

Metode-Metode Prediksi Arus Lalu Lintas Berbasis *Internet of Things* dan Ponsel Pintar

Qornain Aji
Departemen Teknik Elektro dan
Teknologi Informasi,
Universitas Gadjah Mada
Yogyakarta, Indonesia
qornain.aji@mail.ugm.ac.id

Abstract—Telematika kendaraan berbasis IOT dengan ponsel pintar merupakan bidang yang sangat berkembang pesat. Berbagai metode telah digunakan untuk menyelesaikan permasalahan memprediksi arus lalu lintas dengan *dataset* yang besar. Dari empat kategori metode-metode prediksi arus lalu lintas, kategori metode hibridalah yang paling sesuai karena mampu memroses fitur non-linier dan lapisan-lapisan neuronnya mampu mencari pola-pola tersembunyi. Dari keempat metode yang ada pada kategori tersebut, CNN-LSTM, LSTM SAE, dan CNN + LSTM merupakan metode-metode yang direkomendasikan karena nilai MAPE yang rendah, yaitu secara berturut-turut sebesar 7,29%; 12,54%; dan 17,95%.

Keywords—*Internet of Things*, prediksi lalu lintas, kendaraan, ponsel pintar, metode hibrida

I. PENDAHULUAN

Bidang IoT (*Internet of Things*) adalah bidang yang berkembang dengan sangat pesat. Dengan IoT, peralatan rumah tangga dan peralatan kesehatan, hingga ke kendaraan dan alat pemantau lingkungan akan bisa terhubung dengan internet [1]. Perangkat atau peralatan yang terhubung ke internet tersebut dapat mengirimkan laporan data dan bisa mengambil tindakan yang bersesuaian. Laporan tersebut dapat menghasilkan sebuah data yang sangat besar sehingga dapat digunakan untuk keperluan analisa pola kebiasaan perangkat, kondisi lingkungan, dan performa dari perangkat itu sendiri. Studi menyebutkan bahwa pada tahun 2020, perangkat yang terhubung ke internet diprediksi mencapai 13 miliar [1]. Studi lanjut menyebutkan bahwa pemicu dari peningkatan perangkat yang terhubung ke internet adalah ponsel pintar [1].

Perangkat IoT juga sering disematkan pada ponsel pintar. Ponsel pintar sendiri memanfaatkan *software* aplikasi untuk menyamai kemampuan komputer dan mobilitas dari telepon seluler. Jika kita melihat perkembangan ponsel pintar dengan kemunculan iPhone generasi pertama dan Nokia N95, ponsel pintar telah menjadi bagian dari kehidupan sosial kita. Perkembangan ponsel pintar yang sangat pesat ini memiliki penjelasan, terutama karena ponsel pintar memiliki banyak sensor yang ditanamkan pada perangkatnya serta memberikan pengiriman data secara *wireless* dan interaksi sosial di dalamnya. Pemanfaatan ponsel pintar sebagai perangkat untuk mengambil koleksi data salah satunya berada pada bidang telematika.

Berkat dari perkembangan ponsel pintar yang tidak berhenti, industri kendaraan dan navigasi telah menemukan cara baru untuk mengambil data yang akan bermanfaat bagi pemilik kendaraan dan masyarakat. Dengan perolehan data baru tersebut, dilakukan penelitian secara besar-besaran baik dalam proyek industri maupun proyek akademik. Menurut

studi yang dilakukan beberapa waktu lalu, pada 2020 diprediksi sebanyak 263 miliar dollar nilai pasar dari bidang IoT [1], 45 miliar dollar merupakan bidang telematika kendaraan [1]. Hal ini menunjukkan potensi industri di bidang telematika kendaraan.

Keuntungan dari telematika kendaraan berbasis ponsel pintar adalah murah, mudah ditingkatkan, dan mudah dilakukan penskalaan. Ponsel pintar juga merupakan perangkat yang dapat memberikan masukan kepada pengemudi secara audio-visual serta memberikan integrasi pada layanan telematika dengan *social network* yang telah ada. Pendeknya umur dari jangka waktu penggunaan ponsel pintar dibandingkan dengan jangka waktu penggunaan kendaraan menjadikan ponsel pintar sebagai media untuk menawarkan berbagai kemanfaatan.

