

Urban Grammar

“Learning from Deep Learning”

Lessons from using computer vision to identify (urban) form and function in open data satellite imagery

#AAG2023

Dani Arribas-Bel

@darribas

Martin Fleischmann

@martinfleis



The
Alan Turing
Institute



Geographic
Data Science
Lab

“Previous season...”



The Alan Turing Institute



Menu

[Home](#) + [Research](#) + [Research projects](#)

Urban Grammar

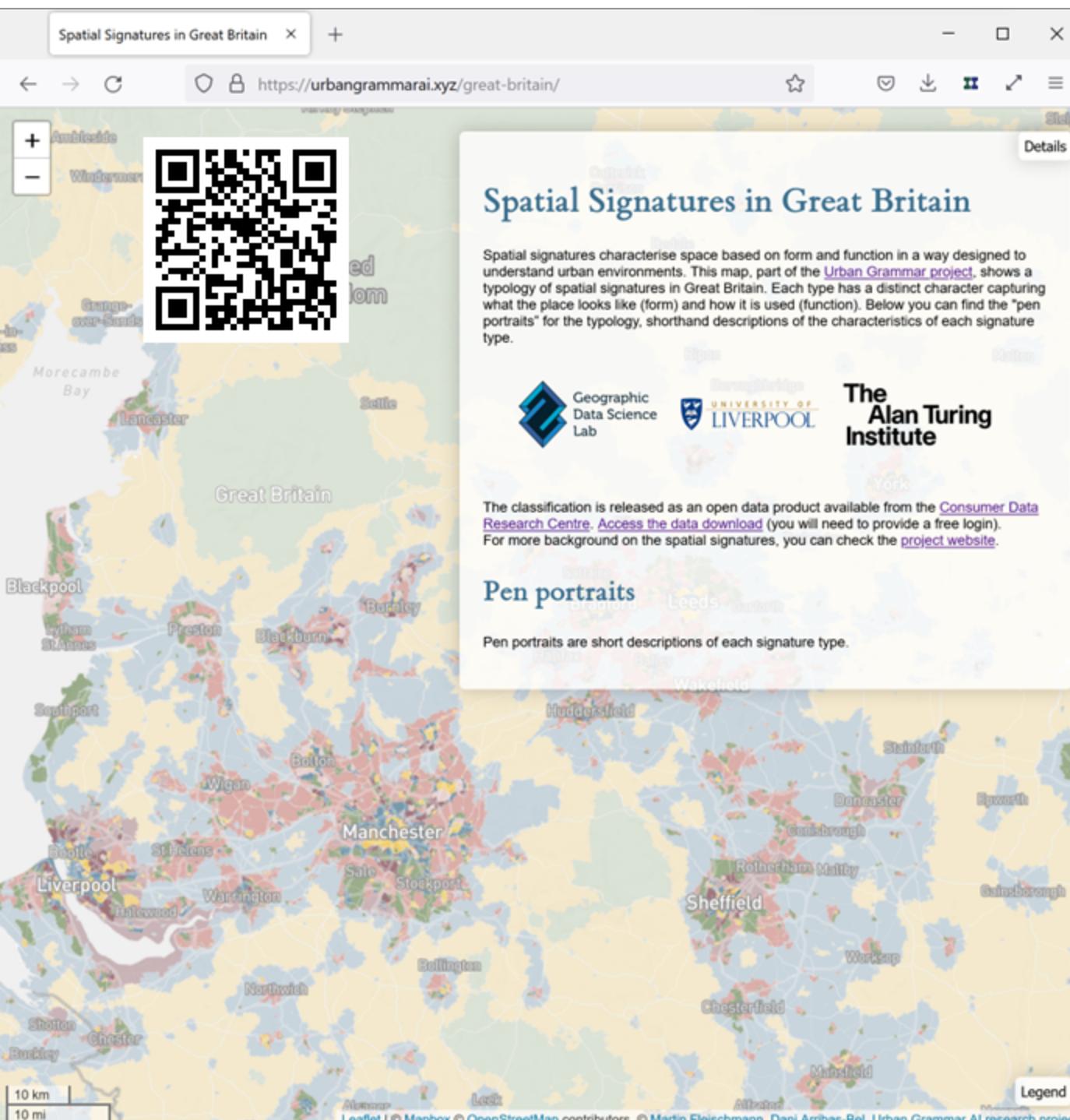
Learning an urban grammar from satellite data through AI

