POLSKIE TOWARZYSTWO INFORMACJI PRZESTRZENNEJ
ROCZNIKI GEOMATYKI 2021 ○ Tom XIX ○ Zeszyt 2(93): 31–43

Artykuł naukowy

Modelling population density using artificial neural networks from open data

Modelowanie gęstości ludności z wykorzystaniem sztucznych sieci neuronowych na podstawie otwartych danych

Adam Nadolny

Wroclaw University of Environment and Life Sciences, Institute of Geodesy and Geoinformatics
Wroclaw Institute of Spatial Information and Artificial Intelligence

Abstract

This paper introduces the concept of creating a model for population density prediction and presents the work done so far. The unit of reference in the study is more the population density of a location rather than tracking human movements and habits. Heterogeneous open data, which can be obtained from the World Wide Web, was adopted for the analysis. Commercial telephony data or social networking applications were intentionally omitted. Both for data collection and later for modeling the potential of artificial neural networks was used. The potential of detection models such as YOLO or ResNet was explored. It was decided to focus on a method of acquiring additional data using information extraction from images and extracting information from web pages. The BDOT database and statistical data from the Central Statistical Office (polish: GUS) were adopted for the base model. It was shown that the use of street surveillance cameras in combination with deep learning methods gives an exam.

Keywords: population density, artificial neural networks, detection models, image information extraction
Słowa kluczowe: gęstość ludności, sztuczne sieci neuronowe, modele detekcyjne, ekstrakcja informacji z obrazu

Introduction

The development and availability of new technologies and applications make it possible to track humans more and more accurately and collect huge amounts of data (Zheng, 2015). There have already been many studies on population flow prediction of people's behavior and habits (Xia et al, 2015; Xia et al. 2016, Kong et al, 2016; Kong et al, 2017; Kong et al, 2018). There are often dilemmas about encroaching too boldly on people's privacy (Smolak et al, 2020). However, accessing datasets from, for example, cell phones in Poland is expensive. Fortunately, there has been a trend towards opening up a variety of data resources. A review of the literature identified several aspects to address in this topic (Wang et al 2019). Data acquisition techniques: it is necessary to design new data acquisition technologies to collect multi-source data in the era of Big Data (Yao, Leh & Gil, 2021). Multi-source heterogeneous data provides an excellent opportunity to discover the mobility patterns of people in cities: Data analytics based on multi-source datasets still needs in-depth research. This paper will move in the realm of these two issues however in terms of population density. To assess the population in one urban zone in Vienna, the residential population from the local and national census was disaggregated at the level of individual building blocks. Downscaling is based on a dasymetric approach using an urban land use map as supporting information (Tenerelli, Gallego & Ehrlich, 2015). Following a similar approach, a basic population density model was created for further research purposes. Based on street monitoring, drones, PoI locations (DataFromSky), tools were created to support traffic control, which is in line with the Smart City concept and leads to a reduction of the emission load from traffic (Adamec et al, 2020). In the era of widespread use of neural networks, research has been undertaken on this issue as well. Literature analysis has indicated good results when combining multiple models to address the weakness of each model (called Ensembling). The results are better by a few percentage points, making it increasingly used (Shafizadeh-Moghadam et al 2018; Pham et al, 2019; Pham et al, 2020; Costache et al, 2020; Sachdeva, Bhatia & Verma, 2020; Smolak et al 2020). In recent years, an adaptation of the well-known "word2vec" method from natural language processing has been used. It has been called "loc2vec" in the context of location prediction (Tsubouchi, Kobayashi & Shimizu, 2020; Sassi, 2019; Tang 2015; Panphattarasap & Calway, 2018). This method gives good results in machine learning. It is a different perspective on how to represent location, especially when it comes to population movement prediction. In later stages of this work, this method will be adapted for population density prediction. Different types of neural networks can be used to extract information from an image. The most common in the literature and widely implemented in Python language are detection networks. Recent work on models is mainly to improve the prediction speed, its accuracy and real-time computation capability. For real-time object detection in very short time and with relatively cheap devices, it is hard to use YOLOv4 networks. A shallow model called Concatenated Feature Pyramid Network (CFPN) (Chen et al, 2019) was intended to fulfill such a task. Work on models that are fast and lightweight but still efficient is having a positive effect. The ThunderNet model provides a prediction speed of over 20 frames per second with an AP50 of about 41% (Qin et al, 2019). Less than two years later, the Scaled-YOLOv4 model is developed, which with a complex architecture achieves 16 FPS with an AP50 of 73%, while the tiny version achieves as much as 443 FPS with the AP50 dropping to 42% (Wang, Bochkovskiy & Liao, 2021). In future work, this method will be adapted for population density prediction. Different types of neural networks can be used to extract information from an image. The most common in the literature and widely implemented in Python language are detection networks. Recent work on models is mainly to improve the prediction speed, its accuracy and real-time computation capability. For real-time object detection in very short time and with relatively cheap devices, it is hard to use YOLOv4 networks. A shallow model called Concatenated Feature Pyramid Network (CFPN) (Chen et al, 2019) was intended to fulfill such a task. Work on models that are fast and lightweight but still efficient is having a positive effect. The ThunderNet model provides a prediction speed of over 20 frames per second with an AP50 of about 41% (Qin et al, 2019). Less than two years later, the Scaled-YOLOv4 model is developed, which with a complex architecture achieves 16 FPS with an AP50 of 73%, while the tiny version achieves as much as 443 FPS with an AP50 dropping to 42% (Wang, Bochkovskiy & Liao, 2021). In addition to the architecture of each layer, the models also operate on various loss metrics. It is based on the metrics that the network can tell if it is learning correctly. The most popular metric is IoU (Intersect over Union). The research space can be seen in this area as well. An alternative is GIoU (generalized IoU) (Rezatofighi et al, 2019). Further improvements to the metric were D-IoU (Distance IoU), and C-IoU (Complete IoU). In the former, a normalized distance between bboxes was added. In the latter, more parameters were added: center and aspect ratio (Zheng et al 2019). The following paper will outline the idea of using publicly available data to predict instantaneous population density and present the work done so far and the results obtained. For extraction of some information artificial neural networks were used, others were taken directly from open resources. In the further process of modeling the phenomenon also the potential of artificial neural networks will be used and the concept of using loc2vec transformation and ensembling will be presented.