Salah satu pemanfaatan ponsel pintar yang telah disematkan perangkat IoT adalah untuk memprediksi arus lalu lintas. Prediksi arus lalu lintas dianggap sebagai masalah yang menantang dalam perencanaan transportasi dan sistem navigasi mobil. Arus lalu lintas dalam jaringan dapat diperkirakan menggunakan data arus lalu lintas historis [2]. Namun, prediksi arus lalu lintas tidak dapat hanya mengandalkan data lalu lintas masa lalu karena alasan berikut: 1) Peristiwa lalu lintas di jalan seperti kecelakaan, penutupan jalan, dll., mempengaruhi arus lalu lintas di jaringan, dan belum jelas diketahui penyebabnya, 2) parameter-parameter yang terjadi di luar jalan dapat berdampak besar pada arus lalu lintas dan mungkin tidak dimasukkan dalam data historis, dan 3) data lalu lintas tidak semuanya tersedia karena sebagian besar titik-titik jalan tidak dilengkapi dengan sensor lalu lintas [2].

Gangguan arus lalu lintas yang mempengaruhi perkiraan aliran kendaraan pada lalu lintas dapat dikategorikan sebagai dapat diprediksi dan tidak dapat diprediksi. Gangguan yang dapat diprediksi termasuk sinyal lalu lintas, rambu berhenti, layanan angkutan umum, acara olahraga terjadwal, konser musik, pembangunan/perbaikan jalan, dll. Gangguan yang tidak dapat diprediksi termasuk kecelakaan lalu lintas, kerusakan, dan penutupan jalan darurat. Dampak gangguan terhadap arus lalu lintas tergantung pada lokasi, durasi gangguan, dan permintaan pada saat gangguan. Studi mengenai dampak gangguan jenis ini terhadap arus lalu lintas meliputi [4]–[5][6].

Masalah yang muncul adalah apakah kita dapat memprediksi arus lalu lintas sebelumnya dengan memberikan informasi historis lalu lintas, informasi tentang acara terjadwal, dan data lalu lintas waktu nyata jika tersedia. Pada prinsipnya, karena gangguan yang tidak dapat diprediksi, prediksi jangka panjang mungkin tidak cukup akurat untuk

penggunaan praktis yang handal. Namun, prediksi lalu lintas jangka pendek, jika dilakukan dengan benar, dapat mencapai tingkat akurasi yang berguna untuk beberapa aplikasi jika dibandingkan dengan tanpa prediksi atau prediksi yang tidak akurat.

Ada banyak penelitian dalam literatur mengenai prediksi arus lalu lintas jangka pendek. Metode prediksi jangka pendek termasuk metode non-linier seperti metode jaringan saraf [7]–[8] [9] [10] [11] dan metode linier seperti filter Kalman [12]–[13] [14] [15] [16] dan metode rata-rata bergerak terintegrasi autoregresif (ARIMA) [17]–[18] [19] [20]. Metode ARIMA merupakan penduga linier berdasarkan nilai masa lalu dari deret waktu yang dimodelkan [21]. Sifat data dan jenis aplikasi menentukan metode yang digunakan untuk prediksi lalu lintas. Schmitt dan Jula [22] menyelidiki keterbatasan metode linier yang umum digunakan dan mengamati bahwa "waktu perjalanan masa datang dapat lebih baik diprediksi oleh prediktor gabungan." Prediktor gabungan adalah kombinasi linier dari prediktor rata-rata historis dan prediktor real-time saat ini. Gua dkk. [23] membandingkan metode pemodelan yang berbeda untuk prediksi lalu lintas jangka pendek dan menyimpulkan bahwa menggunakan metode umpan balik kesalahan dari prediksi meningkatkan akurasi prediksi dalam kondisi normal dan abnormal. Dalam studi lain, Smith et al. [24] membandingkan metode parametrik (ARIMA musiman) dan nonparametrik (regresi berbasis data) yang menunjukkan bahwa "data kondisi lalu lintas bersifat stokastik, bukan *chaos*." Selain itu, dikatakan bahwa metode ARIMA musiman memiliki kinerja yang lebih baik daripada metode regresi nonparametrik. Selanjutnya, studi eksperimental menunjukkan bahwa metode ARIMA mengungguli tolok ukur perkiraan heuristik [18]. Kinerja metode ARIMA dapat ditingkatkan dengan mempertimbangkan korelasi temporal-spasial. Metode multivariat diperkenalkan untuk memperhitungkan korelasi ini. Kamarianakis dan Prastacos [25] dan Min dan Wynter [26] mengusulkan metode rata-rata bergerak terintegrasi autoregressive ruang-waktu untuk memenuhi keterkaitan antar tautan.