Learn more ↓

Project status
Ongoing

Related programmes
[Urban analytics](#)





Habitat International

Volume 128, October 2022, 102641



Spatial Signatures - Understanding (urban) spaces through form and function

Daniel Arribas-Bel^{a,b} , Martin Fleischmann^b

Show more ▾

+ Add to Mendeley Share Cite

<https://doi.org/10.1016/j.habitint.2022.102641>

Get rights and content ▾

open access

scientific data

Explore content ▾ About the journal ▾ Publish with us ▾

[nature](#) > [scientific data](#) > [data descriptors](#) > article

Data Descriptor | Open Access | Published: 07 September 2022

Geographical characterisation of British urban form and function using the spatial signatures framework

Martin Fleischmann & Daniel Arribas-Bel

[Scientific Data](#) 9, Article number: 546 (2022) | [Cite this article](#)

1779 Accesses | 2 Citations | 20 Altmetric | [Metrics](#)

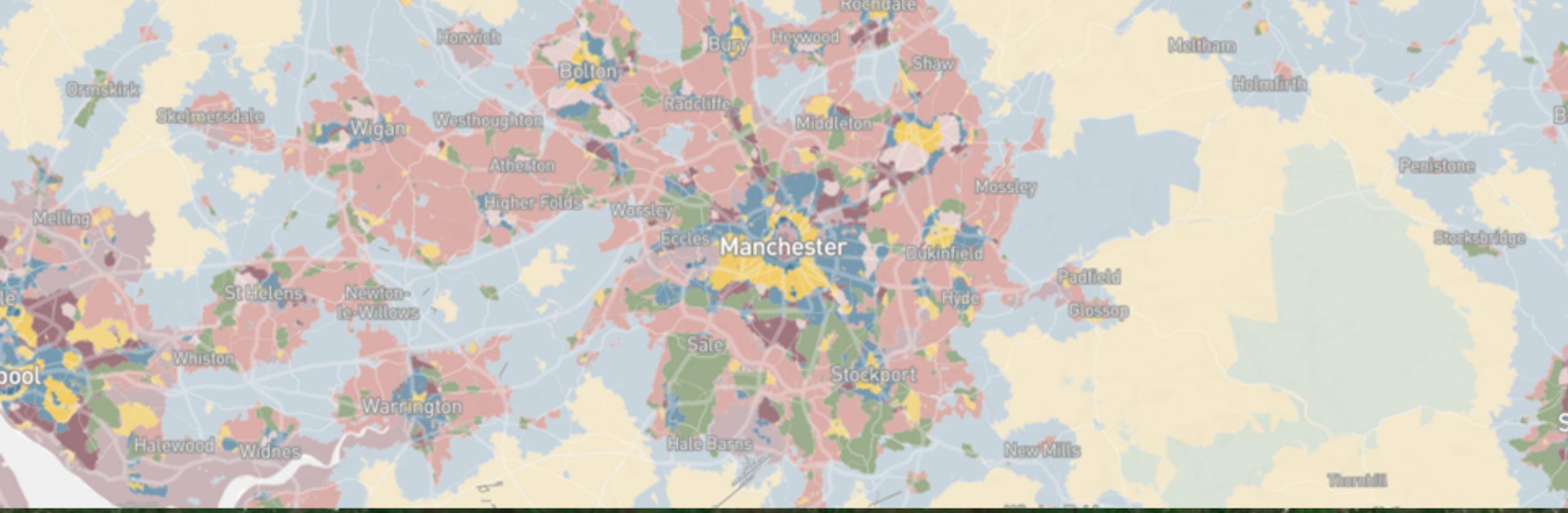
Experiments setup

What

Explore the extent to which neural networks can recognise spatial signatures from satellite imagery

