kerja penilaian keselamatan lalu lintas yang mempertimbangkan efek jaringan tersebut, sehingga memungkinkan evaluasi risiko yang lebih komprehensif (Hu, L. et al., 2023). Pendekatan ini menegaskan bahwa peningkatan keselamatan tidak hanya bergantung pada pemahaman lokasi spesifik, tetapi juga pada bagaimana lokasi tersebut terhubung dan berinteraksi dalam sistem transportasi yang lebih luas. GIS berperan penting dalam analisis tingkat jaringan ini dengan menyediakan kemampuan visualisasi pola risiko spasial, identifikasi hotspot sistematis, dan dukungan pengambilan keputusan berbasis data geospasial yang terintegrasi.

4.4. Analisis Perilaku Pengemudi dan Faktor Risiko

Penelitian keselamatan jalan terus berfokus pada pemahaman perilaku pengemudi serta identifikasi faktor-faktor risiko spesifik, dan studi-studi terbaru telah memanfaatkan sumber data serta metode analisis yang lebih mutakhir untuk memperoleh wawasan yang lebih mendalam. (Fu, C., et al., 2023), Mengembangkan teknik untuk mengidentifikasi risiko keselamatan jalan dengan memanfaatkan data perilaku berkendara yang tidak normal (*anomalous driving behavior*). Berdasarkan analisis data mereka, sejumlah perilaku tertentu, seperti pengereman mendadak, akselerasi cepat, manuver menikung tajam, dan perpindahan lajur secara tiba-tiba, terkait dengan peningkatan risiko kecelakaan.

Sistem tersebut mampu menghasilkan skor risiko pengemudi serta mengidentifikasi lokasi berisiko tinggi di mana perilaku menyimpang terkonsentrasi, melalui pengukuran frekuensi dan intensitas kejadian tersebut. Pendekatan berbasis perilaku ini memungkinkan intervensi yang lebih terarah, seperti peningkatan infrastruktur,

penegakan aturan lalu lintas, serta edukasi bagi pengemudi untuk mengurangi perilaku berisiko.

Analisis faktor manusia secara khusus dimasukkan ke dalam model hybrid FMEA-HFACS yang dikembangkan oleh Lo untuk kecelakaan bus antarkota (Lo, H.-W., 2022). Kerangka HFACS (Human Factors Analysis and Classification System) secara sistematis mengklasifikasikan kesalahan manusia dan faktor organisasi yang berkontribusi terhadap kecelakaan ke dalam beberapa tingkatan, yaitu tindakan tidak aman (unsafe acts), kondisi prasyarat tindakan tidak aman (preconditions for unsafe acts), pengawasan yang tidak aman (unsafe supervision), serta pengaruh organisasi (organizational influences).

Pendekatan ini memberikan pemahaman yang lebih komprehensif mengenai penyebab kecelakaan dengan mengintegrasikan perspektif faktor manusia bersama penilaian risiko teknis melalui FMEA. Mengingat sebagian besar kecelakaan lalu lintas disebabkan oleh kesalahan manusia, pendekatan komprehensif ini menjadi sangat penting dalam analisis keselamatan jalan modern.

Metode pengenalan wajah yang dikembangkan oleh Huang dkk., untuk mendeteksi pengemudi yang lelah dan terdistraksi merupakan langkah penting dalam pemantauan pengemudi secara real-time (Huang, Y., et al., 2024). Sistem tersebut mampu mengidentifikasi kondisi seperti kelelahan, kurangnya perhatian, atau gangguan lainnya, kemudian memberikan peringatan, menyarankan waktu istirahat, atau bahkan pada kasus yang lebih lanjut dapat mengambil alih kendali kendaraan. Salah satu tantangan terbesar dalam keselamatan jalan adalah mengidentifikasi pengemudi yang berada di bawah pengaruh alkohol sebelum mereka menyebabkan kecelakaan.

Penggunaan metode explainable AI, khususnya SHAP, telah memungkinkan pengukuran kontribusi faktor risiko secara lebih akurat. Studi Yang dkk., mengenai kecelakaan truk angkutan barang menemukan bahwa kelelahan pengemudi menyumbang sekitar 30% terhadap risiko kecelakaan, masalah pemeliharaan kendaraan sekitar 20%, dan karakteristik geometri jalan tertentu sekitar 15% (Mussah, A. R., et al., 2021). Temuan kuantitatif ini membantu pengambilan keputusan, terutama dalam menentukan prioritas intervensi serta alokasi sumber daya untuk menangani faktor-faktor yang paling berdampak terhadap keselamatan jalan.