Stages completed and plans

Up to now, scripts have been made to collect cyclic data from web pages. Of course, this phase will be continually developed with new data sources constantly emerging. A graph database from the network of streets, paths and sidewalks has been made. The next stage, and the most important in the context of this paper, was to test the effectiveness of the method for extracting traffic data from intersection monitoring images. A key aspect of this stage is the verification of an alternative data source and acquisition method with subsequent analysis of the time dependence of traffic. There are plans to create a more detailed model sensitive to subsequent data sources. The model is expected to provide greater spatial and temporal resolution and provide greater predictive accuracy. Eventually, a dynamic population density model is to be developed.

Data collection

The background for all collected data will be the database of topographic objects (polish: BDOT). It provides the most basic information about buildings, roads, sidewalks and bicycle paths. Land use, functions of buildings and number of storeys are also useful information. The road network provides a bus for data on car, bicycle and pedestrian transport. Wroclaw was chosen as the test site. The city has been providing more and more interesting information for quite some time now. The basic information is the number of inhabitants by sex and age with the spatial resolution corresponding to the statistical area (size of a small district). Other data more related to population movement are: routes, schedules and even locations of public transport vehicles; number of rentals of city bicycles with travel times, starting and ending stations; occupancy of paid and guarded parking lots throughout the city; detailed map of current bicycle paths, available images from surveillance cameras at many major intersections, information about current defects and road works. A great deal of data can be collected. It will be crucial to analyze the usefulness and impact of the data on the accuracy of the model. The main focus of this paper is on extracting information from images (in this case city surveillance cameras).