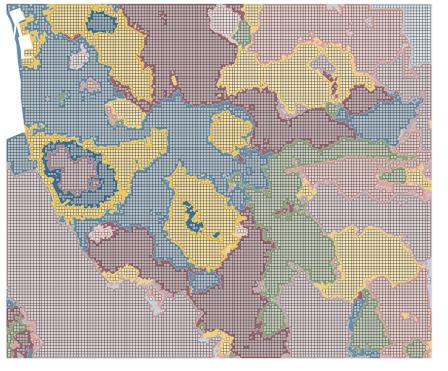
Why

- Learn about Spatial Signatures (scale, context)
- Explore the potential of NNs for cities
- Work towards more frequent Spatial Signatures

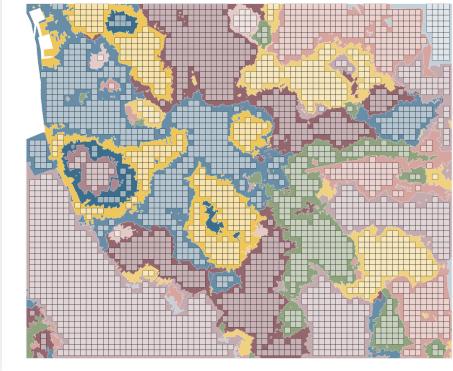


Dimensions to explore

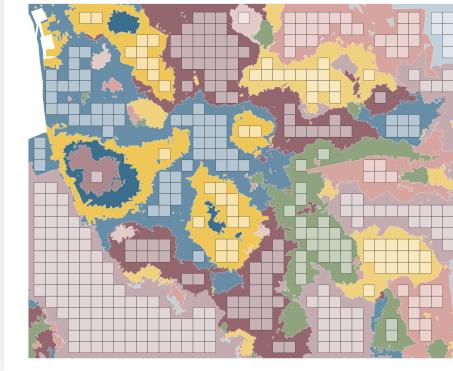
Chip size



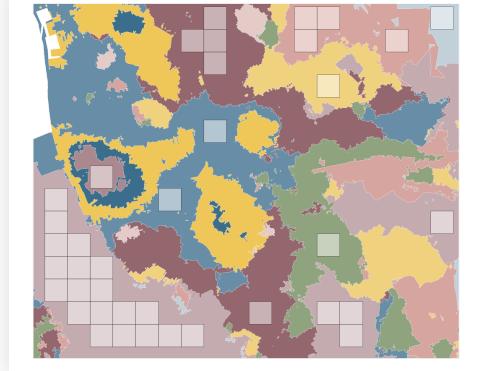
[74%]



[57%]