Metode prediksi lalu lintas yang diberikan menjadi tidak akurat dalam kondisi sebagian datanya hilang. *Missing data* menunjukkan tidak tersedianya data lalu lintas untuk jangka waktu tertentu di suatu bagian jaringan transportasi karena kerusakan sensor atau data yang tercemar *noise*. Masalah ini sering terjadi pada jaringan transportasi [27]–[28] [29]. Banyak penelitian membahas masalah prediksi lalu lintas dengan sebagian data lalu lintas yang hilang. Misalnya, van Lint et al. [7] menyajikan jaringan saraf untuk prediksi waktu perjalanan dalam kondisi data lalu lintas yang hilang. Sundkk. [30] memperkenalkan metode Bayesian untuk meramalkan arus lalu lintas di mana periode tertentu data historis hilang untuk beberapa tautan jaringan transportasi. Bagian yang hilang dari data histori lalu lintas diperkirakan dengan menggunakan metode campuran Gaussian. Selain itu, metode statistik dan probabilitas lainnya juga digunakan untuk mengatasi masalah data lalu lintas yang hilang [31]–[32] [33] [34] [35] [36] [37]. Namun, muncul permasalahan yaitu metode-metode di atas tidak berlaku untuk jaringan transportasi dengan data arus lalu lintas yang besar, karena hanya menggunakan metode berbasis *machine learning* konvensional yang belum mampu menangkap pola-pola tersembunyi pada *dataset* yang besar. Oleh karena itu, penelitian ini bertujuan mencari metode yang sesuai yang bisa

menyelesaikan prediksi arus lalu lintas pada *dataset* yang besar.

Penelitian ini kemudian dibagi menjadi empat bagian, yakni Bab 1 adalah Pendahuluan, Bab II adalah Metodologi Penelitian, Bab III merupakan Hasil dan Pembahasan, dan Bab IV adalah Kesimpulan.

II. METODOLOGI PENELITIAN

A. Pertanyaan Penelitian

Tujuan utama dari penelitian ini adalah untuk mengetahui perkembangan terkini dari telematika sistem cerdas kendaraan berbasis IoT dan ponsel pintar di bidang navigasi. Oleh karena itu, pertanyaan penelitian berikut digunakan untuk melakukan tinjauan pustaka ini.

RQ1: Apakah metode-metode prediksi arus lalu lintas berbasis IoT dan ponsel pintar yang mampu menangani *dataset* yang besar?

Tujuan: Mengidentifikasi dan mengklasifikasikan metode-metode yang telah digunakan pada penelitian sebelumnya. Penelitian ini juga bermaksud untuk meninjau kelebihan dan kekurangan metode, serta pembahasannya.

B. Strategi Pencarian

Kata kunci juga digunakan seperti "Internet of Things" dan dengan manual memfilter dengan memilih rentang tahun studi mulai dari 2017 hingga 2021. Substring yang digunakan adalah "Traffic Flow Prediction". Penyortiran dilanjutkan dengan memilih jurnal IEEE Transactions on Intelligent Transportation Systems. Tabel string pencarian dapat dibuat, seperti Tabel 1.

TABLE I. SEARCH STRING

Concept	Sub-string	Connector
Internet of Things	Traffic Flow Prediction	-

C. Kriteria Inklusi dan Eksklusi

Beberapa syarat inklusi dan eksklusi dalam tinjauan pustaka ini juga digunakan untuk memfokuskan topik penelitian. Metode inklusi dan eksklusi sangat penting dilakukan agar tinjauan pustaka menjadi lebih efektif sehingga waktu yang dibutuhkan untuk mereview literatur menjadi lebih singkat dan efisien. Beberapa dari inklusi adalah:

IC1. Artikel mengenai telematika berkendara berbasis ponsel pintar.

IC2. Artikel yang diterbitkan dalam rentang 2017 hingga 2021.

IC3. Artikel yang diakses dari IEEE Xplore.

IC4. Artikel yang diterbitkan pada judul publikasi IEEE Transactions on Intelligent Transportation Systems.

IC5. Artikel yang menggunakan solusi Deep Learning dan Machine Learning sebagai solusi memprediksi arus lalu lintas.

Beberapa dari kriteria eksklusi adalah sebagai berikut:

- EC1. Artikel yang diterbitkan oleh selain dari IEEE Xplore.
- EC2. Artikel yang bukan memiliki judul publikasi Transactions on Intelligent Transportation Systems.
- EC3. Sumber studi yang berasal dari editorial, prolog, poster, wawancara, dan berita.
- EC4. Artikel yang membahas di luar topik telematika kendaraan berbasis ponsel pintar.
- EC5. Studi yang berbahasa selain Bahasa Inggris.

D. Proses Penseleksian

Dalam tinjauan pustaka ini, perpustakaan digital yang digunakan adalah IEEE Xplore dengan pilihan jurnal IEEE Transactions on Intelligent Transportation Systems. Pencarian literatur studi diawali dengan menemukan 76 studi. Proses penyaringan dilakukan dengan metode inklusi dan eksklusi hingga akhirnya mendapatkan 3 studi relevan yang berkaitan dengan penggunaan ponsel pintar pada layanan telematika kendaraan. Lalu, dilakukan analisa dan metode sintesis terhadap studi tersebut untuk menjawab pertanyaan di atas.