V. INOVASI TEKNOLOGI

5.1. Explainable AI (XAI) dan Interpretabilitas

Seiring dengan semakin kompleksnya dan kuatnya model machine learning, kebutuhan akan interpretabilitas dan explainability juga semakin meningkat, terutama dalam domain yang bersifat kritis seperti keselamatan lalu lintas. Yang dkk., merupakan salah satu studi awal yang menggunakan SHAP (Shapley Additive exPlanations) untuk menganalisis kecelakaan truk angkutan barang dan keselamatan jalan (Mussah, A. R., et al., 2021).

SHAP menyediakan kerangka kerja terpadu untuk menjelaskan prediksi model dengan mengkuantifikasi kontribusi setiap fitur terhadap hasil prediksi tertentu. Metode ini mampu mengatasi permasalahan "black box" yang melekat pada model kompleks seperti XGBoost dan jaringan saraf. Dengan demikian, pemangku kepentingan dapat memahami alasan di balik prediksi peningkatan risiko kecelakaan pada kondisi tertentu. Kejelasan yang dihasilkan oleh SHAP tidak hanya meningkatkan kepercayaan terhadap hasil model, tetapi juga membantu mengidentifikasi faktor risiko yang dapat ditindaklanjuti untuk intervensi keselamatan. Beberapa penelitian telah menggunakan teknik feature importance untuk mengidentifikasi variabel yang paling berpengaruh terhadap risiko kecelakaan. Model Random Forest dan Gradient Boosting secara inheren menyediakan skor feature importance yang menunjukkan seberapa besar kontribusi setiap variabel dalam meningkatkan akurasi prediksi. Sangare dkk., memanfaatkan skor ini untuk mengidentifikasi faktor-faktor paling kritis dalam peramalan kecelakaan lalu lintas, dan menemukan bahwa volume lalu lintas, kondisi cuaca, serta waktu kejadian merupakan prediktor yang paling berpengaruh (Sangare, M., et al., 2023). Informasi ini dapat digunakan untuk mengarahkan prioritas pengumpulan data serta merumuskan strategi intervensi keselamatan yang lebih tepat sasaran.

Model attention-based residual dilated network yang dikembangkan oleh Zhang dkk., menggunakan mekanisme attention untuk menunjukkan aspek spatiotemporal mana yang menjadi fokus model saat melakukan prediksi (Zhang, X., et al., 2023). Visualisasi bobot attention menunjukkan bahwa model cenderung memusatkan perhatian pada lokasi dan waktu dengan tingkat risiko yang tinggi. Hal ini memberikan gambaran yang jelas tentang cara model membuat keputusan. Kemampuan untuk

memahami proses pengambilan keputusan ini sangat penting dalam membangun kepercayaan terhadap model deep learning, serta dalam mengidentifikasi potensi bias atau keterbatasan yang mungkin terdapat dalam model tersebut.

Tren menuju explainable AI mencerminkan pengakuan bahwa transparansi model merupakan elemen penting dalam implementasi operasional pada aplikasi keselamatan lalu lintas. Instansi transportasi dan pembuat kebijakan perlu memahami alasan di balik prediksi model agar dapat memberikan justifikasi terhadap intervensi, mengalokasikan sumber daya secara tepat, serta menjaga kepercayaan publik. Studi-studi yang ditinjau menunjukkan bahwa aspek keterjelasan (explainability) tidak harus mengorbankan akurasi, karena teknik-teknik modern saat ini memungkinkan tercapainya kinerja tinggi sekaligus interpretabilitas yang baik (Mussah, A. R., et al., 2021).