Data collection: obtaining directly

From the Wroclaw Open Data portal, it is possible to obtain tabular information on the number of free/occupied parking spaces, city bicycle rentals, and public transport timetables with stops. Such data needed to be geolocated. Point vector layers were created that corresponded to the collected data. Using Python libraries, a script was created to automatically collect the above data into a database. In order to well represent and locate the data sources in space, a vector layer of each source was created. Public transportation stops have been described in such a way, so that the route of a particular line is known. Bicycle stops are independent from each other. Rentals can occur at any station and lead to any station as well. It is possible to rely only on rental nodes and return nodes. The same is true for parking. It is only known if a car has entered or left the parking lot. Additional information can only be the direction of the parking exit (into a specific street). A point vector layer was also created from this information. In order to connect the whole city and all source elements, the street grid had to be properly prepared. For this purpose, a graphical data model was created based on road/bike lines along with points that feed the model with data.

Data collection: street monitoring

Artificial neural networks were used to collect statistical data from intersection monitoring. For several days the material was collected from about 30 cameras. Each of them has different resolution and image quality, also the perspective and distance from the intersection is different. The frequency of downloading the view from a camera is one minute. It is easy to imagine the number of acquired images. About a thousand images from different cameras were used to prepare the learning data. It was decided to determine only one class at the beginning: vehicles. In later studies, classes such as cars, motorcycles, bicycles, and pedestrians are likely to be extracted. In each image there were from a few to a few dozen objects. The purpose of learning the model is to automatically and continuously collect information from intersections. The model counts the objects present in the image and stores this number into a working database (Fig. 1).

Fig.1. On the left : A view from a camera; On the right: A map presenting cameras location

The acquired data along with the geolocation of the camera/intersection will be important information for the population density study. This is information that in near real time can add to the model. Another way to use this data is to study the traffic volume as a function of time (time of day, time of week, time of year) and then use these relationships in the model. The above issue was tried to be solved with several detection network models. Yolo-v4 models in two variations, and ResNet model were tested. The learning results (learning time, prediction time, prediction accuracy and model size are shown in Tab. 1).

Tabele 1: Comparing of models results

	training time	model weight	mAP@0.50/mAP@0.75	prediction time
yolov4	7 h 32 m	244 MB	79.51%/35.10%	1 sek.
csdarknet	7 h 39 m	244 MB	79.51%/35.10%	1 sek.
csresnet	6 h 8 m	221MB	76.93%/29.64%	1 sek.

According to the tests, several important parameters were obtained that influenced the selection of a particular neural network model. Learning time, which is the time it took to learn the model to a particular epoch. This does not have much impact during later prediction. It can make a difference during model learning. The prediction time is an important parameter if we want to apply the model in real time. When the image is acquired once per minute, this parameter has no significant impact. On the other hand, prediction accuracy is important. It allows to determine the effectiveness of model detection, i.e. the probability with which the model acquires data for further purposes. The last element is the size of the model. In other words, how much disk space a stored model takes up. This affects the computer memory used during prediction. Like prediction time, this parameter is of marginal importance if we use a powerful desktop computer with a graphics card for prediction. It matters if we would like to use the model on mobile devices. Analyzing the obtained results, we see that two of the proposed models are practically no different from each other. Comparing yolov4 with resnet we can be tempted to make a comparison. Resnet learned slightly shorter. Finally, the model takes up less space on the disk. The element that determines the use of the model is the achieved accuracy. In the table it is expressed as mAP (mean average precision) based on IoU parameter. Resnet showed a much lower accuracy, so the YOLOv4 model was used for further work.

Data usage

Model building began with the preparation of BDOT data and statistical data provided by the Central Statistical Office. The dataset includes information on population by statistical unit, age, and gender. The data was used to calculate the average population density. The area of residential buildings (information from BDOT database - building functions) multiplied by the number of stories was used for the calculation. In this way, the theoretical residential area expressed in square meters was obtained. From statistical data, information about the population in the area was taken. By dividing the population by the residential area, the average density per square meter was obtained. Knowing the square meters of individual buildings, a specific number of inhabitants was assigned to them. This data was used as a reference unit. In this way, a base model was made (Fig. 2), on which the next planned stages will be carried out. Density is assigned to a building, which will be the primary source to begin density work. BDOT database is also information about land cover and land use. To cover the whole area with potential population density, the class "PT" was adopted. (Land Cover). This is a series of layers that complement each other and closely fill the entire area of the country (Tab. 2). Each class has at least one detailed class. One of the land cover classes is land use for buildings (residential, commercial, retail, education, religious, etc.). The information about the population of a building was used to give the population density in each built-up area. In addition to residential buildings, the density in buildings with other functions was also estimated. It is important to clarify here that the model presented assumes the presence of people in several places simultaneously. It certainly happens that when looking at an area in aggregate, the population is greater than the statistics indicate. In the next stages of work the model will take more dynamic shape and values closer to real ones. Different population densities have been assigned for different land cover areas. In addition to areas under development, there are land classes less frequented by people. Assuming different densities depending on the class, an overall model of potential population density was obtained.