III. HASIL DAN PEMBAHASAN

Hasil dan pembahasan dari tinjauan pustaka ini dilakukan secara menyeluruh. Pertanyaan-pertanyaan penelitian dijawab dalam Bab Hasil dan Pembahasan. Pembahasan dilakukan dengan cara melakukan kategorisasi terhadap metode-metode yang digunakan pada penelitian telematika kendaraan berbasis IoT dan ponsel pintar. Setelah itu dicari kelebihan dan kelemahan pada masing-masing kategori metode yang digunakan. Berdasarkan kelebihan dan kelemahan tersebut, kategori dari metode penelitian yang terbaik akan direkomendasikan. Berdasarkan kategori metode penelitian terbaik tersebut, tabel dan grafik akan digunakan untuk mengilustrasikan metode-metode dengan akurasi yang tertinggi. Tiga buah metode akan direkomendasikan dalam tinjauan pustaka ini.

A. RQ1: Apakah metode-metode prediksi arus lalu lintas berbasis IoT dan ponsel pintar yang mampu menangani dataset yang besar?

Beberapa penelitian telah mengusulkan metode prediksi arus lalu lintas. Namun untuk mencari penelitian metode-metode prediksi arus lalu lintas berbasis IoT dan ponsel pintar tidaklah mudah. Hanya penelitian terdekatlah yang ditemukan, yaitu secara umum membahas telematika prediksi arus lalu lintas kendaraan. Metode prediksi arus lalu lintas yang ditemukan mencakup empat kategori utama yaitu metode statistik berbasis *time series*, metode nonparametrik, metode *Deep Learning*, dan metode hibrida [3].

Metode statistik berbasis deret waktu juga dikenal sebagai metode parametrik. Metode parametrik yang paling terkenal adalah ARIMA. ARIMA telah digunakan untuk prediksi arus lalu lintas pada tahun 1979 [3]. Ahmed dan Cook menggunakan teknik Box-Jenkins untuk meramalkan arus lalu lintas jalan bebas hambatan jangka pendek. Sedangkan untuk metode nonparametrik, Fu et al. [38] menggunakan SVM dari metode nonparametrik untuk mempelajari data arus lalu lintas dimensi tinggi dengan memperkenalkan fungsi kernel. Wang dkk. [39] menggunakan metode prediksi

Bayesian baru yang dikombinasikan dengan skala abu-abu yang membentuk metode hibrida untuk menganalisis koreksi historis arus lalu lintas. Namun, baik metode parametrik maupun nonparametrik tidak dapat menangani kumpulan data besar yang mengarah pada hasil yang kurang tepat [3].

Dalam beberapa tahun terakhir, metode *Deep Learning* mampu dengan baik mengekstraksi fitur implisit dan kompleks dari sejumlah besar data mentah [3]. Dengan demikian, *Deep Learning* akan menjadi metode utama dalam pemrosesan fitur non-linier dan lapisan-lapisan neuronnya mampu mencari pola-pola tersembunyi. Banyak ahli menggunakan berbagai metode *Deep Learning* untuk memprediksi arus lalu lintas dalam berbagai skenario [40], [41]. Metode *Deep Learning* dapat mempelajari sampel yang besar dan bersesuaian dengan karakteristik non-linier sampel melalui lapisan tersembunyi dan neuron. Lv dkk. [42] pertama-tama menggunakan metode penyandian otomatis yang bertumpuk untuk mempelajari fitur arus lalu lintas. Korelasi spasial dan temporal dapat dipertimbangkan dalam pemodelan ini. Hasilnya menunjukkan kinerja yang unggul. Wang dkk. [43] mengusulkan satu lapisan tersembunyi CNN berdasarkan umpan balik kesalahan untuk memprediksi kecepatan lalu lintas. Zhan dkk. [44] mengusulkan metode berbasis CNN untuk prediksi arus lalu lintas yang menggunakan data GPS taksi untuk memprediksi arus lalu lintas. Lu dkk. [45] mempertimbangkan karakteristik spasial-temporal jaringan jalan perkotaan dan menggunakan jaringan saraf BP (Back Propagation) untuk memprediksi arus lalu lintas dalam waktu singkat. Untuk lebih meningkatkan akurasi ramalan, para peneliti menggunakan jaringan LSTM yang mampu menyeimbangkan fitur jangka panjang dan jangka pendek, dan sering digunakan untuk melakukan masalah deret waktu seperti prediksi arus lalu lintas atau pemrosesan bahasa alami. Zhao dkk. [46] mengusulkan jaringan LSTM mempertimbangkan korelasi temporal-spasial dalam sistem lalu lintas yang mencapai kinerja yang lebih baik.