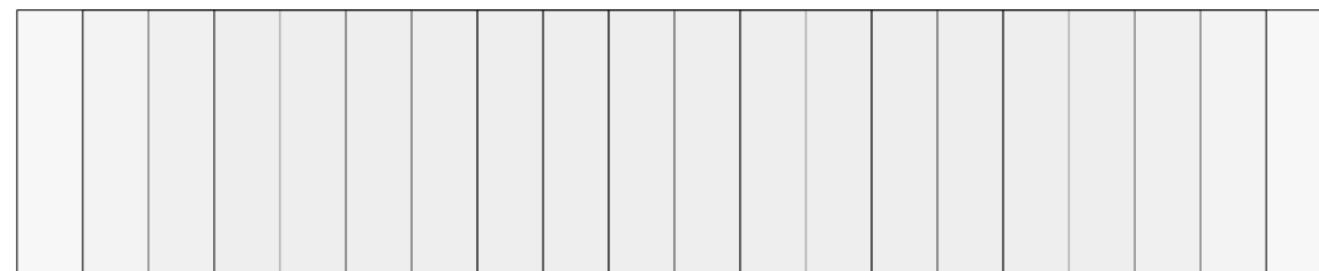
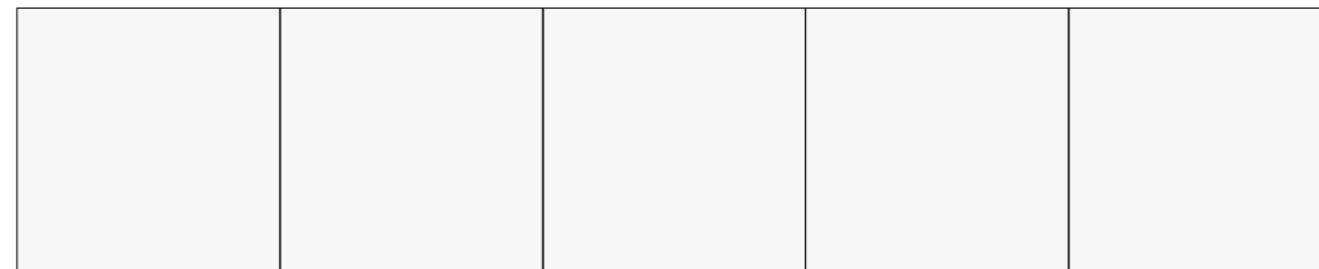


[35%]



[13%]

(Spatial) data augmentation



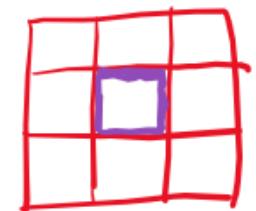
Model architecture

EfficientNetB4

- Image Classification
- Multi-Output Regression

$$S_i = f \left(\sum_k P_k + \sum_k w_k P_k \right)$$

- f {
1. Argmax
2. (MN) Logit⁺
3. Random Forest
4. Grad. Boosted Trees



$$wP_{ii} = \sum_j w_{ij} P_{ik}$$

Evaluation

Metrics

- Standard
 κ , accuracy, F1
- Spatial
joint counts

Summarisation

$$\begin{aligned} \text{Perf}_i = & \alpha + \sum_m \delta_m M_i + \\ & \sum_a \gamma_a A_i + \beta_1 \text{Chip Size}_i \\ & + \beta_2 W_i + \epsilon_i \end{aligned}$$

Results

	κ	Global Accuracy	Macro F1 w.	Macro F1 avg.
Intercept	0.2185*** (0.0209)	0.3236*** (0.0175)	0.2790*** (0.0174)	0.1798*** (0.0375)
(M) Logit E.	-0.0245 (0.0168)	-0.0256* (0.0141)	-0.0324** (0.0141)	-0.0325 (0.0302)
(M) Max. Prob.	-0.0559** (0.0222)	-0.0606*** (0.0187)	-0.0421** (0.0186)	-0.0296 (0.0399)
(A) M.O.R.	0.0227 (0.0184)	-0.0357** (0.0155)	-0.0278* (0.0154)	0.1787*** (0.0331)
(A) S.I.C.	0.0232 (0.0184)	-0.0247 (0.0155)	-0.0171 (0.0154)	0.1101*** (0.0331)
Chip Size	0.0036*** (0.0004)	0.0043*** (0.0003)	0.0048*** (0.0003)	0.0014** (0.0006)
W	0.0572*** (0.0168)	0.0468*** (0.0141)	0.0531*** (0.0141)	0.0392 (0.0302)
R^2	0.7214	0.8281	0.8514	0.4191
R^2 Adj.	0.6899	0.8086	0.8346	0.3533
N.	60	60	60	60

Table 2: Regression outputs explaining global non-spatial performance scores. Explanatory variables with a preceding (M) and (A) correspond to binary variables for the type of model (with histogram-based boosted classifier, or HGBC, as the baseline) and architecture (with baseline image classification, or BIC, as the baseline), respectively. Standard errors in parenthesis. Coefficients significant at the 1%, 5%, 10% level are noted with ***, **, and *, respectively.