Fig.2. Wroclaw's population density based on census data – centroids

Table 2 Category of surface layers used for the initial model (Own source)

GROUP	CLASS	NAME
PT	PTWP	surface water
	PTZB	buildings
	PTLZ	forest and wooded area
	PTRK	shrubby vegetation
	PTUT	permanent crop
	PTTR	grassland and agricultural land
	PTKM	land under roads, railways and airports
	PTGN	unused land
	PTPL	yard
	PTSO	landfill
	PTWZ	pit and heap
	PTNZ	remaining undeveloped land

Use of the model

From the prepared detection network models, YOLOv4 was selected. Then a Python script was prepared to collect statistics from cameras placed at intersections. The model performed vehicle detection on camera images and counted them. It collected information about passing vehicles approximately every minute for 24 hours. Such monitoring is run all the time to collect long term data. As the months go by, the model will be trained on new data. Over the course of a day/month/year, the greening of the environment, the light (its color and how long it shines during the day) changes. In general, the characteristics of the image provided by the cameras change. Therefore it is important to constantly improve the model.

Observations will serve as additional information modeling the variability of traffic on the streets in time or directly as an input to the base. Monitoring the phenomenon so far allowed us to observe and confirm the generally accepted correlations of "rush hour" with road traffic. Daily dependencies are definitely visible (Fig.3.).

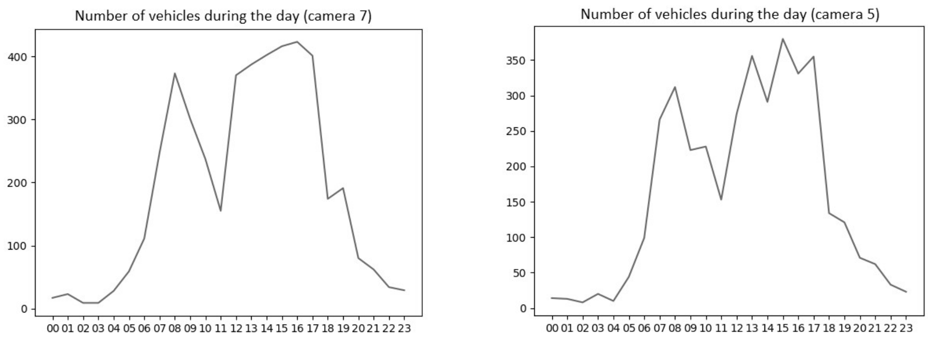

Figure 3: Graph of the number of vehicles during the day

Conclusions

The use of artificial neural networks to obtain additional sources of population movement data is reasonable. Experiments have demonstrated the model's ability to count passing vehicles. This, in turn, can be a valuable data source in the context of studying population density at a given location at a given time. Further work should be carried out to improve the model's operation and effectiveness. One might be tempted to specify more classes in order to collect more detailed information. Further exploitation of the potential of this data source will be subject to further experimentation and publication.

Acknowledgments

Thanks to Wroclaw Institute of Spatial Information and Artificial Intelligence (WIZIPISI Sp. z o.o.) for the opportunity to acquire skills in using artificial intelligence.

Funding

The research was primarily created with own funds. Only for learning neural networks hardware of Wroclaw Institute of Spatial Information and Artificial Intelligence was used.