Meskipun metode *Deep Learning* telah memperoleh kinerja yang sangat baik daripada metode tradisional, setiap metode seperti CNN atau LSTM atau SAE menghadapi kelemahan mereka sendiri. Untuk mengatasi masalah ini, para peneliti akan menggabungkan metode yang berbeda bersamaan yang disebut *ensemble learning* [3]. Wu dkk. [47] mengusulkan arsitektur mendalam baru gabungan CLTFP + CNN-LSTM untuk meramalkan arus lalu lintas. Zheng dkk. [48] juga menggabungkan CNN dan LSTM untuk menangkap informasi lalu lintas yang halus. Zhao dkk. [49] menggunakan teori *no negative constraints theory* (NNCT) integrasi bobot dan algoritma optimasi populasi ekstrem (PEO) untuk menentukan jeda waktu dalam jaringan LSTM *ensemble*. Banyak karya telah membuktikan potensi pembelajaran *ensemble* atau metode hibrida. Namun, bagaimana memilih metode *Deep Learning* yang tepat untuk memecahkan masalah prediksi lalu lintas terhadap jaringan jalan perkotaan masih menjadi tantangan. Dengan skala kota yang terus meningkat, jaringan jalan menjadi semakin kompleks. Seperti disebutkan di atas, berkat strukturnya yang unik, LSTM mampu mempertahankan memori jangka panjang dalam peristiwa runtun waktu. Sehingga LSTM selalu menampilkan performa yang luar biasa dalam cara prediksi waktu. Lv dkk. [50] telah membuktikan bahwa SAE mampu mempelajari fitur arus lalu lintas generik dari aspek spasial. Akibatnya, menggabungkan LSTM dan SAE bersama-sama sebagai

metode hibrida adalah cara yang terbaik untuk lebih meningkatkan kinerja prediksi jaringan jalan [3].

Dari keempat kategori metode-metode prediksi arus lalu lintas yang sudah dibahas, ternyata kategori metode terakhirlah yang mampu menangani *dataset* yang besar, yaitu kategori metode hibrida. Dengan catatan penelitian-penelitian tersebut masih bersifat umum, yaitu menangani permasalahan telematika navigasi kendaraan. Hal ini diterapkan karena tinjauan pustaka yang dilakukan tidak memperoleh penelitian-penelitian prediksi arus lalu lintas berbasis IoT dan ponsel pintar yang mampu menangani *dataset* yang besar. Oleh karena itu, penelitian ini mengambil data dari metode-metode hibrida dan performanya diilustrasikan pada Tabel II.

TABLE II. PERFORMA METODE

no	Metode	<i>Dataset</i>	MAPE	MAE	RMSE
1	LSTM SAE	TDRL	12,54%		
2	CNN-LSTM + CLTFP	PeMS	7,29%	19,32%	
3	CNN+LSTM	China-KDD CUP 2017	17,95%	25,39%	44,82%
4	EnLSTM-WPEO	Freidman Test		72,15%	93,30%

Dari Tabel II, terlihat bahwa metrik yang paling sesuai digunakan adalah MAPE (Mean Absolute Percentage Error). Hal ini dilakukan karena metrik RMSE hanya dimiliki oleh dua metode serta nilainya menunjukkan error yang terlalu besar. Sedangkan dengan metrik MAE hanya menghasilkan dua buah metode saja yang mempunyai MAE dibawah 30%. Sehingga terpilih lah tiga metode, yaitu LSTM SAE, CNN-LSTM + CLTFP, dan CNN+LSTM yang grafik performanya dapat dilihat pada Fig.1.

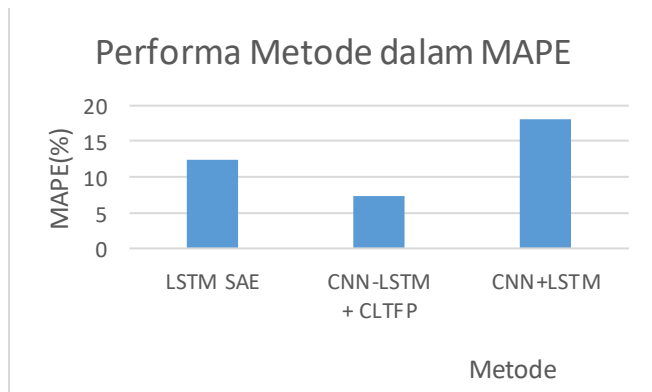


Fig.1. Performa metode dalam MAPE

Dari Fig.1. dapat dilihat bahwa urutan metode terbaik yang memiliki nilai MAPE dari terendah sampai dengan tertinggi adalah sebagai berikut: 1) CNN-LSTM 2) LSTM SAE, dan 3) CNN-LSTM + CLTFP. Nilai MAPE yang kecil menunjukkan error yang kecil yang juga mengindikasikan akurasi yang besar. Meninjau dari hasil MAE, urutan metode yang terbaik, yaitu dari yang terendah hingga yang tertinggi secara berturut-turut adalah 1) CNN-LSTM + CLTFP; 2) CNN+LSTM dan 3) EnLSTM-WPEO. Dengan menggabungkan data nilai MAPE dan nilai MAE, dapat disimpulkan bahwa metode CNN-LSTM + CLTFP memiliki akurasi yang paling besar dilanjutkan dengan metode CNN+LSTM di urutan kedua.

