

AI-Based Sensor Information Fusion

Subjects: Computer Science, Information Systems

Contributor: Carson Leung

In recent years, artificial intelligence (AI) and its subarea of deep learning have drawn the attention of many researchers. At the same time, advances in technologies enable the generation or collection of large amounts of valuable data (e.g., sensor data) from various sources in different applications, such as those for the Internet of Things (IoT), which in turn aims towards the development of smart cities. With the availability of sensor data from various sources, sensor information fusion is in demand for effective integration of big data.

Keywords: sensor ; information fusion ; sensor fusion ; artificial intelligence (AI) ; deep learning ; supervised learning ; data mining ; transportation ; geographic information system (GIS) ; global navigation satellite system (GNSS) ; global positioning system (GPS)

1. Introduction

Recent advances in technology have increased the popularity of the area of artificial intelligence (AI) ^{[1][2]}, which aims to build “intelligent agents” with the ability to correctly interpret external data, learn from these data, and use the learned knowledge for cognitive tasks ^[3] like reasoning, planning, problem solving, decision making, motion and manipulation. Subareas of AI include robotics, computer vision, natural language processing (NLP), and machine learning ^{[4][5][6][7]}. Within the latter, deep learning ^{[8][9][10]} has attracted the focus of many researchers. For instance, the development of AlphaGo (which uses deep reinforcement learning) for the board game of Go ^[11] has drawn the attention of researchers and the general public. In general, deep learning uses deep neural networks (DNNs), convolutional neural networks (CNNs), as well as recurrent neural networks (RNNs) like long short-term memory (LSTM) for supervised, semi-supervised, or unsupervised learning tasks ^{[12][13][14]} in various application areas like computer vision and NLP. Recently, deep learning has also been applied to the transportation domain ^{[15][16]}, but for tasks like traffic flow forecasting, automatic vehicle detection, autonomous driving, and classification of speeding drivers.

Moreover, recent advances in technology have also enabled the generation or collection of large amounts of valuable data from a wide variety of sources in different real-life applications ^{[17][18][19][20][21][22]}. For instance, different types of sensor data can be easily generated and collected in various Internet of Things (IoT) ^{[23][24]} applications—such as smart homes, smart grids, smart retail, smart cars, and smart cities ^{[25][26]}. As an example, sensors (e.g., cameras; digital scanners; light imaging, detection, and ranging (LIDAR) ^[27]) mounted on aircrafts, small unmanned aerial vehicles (UAVs) ^[28], and other moving objects such as vehicles ^[29] have created large volumes of remotely sensed data, geospatial data, spatial-temporal data, and geographic information for the geographic information system (GIS) ^{[30][31]}. As another example, sensors on the global navigation satellite system (GNSS) ^[32]—such as the Global Positioning System (GPS) ^[33], GLONASS, Gaillieo and Beidou (which are originated in the USA, Russia, the EU and China, respectively), as well as other regional systems—have also created large volumes of geolocation and time information. With big sensory data from these sources and other sensors, sensor information fusion ^{[34][35][36][37]}—which integrates sensor data and information from a collection of these heterogeneous or homogeneous sensors to produce accurate, consistent and useful knowledge—is in demand.

2. Traditional Methods for Urban Data Analytics and Machine Learning

Urban data mining helps discover useful knowledge from urban data, which in turn helps solve some urban issues. For instance, the discovery of popular transportation modes (e.g., bicycles) of residents in a city helps city planners to take appropriate actions (e.g., add more bike lanes). To mine these urban data, researchers have traditionally been using paper-based and telephone-based travel surveys ^[38]. Unfortunately, these travel surveys can be biased and contain inaccurate data about the movements of their participants. For instance, participants tend to under-report short trips, irregular trips, and car trips. They also tend to over-report public transit trips ^{[39][40]}.

Alternatively, researchers have also been using commute diaries [41][42], which capture data about people's daily commutes. Unfortunately, these diaries can be error-prone. When people are asked to use a diary to keep track of their commutes, they often forget to record their commutes throughout the day. When trips are recorded at the end of the day, diary studies can then inherit the same problems as paper-based and telephone-based travel surveys. Moreover, these diaries can also be a mental burden to study the participants, and thus cannot be used long term [43]. Furthermore, as people's willingness to record trips accurately throughout the day declines with each day of participation, the corresponding accuracy of the commute diaries also drops [44].