5.2. Sistem Informasi Geografis (GIS) dan Analisis Spasial

Teknologi Sistem Informasi Geografis (GIS) sangat penting dalam kajian keselamatan jalan karena memungkinkan analisis spasial yang kompleks serta visualisasi data, sehingga mendukung pengambilan keputusan berbasis bukti. Alkhadour dkk., (Alkhadour, W., et al., 2021) menunjukkan bahwa analisis hotspot berbasis GIS merupakan pendekatan yang efektif untuk mengidentifikasi area berisiko tinggi. Dengan menggunakan tools hotspot analysis pada ArcGIS serta Nearest Neighbor Index (NNI), penelitian tersebut menemukan bahwa kecelakaan di Amman, Yordania, tersebar di wilayah yang luas. Informasi ini memungkinkan penerapan fitur keselamatan tambahan di lokasi tertentu serta pada ruas jalan dengan tingkat kecelakaan yang tinggi. Studi tersebut juga mengungkap bahwa hotspot sebagian besar terdapat pada jenis jalan tertentu (jalan dua lajur dengan median) serta pada zona permukiman, komersial, dan industri. Temuan ini memberikan dasar penting bagi perencanaan dan peningkatan infrastruktur keselamatan jalan yang lebih tepat sasaran.

GIS dapat digunakan untuk menganalisis pola dan hubungan spasial yang sulit dideteksi dengan metode lain. Alkhadour dkk., mencatat bahwa kecelakaan fatal cenderung menunjukkan pola pengelompokan (clustering) yang signifikan, yang mengindikasikan bahwa lokasi tertentu memiliki karakteristik yang meningkatkan baik tingkat keparahan maupun probabilitas terjadinya kecelakaan (Alkhadour,

W., et al., 2021). Temuan ini memiliki implikasi penting dalam penentuan prioritas peningkatan keselamatan, di mana wilayah dengan tingkat kecelakaan fatal yang tinggi memerlukan tindakan cepat serta penerapan langkah-langkah intervensi yang lebih komprehensif.

Integrasi data kecelakaan dengan berbagai dataset geospasial lainnya, seperti geometri jalan, volume lalu lintas, tata guna lahan, demografi, dan variabel lingkungan, kini dimungkinkan oleh sistem GIS modern. Al-Mistarehi dkk., memanfaatkan model machine learning yang diintegrasikan dengan ArcGIS untuk memprediksi tingkat keparahan kecelakaan menggunakan data geografis terkait parameter jalan dan kondisi lingkungan (Al-Mistarehi, B.-W. et al., 2022).

Integrasi ini memungkinkan analisis yang lebih komprehensif karena mempertimbangkan keseluruhan konteks lingkungan tempat terjadinya kecelakaan, sehingga menghasilkan pemahaman yang lebih akurat terhadap faktor-faktor yang memengaruhi tingkat keparahan kecelakaan. Kemampuan visualisasi yang kuat dalam GIS memudahkan penyampaian temuan keselamatan kepada berbagai pemangku kepentingan. Kerangka ESRM yang dikembangkan oleh Huang dkk., menghasilkan peta panas risiko (risk heatmaps) yang secara jelas menggambarkan tingkat risiko di seluruh jaringan jalan (Huang, Y., et al., 2024).

Representasi visual tersebut memungkinkan pemahaman yang intuitif terhadap pola bahaya spasial, sehingga dapat mendukung kampanye peningkatan kesadaran publik, perencanaan rute perjalanan, serta penempatan aparat penegakan hukum. Kemampuan untuk memvisualisasikan hasil analisis yang kompleks secara jelas merupakan faktor penting dalam menjembatani penelitian dengan implementasi kebijakan di lapangan.

5.3. Grafik Pengetahuan dan Analisis Semantik

Metode baru memanfaatkan analisis semantik dan representasi pengetahuan untuk menangkap keterkaitan kompleks antara faktor-faktor penyebab kecelakaan serta memungkinkan penalaran yang lebih canggih dalam keselamatan lalu lintas. Zhang dkk., mengembangkan sebuah knowledge graph untuk menganalisis kecelakaan lalu lintas dengan menggunakan jaringan semantik terstruktur yang merepresentasikan entitas (seperti kendaraan, pengemudi, karakteristik jalan, dan

kondisi cuaca) beserta hubungan di antaranya (Zhang, Y., et al., 2022).

