Comparing the urban environment with socioeconomic characteristics using features extracted from aerial imagery

Melanie Green*[1] and Dani Arribas-Bel†[1]

[1]Department of Geography and Planning, University of Liverpool

17 January 2019

Summary

Information about the characteristics and activities of humans becomes encoded in the landscape as it is continually shaped by the people who live there. These socioeconomic and demographic characteristics are traditionally measured using census data and represented using geodemographic classifications. Remotely sensed images at a high resolution contain a huge amount of detail on the urban environment and its potential to reveal complex patterns is unlocked by deep learning. This paper uses convolutional neural networks to extract features from aerial imagery and uses clustering to classify areas and compare their physical features to their socioeconomic features.

KEYWORDS: socioeconomic characteristics, urban environment, aerial imagery, clustering, convolutional neural networks

1. Introduction

This paper investigates the relationship between socioeconomic characteristics and the urban landscape using high-resolution aerial imagery and deep learning. This will be done by classifying areas based on features extracted from aerial imagery and examining the differences in socioeconomic characteristics between clusters. The people who live in a neighbourhood play a part in shaping its landscape, and the differences are often visible between neighbourhoods with different socioeconomic characteristics. High-resolution aerial imagery holds a vast amount of data on the features of the urban environment, including individual objects and their arrangement in space, and this information can be extracted using deep learning. Classifying neighbourhoods using this information can complement the information on the built environment that is traditionally used in geodemographic classifications, and applies methods that are often used for supervised problems to an unsupervised one.

2. Understanding Neighbourhoods

Two of the main elements which constitute a neighbourhood according to the literature include the similarity of residents' characteristics and the individuality of the physical appearance of the neighbourhood (Batey and Brown, 1995). The former arises due to the tendency for people with similar characteristics and values to cluster spatially (Cheshire, 2012), or as it is often said, 'birds of a feather flock together'. The physical appearance differs between neighbourhoods as it reflects the characteristics of the society that created it (Taubenbock et al, 2009).

Understanding the characteristics of neighbourhoods is relevant to many academic fields including urban economics, sociology and public health, as well as policymakers for targeting facilities and resources where they are most needed, and companies who want to know where their customers are most likely to be located. To aid in the understanding of the characteristics of neighbourhoods, cluster analysis is often used to reduce the data from many variables into one (Webber and Burrows, 2018).

* m.o.green@liverpool.ac.uk

† d.arribas-bel@liverpool.ac.uk

Geodemographic classifications are the primary example of this. Census data is usually used as it is the most complete and reliable source of socioeconomic data (Vickers and Rees, 2007). Often some information about the built environment is included in geodemographic classifications (housing type) but this has been restricted by the limited availability and detail of data.

Although they cannot directly measure socioeconomic characteristics, aerial images hold a large amount of information on the physical environment which can be exploited using deep learning. Previous research has found relationships between the physical environment of neighbourhoods and their socioeconomic characteristics, for example the Liverpool study mentioned by Webber and Burrows (2018, p.61) found that housing stock was one of the main differences between categories of area which were classified using census data. Further examples can be found in the review by Patino and Duque (2013, p.5-6).

3. Social Science Applications of Remote Sensing

Remotely sensed imagery has been used in social science research to investigate the relationship between the physical features of cities and human characteristics, originating in the 1950s (Monier and Green, 1957). Applications have included the detection of deprivation hot spots, quality of life assessment, house value estimation and social vulnerability assessment among others (Patino and Duque, 2013). Satellite imagery has often been used to measure socioeconomic characteristics of the population in countries where survey data is scarce or inaccurate, for example Jean et al (2016), but it also offers the opportunity to complement survey data where it already exists.

The use of convolutional neural networks (CNNs) – a form of deep learning used to analyse images – is a recent development in remote sensing and, combined with high-resolution imagery, has the potential to analyse more information than ever on the features of the physical environment of cities. For example, Albert et al (2017) quantified the similarities/differences between urban environments with the same land use in different European cities. Transfer learning – using the CNN for a different problem than it was originally trained to solve – is generally used to avoid the time required to train a CNN from scratch. This can be done by fine-tuning the weights of a pre-trained network, for example (Jean et al, 2016; Albert et al, 2017), or simply using a pre-trained network, which is the least time-consuming option. Within remote sensing, features extracted from pre-trained networks have mainly been used for land use classification (Hu et al, 2015; Marmanis et al, 2016). There is potential to apply these methods to understanding the socioeconomic landscape of cities.

4. Methods

A pre-trained ResNet50 (He et al, 2016) convolutional neural network (CNN) was used to extract features from 25cm resolution aerial imagery of Liverpool. A regular grid of images was used as input to the CNN, and labelled with the corresponding Output Area Classification (OAC) supergroup from the output area covering the majority of the square. This allowed for comparison with socioeconomic characteristics. The OAC is a geodemographic classification based on census data, which classifies areas into one of eight supergroups based on their socioeconomic characteristics (Gale et al, 2016).

Two different sets of pre-trained weights were used in the CNN: one trained on ImageNet and one trained on aerial imagery and Ordnance Survey topographic data (TopoNet) (Sargent et al, 2017). The features were extracted from the penultimate layer of the network and used as input to a k-means clustering algorithm. This grouped together areas which were found by the network to have similar features in the images. Clustergrams were used to determine the optimal number of clusters for each classification. To show for each cluster how the presence of each supergroup compares to the average across Liverpool as a whole, the location quotient was calculated.