3. Sensor-Based Methods for Urban Data Analytics and Machine Learning

Recent advances in technology have led to the availability of sensor data, which in turn have led to better approaches for urban data mining. To elaborate, sensors enable users to track a large number of movement trajectories that are collected by participants of a study who use GNSS/GPS trackers or other sensors (e.g., accelerometers, barometers, Bluetooth, Wi-Fi, etc.). Hence, these *GPS-based travel surveys* [45][46] are more accurate than the travel surveys and commute diaries. However, the challenge of labeling trip purposes and classifying transportation modes persists. For instance, the manual segmentation of trajectories based on transportation mode can be labor intensive and is likely to be impracticable for big data [47]. Any AI approach to automating such a task would obviously be beneficial to travel studies and other applications (e.g., IoT applications) that rely on contextual knowledge (e.g., the current travel mode of a person). For example, a driver would benefit from receiving a notification from his smartphone or smartwatch about an estimated arrival time for his trip (computed based on his current location, destination, and his interaction or saved frequently visited locations). As another example, urban analysts would benefit from an automatic trip transportation mode labeling method in a way is similar to timeline in Google Maps (which keeps track of a user's location history and attempts to automatically classify trips with the major transportation mode). However, existing trip transportation mode labeling methods were not very accurate, needed corrections by the user, and do not track when transportation modes were changed. Hence, a more accurate method is needed.

Consider the use of standalone tracking and logging devices, which enables the participants of travel surveys to log sensor data accurately, reliably, and consistently as they have full control over the device and the hardware and software platforms are the same on every device. These devices can log data to local device storages, which are then collected for data retrieval. These devices can also connect to a smartphone application on a participant's phone via Bluetooth and collect data regularly at intervals for further processing. To a further extent, transportation mode classification could happen on a smartphone, which then could reduce the computational burden on the logger device, decrease both architecture cost (as it requires weaker processing units) and power consumption, and thus increase the battery life. Among related works, Zheng et al. [48] used supervised decision trees and graph-based post-processing after classification to classify transportation modes from GPS data only.

In contrast, Hemminki et al. [49] used only accelerometer data to classify transportation modes ("stationary", "walk", "bus", "tram", "train", "metro"). To elaborate, three different classifiers were trained with a combination of AdaBoost and Hidden Markov Model (HMM) for three different classes of modes. Shafique and Hato [50] also used accelerometer data only. They applied multiple machine learning algorithms to perform transportation mode classification and found that the Random Forest algorithm [51] gave accurate classifications.

4. Sensor Fusion-Based Methods for Urban Data Analytics and Machine Learning

Instead of using only GPS data or only accelerometer data, Ellis et al. [52] applied the Random Forest to both GPS data and accelerometer data for successful transportation mode classification.

Other than using both GPS data and accelerometer data, Hosseinyalamdary et al. [29] used both GIS and GPS data (together with an inertial measurement unit (IMU)). However, they used these data for tracking three-dimensional (3D) moving objects rather than classifying transportation modes. On the hand, Chung and Shalaby [53] developed a system that uses both GPS and GIS data to classify four transportation modes—"walk", "bicycle", "bus" and "car"—for GPS-based travel surveys by using a rule-based algorithm and a map-matching algorithm [54] to detect the exact roads people moved on. However, the accuracy of the system is dependent on the corresponding GIS data. Similarly, Stenneth et al. [55] also used both GPS and GIS data when building their real-time transportation mode classification system. To perform the classification, they used the Random Forest as the supervised learning algorithm to identify a person's current transportation mode.

5. Summary

To recap, traditional methods for urban data mining include paper-based and telephone-based travel surveys [38][39][40], as well as commute diaries [41][42][43][44]. To reduce the human workload and to utilize sensors and AI technologies for automatic processes, GPS-based travel surveys [45][46][47] were used. In recent years, advances in technologies have enabled the use of some combinations of data from different sensors (e.g., GNSS/GPS, accelerometers) and other modern smartphone sensors (e.g., barometer, magnetometer, etc.). Some related works [48] use only GPS data, while some others [49][50] use only accelerometer data. In addition, some related works [52] integrate both GPS and accelerometer data (i.e., an example of sensor information fusion), while some others [53][54][55] integrate both GPS and GIS data (i.e., another example of sensor information fusion). However, none of the aforementioned works combines GNSS/GPS, accelerometer, and GIS data in a single system. Hence, a system that integrates GNSS/GPS, accelerometer, and GIS data for urban data analytics and machine learning is in demand.