Grafik pengetahuan ini memungkinkan analisis yang lebih rinci terhadap kausalitas kecelakaan, karena tidak hanya merekam penyebab individual, tetapi juga hubungan dan ketergantungan antarfaktor. Sebagai contoh, grafik tersebut menunjukkan bahwa kondisi jalan basah meningkatkan kemungkinan terjadinya kecelakaan, dan efek tersebut menjadi lebih signifikan ketika dikombinasikan dengan kecepatan tinggi serta jarak pandang yang buruk. Representasi relasional ini memungkinkan penalaran terhadap skenario multifaktor yang kompleks, yang sulit dimodelkan dengan metode konvensional.

Grafik pengetahuan (knowledge graphs) mengintegrasikan pengetahuan domain dan inferensi logis untuk memungkinkan penalaran semantik yang melampaui sekadar pengenalan pola. Algoritma yang dikembangkan oleh Zhang dkk., mampu menyimpulkan potensi bahaya dari kombinasi parameter tertentu, bahkan ketika kombinasi tersebut belum pernah muncul dalam data historis sebelumnya (Zhang, Y., et al., 2022). Kemampuan ini sangat bermanfaat dalam mengidentifikasi ancaman baru serta skenario langka namun berisiko tinggi yang sering kali kurang terwakili dalam data pelatihan. Untuk menggabungkan pengetahuan domain yang terstruktur dengan pembelajaran berbasis data, knowledge graph dapat diintegrasikan dengan model machine learning. Pendekatan hybrid ini mengombinasikan pengetahuan ahli mengenai hubungan kausal dan prinsip keselamatan dengan kemampuan machine learning dalam mengenali pola kompleks. Zhang dkk., menunjukkan bahwa integrasi knowledge graph dan machine learning menghasilkan peningkatan akurasi prediksi sekaligus interpretabilitas yang lebih baik dibandingkan masing-masing pendekatan secara terpisah (Zhang, Y., et al., 2022). Secara struktural, knowledge graphs menyusun entitas dan relasi menggunakan node dan edge sehingga mendukung inferensi logis dan penalaran berbasis pengetahuan; integrasi dengan model machine learning memungkinkan prediksi yang tidak hanya didasarkan pada pola statistik tetapi juga mempertimbangkan hubungan sebab-akibat yang telah diketahui.

VI. TANTANGAN DAN KETERBATASAN

Analisis keselamatan jalan masih menghadapi berbagai tantangan signifikan yang membatasi efektivitas dan penerapan metode

yang ada, meskipun telah terjadi kemajuan pesat. Banyak penelitian menunjukkan bahwa kualitas data masih menjadi masalah utama. Keandalan hasil analisis berbasis laporan kecelakaan dapat terpengaruh oleh data yang tidak lengkap, tidak akurat, atau terlambat tersedia.

Meskipun data kendaraan terkoneksi sangat kaya informasi, Mussah dkk., mencatat bahwa cakupannya masih terbatas karena hanya sebagian kecil kendaraan yang saat ini dilengkapi dengan teknologi konektivitas, sehingga berpotensi menimbulkan bias sampel (Mussah, A. R., et al., 2022). Sejalan dengan itu, He dkk., dalam kajian komprehensifnya, menyoroti bahwa kompleksitas pengumpulan dan pengolahan data masih menjadi tantangan yang berkelanjutan dalam penerapan big data untuk keselamatan transportasi (He, M., et al., 2025). Untuk mengatasi permasalahan tersebut, diperlukan investasi berkelanjutan dalam infrastruktur pengumpulan data, prosedur jaminan kualitas, serta metode untuk menangani data yang tidak lengkap atau tidak jelas.

Model yang dilatih menggunakan data dari yurisdiksi atau periode waktu tertentu sering kali tidak menunjukkan kinerja yang baik ketika diterapkan pada konteks yang berbeda. Variasi regional dalam pola lalu lintas, perilaku pengemudi, infrastruktur jalan, serta strategi penegakan hukum dapat membatasi tingkat transferabilitas model.

Sejumlah penelitian menunjukkan bahwa untuk menilai kemampuan generalisasi, diperlukan validasi dalam berbagai skenario. Al-Mistarehi dkk., menekankan bahwa model sebaiknya divalidasi menggunakan data dari yurisdiksi yang berbeda sebelum diimplementasikan secara operasional (Al-Mistarehi, B.-W. et al., 2022). Untuk membangun model yang lebih generalizable, diperlukan pelatihan menggunakan beragam dataset atau penerapan transfer learning agar model dapat beradaptasi dengan kondisi baru meskipun hanya tersedia sedikit data lokal.