Literatura (References)

Adamec, V., Herman, D., Schullerova, B., & Urbanek, M. (2019). Modelling of Traffic Load by the DataFromSky System in the Smart City Concept. Smart Governance for Cities: Perspectives and Experiences EAI/Springer Innovations in Communication and Computing,135-152. doi:10.1007/978-3-030-22070-9_7

Chen, P., Hsieh, J., Gochoo, M., Wang, C., & Liao, H. M. (2019). Smaller Object Detection for Real-Time Embedded Traffic Flow Estimation Using Fish-Eye Cameras. 2019 IEEE International Conference on Image Processing (ICIP). doi:10.1109/icip.2019.8803719

Costache, R., Pham, Q. B., Sharifi, E., Linh, N. T., Abba, S., Vojtek, M., . . . Khoi, D. N. (2019). Flash-Flood Susceptibility Assessment Using Multi-Criteria Decision Making and Machine Learning Supported by Remote Sensing and GIS Techniques. Remote Sensing,12(1), 106. doi:10.3390/rs12010106

Kong, X., Xia, F., Wang, J., Rahim, A., & Das, S. K. (2017). Time-Location-Relationship Combined Service Recommendation Based on Taxi Trajectory Data. IEEE Transactions on Industrial Informatics,13(3), 1202-1212. doi:10.1109/tii.2017.2684163

Kong, X., Li, M., Li, J., Tian, K., Hu, X., & Xia, F. (2018). CoPFun: An urban co-occurrence pattern mining scheme based on regional function discovery. World Wide Web,22(3), 1029-1054. doi:10.1007/s11280-018-0578-x

Liao, Y., Yeh, S., & Gil, J. (2021). Feasibility of estimating travel demand using geolocations of social media data. Transportation. doi:10.1007/s11116-021-10171-x

Panphattarasap, P., & Calway, A. (2018). Automated Map Reading: Image Based Localisation in 2-D Maps Using Binary Semantic Descriptors. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). doi:10.1109/iros.2018.8594253

Pham, B. T., Phong, T. V., Nguyen-Thoi, T., Trinh, P. T., Tran, Q. C., Ho, L. S., . . . Prakash, I. (2020). GIS-based ensemble soft computing models for landslide susceptibility mapping. Advances in Space Research,66(6), 1303-1320. doi:10.1016/j.asr.2020.05.016

Pham, Q. B., Abba, S. I., Usman, A. G., Linh, N. T., Gupta, V., Malik, A., . . . Tri, D. Q. (2019). Potential of Hybrid Data-Intelligence Algorithms for Multi-Station Modelling of Rainfall. Water Resources Management,33(15), 5067-5087. doi:10.1007/s11269-019-02408-3

Qin, Z., Li, Z., Zhang, Z., Bao, Y., Yu, G., Peng, Y., & Sun, J. (2019). ThunderNet: Towards Real-Time Generic Object Detection on Mobile Devices. 2019 IEEE/CVF International Conference on Computer Vision (ICCV). doi:10.1109/iccv.2019.00682

Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I., & Savarese, S. (2019). Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). doi:10.1109/cvpr.2019.00075

Sachdeva, S., Bhatia, T., & Verma, A. K. (2019). A novel voting ensemble model for spatial prediction of landslides using GIS. International Journal of Remote Sensing,41(3), 929-952. doi:10.1080/01431161.2019.1654141

Sassi, A., Brahimi, M., Bechkit, W., & Bachir, A. (2019). Location Embedding and Deep Convolutional Neural Networks for Next Location Prediction. 2019 IEEE 44th LCN Symposium on Emerging Topics in Networking (LCN Symposium). doi:10.1109/lcnsymposium47956.2019.9000680

Shafizadeh-Moghadam, H., Valavi, R., Shahabi, H., Chapi, K., & Shirzadi, A. (2018). Novel forecasting approaches using combination of machine learning and statistical models for flood susceptibility mapping. Journal of Environmental Management,217, 1-11. doi:10.1016/j.jenvman.2018.03.089

Smolak, K., Rohm, W., Knop, K., & Siła-Nowicka, K. (2020). Population mobility modelling for mobility data simulation. Computers, Environment and Urban Systems,84, 101526. doi:10.1016/j.compenvurbsys.2020.101526

Smolak, K., Kasieczka, B., Fialkiewicz, W., Rohm, W., Siła-Nowicka, K., & Kopańczyk, K. (2020). Applying human mobility and water consumption data for short-term water demand forecasting using classical and machine learning models. Urban Water Journal,17(1), 32-42. doi:10.1080/1573062x.2020.1734947