Metode EnLSTM-WPEO tidak direkomendasikan karena nilai dari MAE terlalu besar yaitu 72,15%. Karena *dataset* yang digunakan berbeda-beda, maka untuk memprediksi arus lalu lintas yang membutuhkan *dataset* yang besar, ketiga metode di atas direkomendasikan untuk digunakan.

IV. KESIMPULAN

Dari studi pustaka yang dilakukan, ditemukan bahwa kategori metode hibrida merupakan kategori metode yang paling sesuai dalam menangani prediksi arus lalu lintas berbasis IoT dan ponsel pintar pada *dataset* yang besar. Hal ini terjadi karena kategori metode hibrida ini mampu memroses fitur non-linier dan lapisan-lapisan neuronnya mampu mencari pola-pola tersembunyi. Dari keempat metode yang ada pada kategori tersebut, CNN-LSTM, LSTM SAE, dan CNN + LSTM merupakan metode-metode yang direkomendasikan karena nilai MAPE yang rendah, yaitu secara berturut-turut sebesar 7,29%; 12,54%; dan 17,95%.

DAFTAR PUSTAKA

- [1] J. Wahlström, I. Skog and P. Händel, "Smartphone-Based Vehicle Telematics: A Ten-Year Anniversary," in IEEE Transactions on Intelligent Transportation Systems, vol. 18, no. 10, pp. 2802-2825, Oct. 2017.
- [2] A. Abadi, T. Rajabioun and P. A. Ioannou, "Traffic Flow Prediction for Road Transportation Networks With Limited Traffic Data," in IEEE Transactions on Intelligent Transportation Systems, vol. 16, no. 2, pp. 653-662, April 2015.
- [3] C. Chen, Z. Liu, S. Wan, J. Luan and Q. Pei, "Traffic Flow Prediction Based on Deep Learning in Internet of Vehicles," in IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 6, pp. 3776-3789, June 2021.
- [4] RAND, Moving Los Angeles: Short-Term Transportation Policy Options for Improving Transportation, RAND Corporation, 2008.
- [5] "Traffic choices study: Summary report," Seattle, WA, USA, 2008.
- [6] T. P. Mamuneas and M. I. Nadri, "Contribution of highway capital to industry and national productivity growth," Transportation Research Board, Washington, DC, USA, 1996.
- [7] J. W. C. van Lint, S. P. Hoogendoorn, and H. J. van Zuylen, "Accurate freeway travel time prediction with state-space neural networks under missing data," Transp. Res. Part C Emerg. Technol., vol. 13, no. 5/6, pp. 347-369, Oct.-Dec. 2005.
- [8] J. Yu, G.-L. Chang, H. W. Ho, and Y. Liu, "Variation based online travel time prediction using clustered neural networks," in Proc. 11th Int. IEEE Conf. Intell. Transp. Syst., Oct. 2008, pp. 85-90.
- [9] K. Y. Chan, T. S. Dillon, J. Singh, and E. Chang, "Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and Levenberg-Marquardt algorithm," IEEE Trans. Intell. Transp. Syst., vol. 13, no. 2, pp. 644-654, Jun. 2012.
- [10] D. Park, L. R. Rilett, and G. Han, "Spectral basis neural networks for realtime travel time forecasting," J. Transp. Eng., vol. 125, no. 6, pp. 515-523, Nov. 1999.
- [11] Q. Y. Q. Ye, S. C. Wong, and W. Y. Szeto, "Short-term traffic speed forecasting based on data recorded at irregular intervals," in Proc. 13th Int. IEEE Conf. ITSC, 2010, pp. 1541-1546.
- [12] I. Okutani, "The Kalman filtering approaches in some transportation and traffic problems," Transp. Res. Rec., vol. 2, no. 1, pp. 397-416, 1987.
- [13] Y. Xie, Y. Zhang, and Z. Ye, "Short-term traffic volume forecasting using Kalman filter with discrete wavelet decomposition," Comput.-Aided Civil Infrastruct. Eng., vol. 22, no. 5, pp. 326-334, Jul. 2007.
- [14] I. Okutani and Y. J. Stephanedes, "Dynamic prediction of traffic volume through Kalman filtering theory," Transp. Res. Part B Methodol., vol. 18, no. 1, pp. 1-11, 1984.
- [15] H. J. H. Ji, A. X. A. Xu, X. S. X. Sui, and L. L. L. Li, "The applied research of Kalman in the dynamic travel time prediction," in Proc. 18th Int. Conf. Geoinformat., 2010, pp. 1-5.