Sistem penilaian keselamatan secara real-time masih dibatasi oleh latensi dan kapasitas pemrosesan. Arsitektur ESRM yang dikembangkan oleh Huang dkk., membutuhkan daya komputasi yang signifikan untuk melakukan inferensi machine learning, analitik video, serta pengenalan wajah (Lo, H.-W., 2022). Implementasi sistem pada skala besar menuntut algoritma yang efisien serta infrastruktur yang mampu menangani volume data yang sangat besar dengan latensi minimal. Meskipun menambah kompleksitas sistem, pendekatan edge computing

yang memproses data secara lokal di dekat sumber data, alih-alih di server pusat, berpotensi menjadi solusi untuk mengatasi keterbatasan tersebut.

Karena kecelakaan relatif jarang terjadi, dataset yang berisi observasi non-kecelakaan jauh lebih banyak dibandingkan dengan data kecelakaan itu sendiri. Ketidakseimbangan ini dapat menyebabkan model menjadi kurang sensitif dalam mendeteksi kejadian kecelakaan yang sebenarnya. Untuk mengatasi permasalahan tersebut, Mussah dkk., menggunakan metode resampling data seperti SMOTE (Synthetic Minority Over-sampling Technique) dan random under-sampling (Mussah, A. R., et al., 2022). Namun, metode-metode ini memiliki keterbatasan, terutama pada jenis kecelakaan yang jarang terjadi, dan belum sepenuhnya mampu menyelesaikan permasalahan ketidakseimbangan data secara menyeluruh.

Pengumpulan data perilaku pengemudi secara komprehensif, pelacakan lokasi, serta penggunaan pengenalan wajah menimbulkan kekhawatiran serius terkait privasi. Sistem yang dikembangkan oleh (Huang, Y., et al., 2024), mengumpulkan data lokasi secara kontinu dan citra wajah pengemudi, yang apabila tidak diamankan dengan baik dapat disalahgunakan (Huang, Y., et al., 2024). Oleh karena itu, pengelolaan data (data governance), prosedur pemberian izin, serta mekanisme keamanan perlu dirancang secara cermat untuk menyeimbangkan manfaat keselamatan dari pemantauan menyeluruh dengan perlindungan hak privasi individu. Namun demikian, kerangka regulasi yang mengatur hal ini masih dalam tahap perkembangan, sehingga sering kali belum memberikan panduan yang cukup jelas bagi pengembang dan operator sistem.

Keragaman sumber data dan teknik analisis yang semakin luas turut menghambat proses integrasi. Berbagai data dan analisis komparatif sering kali terkendala oleh perbedaan sistem yang menggunakan sistem koordinat, skema klasifikasi, atau format data yang tidak kompatibel.

He dkk., menekankan pentingnya integrasi metodologis dan pengembangan kerangka kerja yang terstandar untuk memaksimalkan pemanfaatan big data dalam penelitian keselamatan lalu lintas (He, M., et al., 2025). Untuk mencapai hal tersebut, diperlukan kolaborasi antara peneliti, praktisi, penyedia teknologi, dan lembaga regulasi dalam merumuskan serta menerapkan standar bersama.

VII. ARAH MASA DEPAN DAN PELUANG PENELITIAN

Tinjauan literatur menunjukkan beberapa arah yang menjanjikan untuk penelitian dan pengembangan lebih lanjut dalam analisis keselamatan jalan. Analisis keselamatan jalan perlu beradaptasi dengan kondisi lalu lintas campuran, di mana kendaraan yang dikemudikan manusia berinteraksi dengan kendaraan otonom dan semi-otonom seiring meningkatnya adopsi teknologi tersebut.

Penelitian di masa depan perlu mengkaji bagaimana pola kecelakaan dipengaruhi oleh teknologi kendaraan otonom, mengembangkan metode untuk mengevaluasi data komunikasi antara kendaraan dan infrastruktur (V2I) serta antarkendaraan (V2V), serta menilai dampak keselamatan dari berbagai tingkat otomatisasi. Singh dkk., menekankan bahwa pendekatan analisis keselamatan yang baru perlu dikembangkan untuk mengakomodasi karakteristik unik perilaku kendaraan otonom (Yang, C.-H., et al., 2024).