Tang, K., Paluri, M., Fei-Fei, L., Fergus, R., & Bourdev, L. (2015). Improving Image Classification with Location Context. 2015 IEEE International Conference on Computer Vision (ICCV). doi:10.1109/iccv.2015.121

Tenerelli, P., Gallego, J. F., & Ehrlich, D. (2015). Population density modelling in support of disaster risk assessment. International Journal of Disaster Risk Reduction,13, 334-341. doi:10.1016/j.ijdrr.2015.07.015

Tsubouchi, K., Kobayashi, H., & Shimizu, T. (2020). POI Atmosphere Categorization Using Web Search Session Behavior. Proceedings of the 28th International Conference on Advances in Geographic Information Systems. doi:10.1145/3397536.3422196

Wang, C.-Y., Bochkovskiy, A., & Liao, H.-Y. M. (2021). Scaled-YOLOv4: Scaling Cross Stage Partial Network. arXiv [cs.CV]. Opgehaal van http://arxiv.org/abs/2011.08036

Wang, J., Kong, X., Xia, F., & Sun, L. (2019). Urban Human Mobility. ACM SIGKDD Explorations Newsletter,21(1), 1-19. doi:10.1145/3331651.3331653

Xia, F., Liu, L., Li, J., Ahmed, A. M., Yang, L. T., & Ma, J. (2015). BEEINFO: Interest-Based Forwarding Using Artificial Bee Colony for Socially Aware Networking. IEEE Transactions on Vehicular Technology,64(3), 1188-1200. doi:10.1109/tvt.2014.2305192

Xia, F., Liu, L., Jedari, B., & Das, S. K. (2016). PIS: A Multi-Dimensional Routing Protocol for Socially-Aware Networking. IEEE Transactions on Mobile Computing,15(11), 2825-2836. doi:10.1109/tmc.2016.2517649

Yang, Q., Wang, J., Song, X., Kong, X., Xu, Z., & Zhang, B. (2015). Urban Traffic Congestion Prediction Using Floating Car Trajectory Data. Algorithms and Architectures for Parallel Processing Lecture Notes in Computer Science,18-30. doi:10.1007/978-3-319-27122-4_2

Yun, S., Han, D., Chun, S., Oh, S. J., Yoo, Y., & Choe, J. (2019). CutMix: Regularization Strategy to Train Strong Classifiers With Localizable Features. 2019 IEEE/CVF International Conference on Computer Vision (ICCV). doi:10.1109/iccv.2019.00612

Zheng, Y. (2015). Trajectory Data Mining. ACM Transactions on Intelligent Systems and Technology,6(3), 1-41. doi:10.1145/2743025

Zheng, Z., Wang, P., Liu, W., Li, J., Ye, R., & Ren, D. (2019, November 19). Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression. Retrieved from https://arxiv.org/abs/1911.08287

Streszczenie

W niniejszej pracy przedstawiono koncepcję stworzenia modelu do predykcji gęstości ludności oraz przedstawiono wykonane dotychczas prace. Jednostką odniesienia w badaniach jest bardziej gęstość ludności w danym miejscu niż śledzenie ruchów i nawyków człowieka. Do analizy przyjęto heterogeniczne otwarte dane, które można pozyskać z sieci WWW. Celowo pominięto komercyjne dane telefonii czy aplikacji społecznościowych. Zarówno do gromadzenia danych jak i później do modelowania wykorzystano potencjał sztucznych sieci neuronowych. Zbadano potencjał modeli detekcyjnych takich jak YOLO czy ResNet. Postanowiono skupić się na metodzie pozyskiwania dodatkowych danych z wykorzystaniem ekstrakcji informacji z obrazu oraz pozyskiwania informacji ze stron WWW. Do modelu bazowego przyjęto bazę danych BDOT oraz dane statystyczne z GUS. Wykazano, że wykorzystanie kamer monitoringu ulic w połączeniu z metodami głębokiego uczenia daje egzamin.

Dane autorów / Authors details:
mgr inż. Adam Nadolny
ORCID 0000-0002-8886-182X
adam.nadolny@upwr.edu.pl
adam.nadolny@wizipisi.pl

Przesłano / Received	20.10.2021
Zaakceptowano / Accepted	25.11.2021
Opublikowano / Published	05.12.2021