- [16] Y. Wang and M. Papageorgiou, "Real-time freeway traffic state estimation based on extended Kalman filter: A general approach," *Transp. Res. Part B Methodol.*, vol. 39, no. 2, pp. 141–167, Feb. 2005.
- [17] M. Van Der Voort, M. Dougherty, and S. Watson, "Combining Kohonen maps with ARIMA time series models to forecast traffic flow," *Transp. Res. Part C Emerg. Technol.*, vol. 4, no. 5, pp. 307–318, Oct. 1996.
- [18] B. M. Williams, M. Asce, L. A. Hoel, and F. Asce, "Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results," *J. Transp. Eng.*, vol. 129, no. 6, pp. 664–672, Nov. 2003.
- [19] D. Billings and J. S. Yang, "Application of the ARIMA models to urban roadway travel time," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, 2006, pp. 2529–2534.
- [20] C. Chen, J. Hu, Q. Meng, and Y. Zhang, "Short-time traffic flow prediction with ARIMA-GARCH model," in *Proc. IEEE Intell. Veh. Symp.*, 2011, pp. 607–612.
- [21] B. M. Williams, P. K. Durvasula, and D. E. Brown, "Urban freeway traffic flow prediction application of seasonal autoregressive integrated," *Transp. Res. Rec.*, vol. 1644, pp. 132–141, 1998.
- [22] E. J. Schmitt and H. Julia, "On the limitations of linear models in predicting travel times," in *Proc. IEEE Intell. Transp. Syst. Conf.*, Sep. 2007, pp. 830–835.
- [23] F. Guo, J. W. Polak, and R. Krishnan, "Comparison of modeling approaches for short term traffic prediction under normal and abnormal conditions," in *Proc. 13th Int. IEEE Conf. Intell. Transp. Syst.*, 2010, pp. 1209–1214.
- [24] B. L. Smith, B. M. Williams, and R. Keith Oswald, "Comparison of parametric and nonparametric models for traffic flow forecasting," *Transp. Res. Part C Emerging Technol.*, vol. 10, no. 4, pp. 303–321, Aug. 2002.
- [25] Y. Kamarianakis and P. Prastacos, "Space-time modeling of traffic flow," *Comput. Geosci.*, vol. 31, no. 2, pp. 119–133, Mar. 2005.
- [26] W. Min and L. Wynter, "Real-time road traffic prediction with spatiotemporal correlations," *Transp. Res. Part C Emerging Technol.*, vol. 19, no. 4, pp. 606–616, Aug. 2011.
- [27] L. Q. L. Qu, J. H. J. Hu, L. L. Li, and Y. Z. Y. Zhang, "PPCABased missing data imputation for traffic flow volume: A systematic approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 3, pp. 512–522, Sep. 2009.
- [28] C. Chen, Y. Wang, L. Li, J. Hu, and Z. Zhang, "The retrieval of intraday trend and its influence on traffic prediction," *Transp. Res. Part C Emerging Technol.*, vol. 22, pp. 103–118, 2012.
- [29] B. Smith, W. Scherer, and J. Conklin, "Exploring imputation techniques for missing data in transportation management systems," *Transp. Res. Rec.*, vol. 1836, pp. 132–142, 2003.
- [30] S. Sun, C. Zhang, G. Yu, N. Lu, and F. Xiao, "Bayesian network methods for traffic flow forecasting with incomplete data," in *Proc. Mach. Learn., ECML*, 2004, pp. 419–428.
- [31] L. Li, Y. Li, and Z. Li, "Efficient missing data imputing for traffic flow by considering temporal and spatial dependence," *Transp. Res. Part C Emerging Technol.*, vol. 34, pp. 108–120, 2013.
- [32] J. Haworth and T. Cheng, "Non-parametric regression for space-time forecasting under missing data," *Comput. Environ. Urban Syst.*, vol. 36, no. 6, pp. 538–550, Nov. 2012.
- [33] H. Tan et al., "A tensor-based method for missing traffic data completion," *Transp. Res. Part C Emerging Technol.*, vol. 28, pp. 15–27, Mar. 2013.
- [34] A. Khosravi et al., "Prediction intervals to account for uncertainties in travel time prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 537–547, Jun. 2011.
- [35] S. S. S. Sun and C. Z. C. Zhang, "The selective random subspace predictor for traffic flow forecasting," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 2, pp. 367–373, Jun. 2007.
- [36] S. Sun and X. Xu, "Variational inference for infinite mixtures of Gaussian processes with applications to traffic flow prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 2, pp. 466–475, Jun. 2011.