Meskipun metode explainable AI seperti SHAP telah mengalami kemajuan signifikan, aspek interpretabilitas masih memerlukan peningkatan lebih lanjut. Penelitian di masa depan perlu mengembangkan teknik yang tidak hanya memberikan skor pentingnya fitur (feature importance), tetapi juga mampu menjelaskan mekanisme kausal terjadinya kecelakaan. Hal ini menuntut integrasi antara machine learning dengan kerangka kerja causal inference serta pengetahuan domain dalam keselamatan lalu lintas. Peningkatan kemampuan explainability menjadi kunci dalam membangun kepercayaan, mengidentifikasi titik-titik intervensi, serta memastikan bahwa model benar-benar menangkap hubungan kausal yang sebenarnya, bukan sekadar korelasi semu (Mussah, A. R., et al., 2021).

Kebutuhan akan platform yang mengintegrasikan berbagai sumber data dan teknik analisis telah ditekankan dalam tinjauan sistematis (He, M., et al., 2025). Pengembangan platform terbuka dan terstandar yang memungkinkan akademisi dan praktisi mengakses data dari berbagai sumber, menggunakan teknik analisis mutakhir, serta berbagi hasil penelitian perlu menjadi fokus utama pengembangan di masa depan.

Platform semacam ini dapat mendukung pengambilan kebijakan berbasis bukti, mempercepat laju penelitian, serta mempermudah studi komparatif. Dengan

dukungan arsitektur berbasis cloud yang dilengkapi dengan perlindungan privasi yang memadai, platform tersebut juga dapat dimanfaatkan oleh organisasi dari berbagai skala.

Ketika kondisi berisiko tinggi teridentifikasi, sistem di masa depan diharapkan tidak hanya berhenti pada tahap prediksi, tetapi juga secara otomatis memicu langkah-langkah keselamatan (safety countermeasures). Intervensi tersebut dapat mencakup penyesuaian waktu lampu lalu lintas, penyebaran pesan peringatan, pengaturan batas kecepatan dinamis, hingga pengendalian kendaraan secara langsung pada kendaraan terkoneksi atau otonom.

Sistem semacam ini didasarkan pada pendekatan penilaian risiko berbasis peringatan yang dikembangkan (Wang, J., et al., 2022). Dalam pengembangan sistem intervensi yang efektif, pemahaman mengenai jenis tindakan yang paling tepat untuk berbagai kondisi sama pentingnya dengan kemampuan melakukan prediksi yang akurat.

Penelitian di masa depan perlu mengeksplorasi penerapan langkah-langkah keselamatan yang disesuaikan (tailored safety measures) berdasarkan karakteristik dan perilaku masing-masing pengemudi. Identifikasi perilaku berkendara yang tidak normal oleh Fu dkk., dapat diperluas untuk pengembangan program pelatihan serta umpan balik yang bersifat individual (Fu, C., et al., 2023). Dengan mengidentifikasi intervensi yang paling efektif untuk profil pengemudi tertentu, machine learning berpotensi mendukung strategi yang lebih terarah guna memaksimalkan peningkatan keselamatan. Namun demikian, pendekatan personalisasi ini perlu mempertimbangkan isu privasi serta potensi munculnya praktik yang bersifat diskriminatif.

Analisis keselamatan jalan perlu semakin mempertimbangkan dampak perubahan iklim terhadap kondisi cuaca dan lingkungan. Penelitian di masa depan perlu mengkaji bagaimana variasi suhu, pola curah hujan, serta kejadian cuaca ekstrem memengaruhi risiko kecelakaan. Strategi adaptasi yang bersifat proaktif dapat dikembangkan melalui integrasi model keselamatan dengan data meteorologi beresolusi tinggi serta proyeksi iklim, sehingga memungkinkan perumusan kebijakan dan tindakan mitigasi yang lebih responsif terhadap perubahan kondisi lingkungan. Untuk memastikan bahwa peningkatan keselamatan memberikan manfaat yang merata di semua

wilayah, penelitian di masa depan perlu secara khusus mempertimbangkan implikasi kesetaraan (equity implications). Analisis perlu mengkaji bagaimana risiko kecelakaan dan intervensi keselamatan terdistribusi secara adil di antara kelompok demografis maupun wilayah geografis yang berbeda.