References

1. Guo, K.; Lu, Y.; Gao, H.; Cao, R. Artificial intelligence-based semantic Internet of Things in a user-centric smart city. *Sensors* 2018, 18, 1341.
2. Sandino, J.; Pegg, G.; Gonzalez, L.F.; Smith, G. Aerial mapping of forests affected by pathogens using UAVs, hyperspectral sensors, and artificial intelligence. *Sensors* 2018, 18, 944.
3. Deng, D.; Leung, C.K.; Wodi, B.H.; Yu, J.; Zhang, H.; Cuzzocrea, A. An innovative framework for supporting cognitive-based big data analytics for frequent pattern mining. In *Proceedings of the IEEE ICCD 2018, San Francisco, CA, USA, 2–7 July 2018*; IEEE Computer Society: Los Alamitos, CA, USA, 2018; pp. 49–56.
4. Brown, J.A.; Cuzzocrea, A.; Kresta, M.; Kristjanson, K.D.L.; Leung, C.K.; Tebinka, T.W. A machine learning system for supporting advanced knowledge discovery from chess game data. In *Proceedings of the IEEE ICMLA 2017, Cancun, Mexico, 18–21 December 2017*; IEEE Computer Society: Los Alamitos, CA, USA, 2017; pp. 649–654.
5. Leung, C.K.; MacKinnon, R.K.; Wang, Y. A machine learning approach for stock price prediction. In *Proceedings of the IDEAS 2014, Porto, Portugal, 7–9 July 2014*; ACM: New York, NY, USA, 2014; pp. 274–277. [Green Version]
6. Morris, K.J.; Egan, S.D.; Linsangan, J.L.; Leung, C.K.; Cuzzocrea, A.; Hoi, C.S. Token-based adaptive time-series prediction by ensembling linear and non-linear estimators: A machine learning approach for predictive analytics on big stock data. In *Proceedings of the IEEE ICMLA 2018, Orlando, FL, USA, 17–20 December 2018*; IEEE Computer Society: Los Alamitos, CA, USA, 2018; pp. 1486–1491.
7. Zhang, L.; Xiao, N.; Yang, W.; Li, J. Advanced heterogeneous feature fusion machine learning models and algorithms for improving indoor localization. *Sensors* 2019, 19, 125.
8. Islam, M.; Sohaib, M.; Kim, J.; Kim, J. Crack classification of a pressure vessel using feature selection and deep learning methods. *Sensors* 2018, 18, 4379.
9. Xiao, L.; Zhang, Y.; Peng, G. Landslide susceptibility assessment using integrated deep learning algorithm along the China-Nepal Highway. *Sensors* 2018, 18, 4436.
10. Strauß, S. From big data to deep learning: A leap towards strong AI or ‘intelligentia obscura’? *Big Data Cogn. Comput.* 2018, 2, 16.
11. Leung, C.K.; Kanke, F.; Cuzzocrea, A. Data analytics on the board game Go for the discovery of interesting sequences of moves in joseki. *Procedia Comput. Sci.* 2018, 126, 831–840.
12. Castagno, J.; Atkins, E. Roof shape classification from LiDAR and satellite image data fusion using supervised learning. *Sensors* 2018, 18, 3960.
13. Li, M.; Li, Q.; Liu, G.; Zhang, C. Generative adversarial networks-based semi-supervised automatic modulation recognition for cognitive radio networks. *Sensors* 2018, 18, 3913.
14. Wang, J.; Sanchez, J.A.; Ayesta, I.; Iturrioz, J.A. Unsupervised machine learning for advanced tolerance monitoring of wire electrical discharge machining of disc turbine fir-tree slots. *Sensors* 2018, 18, 3359.
15. Bhavsar, P.; Safro, I.; Bouaynaya, N.; Polikar, R.; Dera, D. Machine learning in transportation data analytics. In *Data Analytics for Intelligent Transportation Systems*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 283–307.
16. Nguyen, H.; Kieu, L.; Wen, T.; Cai, C. Deep learning methods in transportation domain: A review. *IET Intell. Transp. Syst.* 2018, 12, 998–1004.