- [37] S. Sun, R. Huang, and Y. Gao, "Network-scale traffic modeling and forecasting with graphical lasso and neural networks," *J. Transp. Eng.*, vol. 138, no. 11, pp. 1358–1367, Nov. 2012.
- [38] G. Fu, G. Han, F. Lu and Z. Xu, "Short-term traffic flow forecasting model based on support vector machine regression", *J. South China Univ. Technol. (Natural Sci. Ed.)*, vol. 41, no. 9, pp. 71–76, 2013.
- [39] J. Wang, W. Deng and Y. Guo, "New Bayesian combination method for short-term traffic flow forecasting", *Transp. Res. C Emerg. Technol.*, vol. 43, pp. 79–94, Jun. 2014.
- [40] Y. Chen, L. Shu and L. Wang, "Traffic flow prediction with big data: A deep learning based time series model", *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, pp. 1010–1011, May 2017.
- [41] N. G. Polson and V. O. Sokolov, "Deep learning for short-term traffic flow prediction", *Transp. Res. C Emerg. Technol.*, vol. 79, pp. 1–17, Jun. 2017.
- [42] Y. Lv, Y. Duan, W. Kang, Z. Li and F.-Y. Wang, "Traffic flow prediction with big data: A deep learning approach", *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 865–873, Sep. 2014.
- [43] J. Wang, Q. Gu, J. Wu, G. Liu and Z. Xiong, "Traffic speed prediction and congestion source exploration: A deep learning method", *Proc. IEEE 16th Int. Conf. Data Mining (ICDM)*, pp. 499–508, Dec. 2016.
- [44] X. Zhan, Z. Yu, X. Yi and S. Ukkusuri, "Citywide traffic volume estimation using trajectory data", *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 2, pp. 272–285, Feb. 2017.
- [45] W. Li, Y. Ge and L. Chen, "Short-term traffic flow forecasting of urban expressway based on space-time fusion", *Comput. Simul.*, vol. 35, no. 9, pp. 136–140, 2018.
- [46] Z. Zhao, W. Chen, X. Wu, P. C. Y. Chen and J. Liu, "LSTM network: A deep learning approach for short-term traffic forecast", *IET Intell. Transp. Syst.*, vol. 11, no. 2, pp. 68–75, Mar. 2017.
- [47] Y. Wu and H. Tan, "Short-term traffic flow forecasting with spatial-temporal correlation in a hybrid deep learning framework", *Comput. Vis. Pattern Recognit.*, vol. 23, no. 1, pp. 1–14, 2016.
- [48] Z. Zheng, Y. Yang, J. Liu, H.-N. Dai and Y. Zhang, "Deep and embedded learning approach for traffic flow prediction in urban informatics", *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 10, pp. 3927–3939, Oct. 2019.
- [49] F. Zhao, G.-Q. Zeng and K.-D. Lu, "EnLSTM-WPEO: Short-term traffic flow prediction by ensemble LSTM NNCT weight integration and population extremal optimization", *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 101–113, Jan. 2020.
- [50] Y. Lv, Y. Duan, W. Kang, Z. Li and F.-Y. Wang, "Traffic flow prediction with big data: A deep learning approach", *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 865–873, Sep. 2014.
- [51] R. Song, W. Sun, B. Zheng and Y. Zheng, "PRESS: A novel framework of trajectory compression in road networks", *Proc. VLDB Endowment*, vol. 7, no. 9, pp. 661–672, 2014.
- [52] N. Mitrovic, M. T. Asif, U. Rasheed, J. Dauwels and P. Jaillet, "CUR decomposition for compression and compressed sensing of large-scale traffic data", *Proc. 16th Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, pp. 1475–1480, Oct. 2013.
- [53] A. Khoshgozaran, A. Khodaei, M. Sharifzadeh and C. Shahabi, "A hybrid aggregation and compression technique for road network databases", *Knowl. Inf. Syst.*, vol. 17, no. 3, pp. 265–286, Dec. 2008.
- [54] Z. Zhao, Y. Zhang, J. Hu and L. Li, "Comparative study of PCA and ICA based traffic flow compression", *J. Highway Transp. Res. Develop. (English Ed.)*, vol. 4, no. 1, pp. 98–102, Jun. 2009.
- [55] J. Gai, J. Wang, S. Cheng and X. Liu, "A synchronous compression method based on compressed sensing for time-interleaved sampling", *Proc. Int. Conf. Meas. Inf. Control*, vol. 1, pp. 59–62, May 2012.