Kajian sistematis oleh Paliotto dkk., mengenai metode analisis keselamatan jaringan jalan memberikan landasan penting untuk pengembangan kerangka evaluasi yang lebih berkeadilan (Paliotto, A., et al., 2024). Kerangka tersebut dapat menjadi dasar dalam merancang pendekatan penilaian keselamatan yang tidak hanya efektif secara teknis, tetapi juga inklusif dan adil bagi seluruh pengguna jalan.

VIII. PENUTUP

Analisis keselamatan jalan mengalami pergeseran signifikan dari pendekatan reaktif berbasis riwayat kecelakaan menuju sistem evaluasi proaktif dan real-time pada periode 2021–2025. Sistem ini ditingkatkan melalui pemanfaatan machine learning, analitik big data, teknologi kendaraan terkoneksi, serta sensor canggih.

Tinjauan terhadap 21 artikel ilmiah menunjukkan adanya kemajuan metodologis yang substansial, dengan kemampuan deteksi kecelakaan dan prediksi tingkat keparahan yang mencapai akurasi sekitar 85–92% berkat perkembangan deep learning. Sementara itu, kerangka explainable AI, khususnya SHAP, telah meningkatkan interpretabilitas model dan memperkuat kepercayaan pemangku kepentingan terhadap hasil prediksi. Metode hybrid dan ensemble juga meningkatkan efektivitas dan stabilitas kinerja model.

Jaringan sensor IoT dan telemetri kendaraan yang terkoneksi semakin memperkuat kemampuan pemantauan real-time melalui otomatisasi deteksi konflik lalu lintas serta penggunaan computer vision untuk menganalisis perilaku pengemudi. Integrasi berbagai sumber data melalui platform big data memungkinkan pelaksanaan studi keselamatan yang lebih komprehensif.

Temuan penting menunjukkan bahwa intervensi proaktif didukung oleh fusi data real-time dan surrogate safety measures, yang memberikan wawasan dari data konflik lalu lintas serta mengungkap dampak sistemik melalui analisis tingkat jaringan. Secara keseluruhan, kemajuan teknologi dalam penelitian keselamatan jalan meningkatkan

kemampuan analisis dan penerapannya, termasuk dalam prediksi, pemahaman mekanisme kecelakaan, dan komunikasi hasil. Namun demikian, tantangan terkait kualitas data, generalisasi model, dan aspek kesetaraan masih memerlukan investasi kolaboratif yang berkelanjutan untuk diatasi. Sebagai rekomendasi strategis, penelitian ke depan perlu: (1) mengembangkan platform data terbuka dan terstandar yang mendukung berbagi data lintas yurisdiksi; (2) memperkuat kerangka regulasi privasi yang adaptif terhadap perkembangan teknologi pemantauan; (3) memprioritaskan pendekatan federated learning untuk mengatasi keterbatasan berbagi data sensitif; serta (4) mendorong kolaborasi internasional dalam harmonisasi standar data keselamatan transportasi.