5. Results and Discussion

Tables 1-2 show the location quotients of OAC supergroups in each cluster. The image features appear to distinguish between socioeconomic groups, as the location quotients vary, sometimes strongly, between clusters. For example, in Table 1, residents of areas in cluster 0 are more likely than average to be from the Urbanites group, while in cluster 2 they are more likely to be from the Hard-Pressed Living group.

Some clusters have one socioeconomic group with a high location quotient, suggesting that the urban environment features are particularly characteristic of that socioeconomic group. Clusters where 2 or more socioeconomic groups have a high location quotient suggest that the differences between socioeconomic groups are not always distinct in the urban environment, and different socioeconomic groups can live in areas which look similar.

In Table 2 the location quotients have more extreme values, suggesting that TopoNet captures a stronger relationship between the image features and socioeconomic characteristics. Also, the number of clusters is greater, suggesting that TopoNet picks up on more detailed differences between different types of area. This is expected as TopoNet is trained on aerial imagery while the images ImageNet is trained on are completely different in nature.

Further work is necessary to understand which features of the urban environment the CNN and clustering algorithm find to be most useful for distinguishing between areas with different socioeconomic characteristics.

Table 1 Location quotients for ImageNet

Cluster	Constrained City Dwellers	Cosmopolitans	Ethnicity Central	Hard-Pressed Living	Multicultural Metropolitans	Suburbanites	Urbanites
0	0.94	1.03	0.96	0.94	0.64	0.91	1.36
1	1.38	1.37	1.23	0.99	1.31	0.57	0.68
2	0.53	0.65	0.18	1.44	0.31	0.96	1.19
3	0.81	0.74	1.06	0.9	1.08	1.47	1.08

Table 2 Location quotients for TopoNet

Cluster	Constrained City Dwellers	Cosmopolitans	Ethnicity Central	Hard-Pressed Living	Multicultural Metropolitans	Suburbanites	Urbanites
0	1.29	1.59	0.58	1.14	0.56	0.71	0.61
1	1.16	0.79	1.16	1.01	1.08	0.83	1.01
2	0.93	0.86	1.3	1.03	1.18	1.08	0.89
3	0.89	0.67	0.75	1.08	0.7	0.78	1.46
4	1.03	1.35	0.85	0.77	0.79	1.36	1.05
5	1.07	0.63	1.41	1.05	1.46	0.93	1.84
6	0.3	0.91	0.18	1.39	0.33	1.11	1.26
7	1.23	1.58	0.7	0.85	0.99	0.77	1.07

Acknowledgements

This work was funded by the Economic and Social Research Council and Ordnance Survey through the Data Analytics and Society Centre for Doctoral Training. Data was provided by Ordnance Survey.

Biographies

Melanie Green is a PhD student at the Geographic Data Science Lab, Department of Geography and Planning, University of Liverpool. Her current research focuses on investigating the relationship between the physical features and human characteristics of cities using machine learning.

References

Albert, A., Kaur, J., & Gonzalez, M.C. (2017). Using Convolutional Networks and Satellite Imagery to Identify Patterns in Urban Environments at a Large Scale. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '17*.

Batey, P., & Brown, P.J.B. (1995). From human ecology to customer targeting: The evolution of geodemographics. In P. Longley & G. Clarke (Eds.), *GIS for Business and Service Planning* (pp. 77–103). Cambridge: GeoInformation International.

Cheshire, P. (2012). Why do Birds of a Feather Flock Together? Social Mix and Social Welfare: A Quantitative Appraisal. In G. Bridge, T. Butler, & L. Lees (Eds.), *Mixed Communities: Gentrification by Stealth?* (pp. 17–24). Bristol: Policy Press.

Gale, C. G., Singleton, A. D., Bates, A. G., & Longley, P. A. (2016). Creating the 2011 area classification for output areas (2011 OAC). *Journal of Spatial Information Science*, *12*, 1–27.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

Hu, F., Xia, G. S., Hu, J., & Zhang, L. (2015). Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. *Remote Sensing*, *7*(11), 14680–14707.

Jean, N., Burke, M., Xie, M., Davis, W.M., Lobell, D.B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, *353*(6301), 790–794.

Marmanis, D., Datcu, M., Esch, T., & Stilla, U. (2016). Deep learning earth observation classification using ImageNet pretrained networks. *IEEE Geoscience and Remote Sensing Letters*, *13*(1), 105–109.

Monier, R.B., & Green, N.E. (1957). Aerial Photographic Interpretation And The Human Geography Of The City. *The Professional Geographer*, *9*(5), 2–5.

Patino, J.E., & Duque, J.C. (2013). A review of regional science applications of satellite remote sensing in urban settings. *Computers, Environment and Urban Systems*, *37*, 1–17.

Sargent, I., Hare, J., Young, D., Wilson, O., Doidge, C., Holland, D., & Atkinson, P.M. (2017). Inference and Discovery in Remote Sensing Data with Features Extracted Using Deep Networks. In M. Bramer & M. Petridis (Eds.), *Artificial Intelligence XXXIV* (pp. 131–136). Cham: Springer International Publishing.

Taubenböck, H., Wurm, M., Setiadi, N., Gebert, N., Roth, A., Strunz, G., Birkmann, J. & Dech, S. (2009). Integrating remote sensing and social science: The correlation of urban morphology with socioeconomic parameters. In *2009 Joint Urban Remote Sensing Event* (pp. 1–7).

Vickers, D., & Rees, P. (2007). Creating the UK National Statistics 2001 output area classification. *Journal of the Royal Statistical Society. Series A: Statistics in Society*, *170*(2), 379–403.

Webber, R., & Burrows, R. (2018). *The Predictive Postcode*. London: Sage.