17. Braun, P.; Cameron, J.J.; Cuzzocrea, A.; Jiang, F.; Leung, C.K. Effectively and efficiently mining frequent patterns from dense graph streams on disk. *Procedia Comput. Sci.* 2014, 35, 338–347.
18. Jiang, F.; Leung, C.K. A data analytic algorithm for managing, querying, and processing uncertain big data in cloud environments. *Algorithms* 2015, 8, 1175–1194.
19. Lakshmanan, L.V.S.; Leung, C.K.; Ng, R.T. The segment support map: Scalable mining of frequent itemsets. *ACM SIG KDD Explor.* 2000, 2, 21–27.
20. Leung, C.K. Frequent itemset mining with constraints. In *Encyclopedia of Database Systems*, 2nd ed.; Springer: New York, NY, USA, 2018; pp. 1531–1536.
21. Li, K.C.; Jiang, H.; Yang, L.T.; Cuzzocrea, A. *Big data: Algorithms, analytics, and applications*; CRC Press: Boca Raton, FL, USA, 2015.
22. Wu, Z.; Yin, W.; Cao, J.; Xu, G.; Cuzzocrea, A. Community detection in multi-relational social networks. In *Proceedings of the WISE 2013, Nanjing, China, 13–15 October 2013*; Springer: Heidelberg, Germany, 2013; pp. 43–56.
23. Braun, P.; Cuzzocrea, A.; Leung, C.K.; Pazdor, A.G.M.; Tanbeer, S.K.; Grasso, G.M. An innovative framework for supporting frequent pattern mining problems in IoT environments. In *Proceedings of the ICCSA 2018, Part V, Melbourne, Australia, 2–5 July 2018*; Springer: Heidelberg, Germany, 2018; pp. 642–657.
24. Drenoyanis, A.; Raad, R.; Wady, I.; Krogh, C. Implementation of an IoT based radar sensor network for wastewater management. *Sensors* 2019, 19, 254.
25. Leung, C.K.; Braun, P.; Pazdor, A.G.M. Effective classification of ground transportation modes for urban data mining in smart cities. In *Proceedings of the DaWaK 2018, Regensburg, Germany, 3–6 September 2018*; Springer: Heidelberg, Germany, 2018; pp. 83–97.
26. Morales Lucas, C.; de Mingo López, L.F.; Gómez Blas, N. Natural computing applied to the underground system: A synergistic approach for smart cities. *Sensors* 2018, 18, 4094.
27. Wang, C.; Ji, M.; Wang, J.; Wen, W.; Li, T.; Sun, Y. An improved DBSCAN method for LiDAR data segmentation with automatic Eps estimation. *Sensors* 2019, 19, 172.
28. Popescu, D.; Dragana, C.; Stoican, F.; Ichim, L.; Stamatescu, G. A collaborative UAV-WSN network for monitoring large areas. *Sensors* 2018, 18, 4202.
29. Hosseinyalamdary, S.; Balazadegan, Y.; Toth, C. Tracking 3D moving objects based on GPS/IMU navigation solution, laser scanner point cloud and GIS data. *ISPRS Int. J. Geo-Inf.* 2015, 4, 1301–1316.
30. Ait Lamqadem, A.; Pradhan, B.; Saber, H.; Rahimi, A. Desertification sensitivity analysis using MEDALUS model and GIS: A case study of the oases of Middle Draa Valley, Morocco. *Sensors* 2018, 18, 2230.
31. Burrough, P.A.; McDonnell, R.A.; Lloyd, C.D. *Principles of Geographical Information Systems*, 3rd ed.; Oxford University Press: Oxford, UK, 2015.
32. Robustelli, U.; Baiocchi, V.; Pugliano, G. Assessment of dual frequency GNSS observations from a Xiaomi Mi 8 Android smartphone and positioning performance analysis. *Electronics* 2019, 8, 91.
33. Zimmermann, F.; Schmitz, B.; Klingbeil, L.; Kuhlmann, H. GPS multipath analysis using Fresnel zones. *Sensors* 2019, 19, 25.
34. Choi, S.; Cho, S. Sensor information fusion by integrated AI to control public emotion in a cyber-physical environment. *Sensors* 2018, 18, 3767.
35. de la Iglesia, D.H.; Villarrubia, G.; de Paz, J.F.; Bajo, J. Multi-sensor information fusion for optimizing electric bicycle routes using a swarm intelligence algorithm. *Sensors* 2017, 17, 2501.
36. Jing, L.; Wang, T.; Zhao, M.; Wang, P. An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox. *Sensors* 2017, 17, 414.
37. Kim, H.; Suh, D. Hybrid particle swarm optimization for multi-sensor data fusion. *Sensors* 2018, 18, 2792.
38. Murakami, E.; Wagner, D.P.; Neumeister, D.M. Using global positioning systems and personal digital assistants for personal travel surveys in the United States. In *Proceedings of the International Conference on Transport Survey Quality and Innovation 1997, Grainau, Germany, 24–30 May 1997*; Transportation Research Board: Washington, DC, USA, 2000; p. III-B.
39. Ettema, D.; Timmermans, H.; van Veghel, L. *Effects of Data Collection Methods in Travel and Activity Research*; SWOV Institute for Road Safety Research: Den Haag, The Netherlands, 1997.
40. Stopher, P.R. Household travel surveys: Cutting-edge concepts for the next century. In *Proceedings of the Conference on Household Travel Surveys 1995, Irvine, CA, USA, 12–15 March 1995*; Transportation Research Board: Washington,