DAFTAR PUSTAKA

- Skaug, T. R., et al., "Road Crash Analysis and Modeling: A Systematic Review of Methods, Data, and Emerging Technologies," **Applied Sciences**, vol. 15, no. 13, 2025. (<https://doi.org/10.3390/app15137115>).
- Sangare, M., et al., "Exploring the Forecasting Approach for Road Accidents: Analytical Measures with Hybrid Machine Learning," **Expert Systems with Applications**, vol. 167, 2021. (<https://doi.org/10.1016/J.ESWA.2020.113855>).
- Yang, C., et al., "The Application of XGBoost and SHAP to Examining the Factors in Freight Truck-Related Crashes: An Exploratory Analysis," **Accident Analysis & Prevention**, vol. 158, 2021. (<https://doi.org/10.1016/J.AAP.2021.106153>).
- Al-Mistarehi, B.-W. et al., "Using Machine Learning Models to Forecast Severity Level of Traffic Crashes by R Studio and ArcGIS," **Frontiers in Built Environment**, vol. 8, 2022. (<https://doi.org/10.3389/fbuil.2022.860805>).
- Zhang, X., et al., "Attention-Based Residual Dilated Network for Traffic Accident Prediction," **Mathematics**, vol. 11, no. 9, 2023. (<https://doi.org/10.3390/math11092011>).
- Mussah, A. R., et al., "Machine Learning Framework for Real-Time Assessment of Traffic Safety Utilizing Connected Vehicle Data," **Sustainability**, vol. 14, no. 22, 2022. (<https://doi.org/10.3390/su142215348>).
- Swathi, M., et al., "Traffic Accident Prediction Using Machine Learning," **International Journal of Computer Science and Mobile Computing**, vol. 11, no. 3, 2022. (<https://doi.org/10.47760/ijcsmc.2022.v11i03.010>).
- Lo, H.-W., "A Hybrid Model Integrating FMEA and HFACS to Assess the Risk of Inter-City Bus Accidents," **Complex & Intelligent Systems**, 2022. (<https://doi.org/10.1007/s40747-022-00657-1>).
- Huang, Y., et al., "Real-Time Driver and Traffic Data Integration for Enhanced Road Safety," **IEEE Transactions on Computational Social Systems**, 2024. (<https://doi.org/10.1109/tcss.2024.3448400>).
- Wang, J., et al., "Traffic Risk Assessment Based on Warning Data," **Journal of Advanced Transportation**, 2022. (<https://doi.org/10.1155/2022/1191239>).
- Fu, C., et al., "Road Safety Risk Identification Driven by Abnormal Driving Behavior Data," **Academic Journal of Science and Technology**, vol. 5, no. 1, 2023. DOI: [10.54097/ajst.v5i1.5601] (<https://doi.org/10.54097/ajst.v5i1.5601>).
- Yang, C.-H., et al., "Application of Big Data Analysis of Traffic Accidents and Violation Reports for Improving Traffic Safety," **Sensors and Materials**, vol. 36, no. 5, 2024. DOI: [10.18494/sam4868] (<https://doi.org/10.18494/sam4868>).
- Singh, D., et al., "Surrogate Safety Analysis-Leveraging Advanced Technologies for Safer Roads," **Suranaree Journal of Science and Technology**, vol. 31, no. 4, 2024. (<https://doi.org/10.55766/sujst-2024-04-e03837>).
- He, M., et al., "Exploring the Applications of Big Data in Transportation Safety Research: A Systematic Review," **Proceedings of the International Conference on Transportation Information and Safety (ICTIS)**, 2025. (<https://doi.org/10.1109/ictis68762.2025.11214953>).
- Houssam, B., et al., "Analysis of the Severity of Transport Vehicle Accidents by a Comparative Study of Machine Learning Models," **International Journal of Advanced Technology and Engineering Exploration**, vol. 10, no. 101, 2023. (<https://doi.org/10.19101/ijatee.2023.10101879>).
- Zheng, L., et al., "Modeling Traffic Conflicts for Use in Road Safety Analysis: A Review of

- Analytic Methods and Future Directions," *Analytic Methods in Accident Research*, vol. 29, 2021. (<https://doi.org/10.1016/J.AMAR.2020.100142>).
- X. Chen et al., "Network-Level Safety Metrics for Overall Traffic Safety Assessment: A Case Study," in *IEEE Access*, vol. 11, pp. 17755-17778, 2023, doi: 10.1109/ACCESS.2022.3223046.
- Alkhadour, W., et al., "Traffic Accidents Detection Using Geographic Information Systems (GIS)," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 4, 2021. (<https://doi.org/10.14569/IJACSA.2021.0120462>).
- Hu, L. et al., "Improving Traffic Safety through Traffic Accident Risk Assessment," *Sustainability*, vol. 15, no. 4, 2023. (<https://doi.org/10.3390/su15043748>)
- Zhang, Y., et al., "Analysis of Traffic Accident Based on Knowledge Graph," *Journal of Advanced Transportation*, 2022. (<https://doi.org/10.1155/2022/3915467>).
- Paliotto, A., et al., "Systematic Review, Evaluation and Comparison of Different Approaches for the Implementation of Road Network Safety Analysis," *Heliyon*, vol. 10, no. 7, 2024. (<https://doi.org/10.1016/j.heliyon.2024.e28391>).



Copyright© Juli 2026. Zubair Saing.