DC, USA, 1995; pp. 11–23.

41. Maat, K.; Timmermans, H.J.P.; Molin, E. A model of spatial structure, activity participation and travel behavior. In Proceedings of the WCTR 2004, Istanbul, Turkey, 4–8 July 2004; Institute for Transport Studies, University of Leeds: Leeds, UK, 2004; pp. 2–14.
42. Stopher, P.R. Use of an activity-based diary to collect household travel data. *Transportation* 1992, 19, 159–176.
43. Schlich, R.; Axhausen, K.W. Habitual travel behaviour: Evidence from a six-week travel diary. *Transportation* 2003, 30, 13–36.
44. Arentze, T.; Dijst, M.; Dugundji, E.; Joh, C.; Kapoen, L.; Krygsman, S.; Maat, K.; Timmermans, H. New activity diary format: Design and limited empirical evidence. *Transp. Res. Rec.* 2001, 1768, 79–88.
45. Forrest, T.; Pearson, D. Comparison of trip determination methods in household travel surveys enhanced by a global positioning system. *Transp. Res. Rec.* 2005, 1917, 63–71.
46. Wolf, J.; Guensler, R.; Bachman, W. Elimination of the travel diary: Experiment to derive trip purpose from global positioning system travel data. *Transp. Res. Rec.* 2001, 1768, 125–134.
47. Biljecki, F.; Ledoux, H.; van Oosterom, P. Transportation mode-based segmentation and classification of movement trajectories. *Int. J. Geogr. Inf. Sci.* 2013, 27, 385–407.[Green Version]
48. Zheng, Y.; Chen, Y.; Li, Q.; Xie, X.; Ma, W. Understanding transportation modes based on GPS data for web applications. *ACM Trans. Web (TWEB)* 2010, 4, 1.
49. Hemminki, S.; Nurmi, P.; Tarkoma, S. Accelerometer-based transportation mode detection on smartphones. In Proceedings of the ACM SenSys 2013, Rome, Italy, 11–14 November 2013; ACM: New York, NY, USA, 2013; p. 13.
50. Shafique, M.A.; Hato, E. Use of acceleration data for transportation mode prediction. *Transportation* 2015, 42, 163–188.
51. Breiman, L. Random forests. *Mach. Learn.* 2001, 45, 5–32.
52. Ellis, K.; Godbole, S.; Marshall, S.; Lanckriet, G.; Staudenmayer, J.; Kerr, J. Identifying active travel behaviors in challenging environments using GPS, accelerometers, and machine learning algorithms. *Front. Public Health* 2014, 2, 36.
53. Chung, E.; Shalaby, A. A trip reconstruction tool for GPS-based personal travel surveys. *Transp. Plan. Technol.* 2005, 28, 381–401.
54. Greenfeld, J. Matching GPS observations to locations on a digital map. In Proceedings of the Transportation Research Board 81st Annual Meeting, Washington, DC, USA, 13–17 January 2002; Transportation Research Board: Washington, DC, USA, 2002.
55. Stenneth, L.; Wolfson, O.; Yu, P.S.; Xu, B. Transportation mode detection using mobile phones and GIS information. In Proceedings of the ACM SIGSPATIAL GIS 2011, Chicago, IL, USA, 1–4 November 2011; ACM: New York, NY, USA, 2011; pp. 54–63.