	Within-Class Accuracy		
Intercept	0.1866*** (0.0308)	-0.0237 (0.0311)	0.0595** (0.0303)
(M) Logit E.	-0.0125 (0.0159)	-0.0125 (0.0141)	-0.0125 (0.0146)
(M) Max. Prob.	-0.0188 (0.0211)	-0.0188 (0.0186)	-0.0188 (0.0193)
(A) M.O.R.	0.1753*** (0.0175)	0.2512*** (0.0163)	0.1753*** (0.0160)
(A) S.I.C.	0.1202*** (0.0175)	-0.0783*** (0.0209)	0.1202*** (0.0160)
Chip Size	0.0014*** (0.0003)	0.0041*** (0.0003)	0.0014*** (0.0003)
1k Obs.		0.0514*** (0.0036)	
% Obs.			0.0156*** (0.0013)
W	0.0365** (0.0159)	0.0365*** (0.0141)	0.0365** (0.0146)
(S)Urbanity	0.2358*** (0.0349)	0.2022*** (0.0309)	0.2574*** (0.0320)
(S)Dense urban neighbourhoods	-0.1420*** (0.0349)	-0.1075*** (0.0309)	-0.0998*** (0.0322)
(S)Dense residential neighbourhoods	-0.1414*** (0.0349)	-0.0836*** (0.0311)	-0.0983*** (0.0322)
(S)Connected residential neighbourhoods	-0.1306*** (0.0349)	-0.0726** (0.0311)	-0.0754** (0.0323)
(S)Gridded residential quarters	-0.0785** (0.0349)	-0.0127 (0.0312)	-0.0049 (0.0326)
(S)Disconnected suburbia	-0.0601* (0.0349)	-0.0103 (0.0311)	-0.0019 (0.0324)
(S)Open sprawl	-0.0845** (0.0349)	-0.0995*** (0.0309)	-0.1143*** (0.0321)
(S)Warehouse park land	-0.0857** (0.0349)	-0.0788** (0.0309)	-0.0817** (0.0320)
(S)Urban buffer	-0.0828** (0.0349)	-0.1382*** (0.0311)	-0.1753*** (0.0330)
(S)Countryside agriculture	0.2236*** (0.0349)	0.1593*** (0.0312)	0.1118*** (0.0334)
(S)Wild countryside	0.3876*** (0.0349)	0.3283*** (0.0311)	0.2925*** (0.0330)
R^2	0.4979	0.6087	0.5794
R^2 Adj.	0.4857	0.5987	0.5686
N.	720	720	720

Table 3: Regression outputs explaining within-class accuracy. Explanatory variables with a preceding (M), (A) and (S) correspond to binary variables for the type of model (with histogram-based boosted classifier, or HGBC, as the baseline), architecture (with baseline image classification, or BIC, as the baseline) and spatial signature (with Accessible suburbia as the baseline), respectively. Standard errors in parenthesis. Coefficients significant at the 1%, 5%, 10% level are noted with ***, **, and *, respectively.

	JC W_thr	$\log(JC)$ W_thr	JC W_union	$\log(JC)$ W_union
Intercept	4.3454*** (0.9507)	1.4617*** (0.1344)	4.7103*** (0.5763)	1.6311*** (0.1080)
(M) Logit E.	-0.1406 (0.4951)	-0.0431 (0.0700)	0.1851 (0.2995)	0.0481 (0.0561)
(M) Max. Prob.	0.1128 (0.6442)	-0.1223 (0.0911)	0.2819 (0.3887)	0.0223 (0.0728)
(A) M.O.R.	-3.1630*** (0.5494)	-0.5744*** (0.0777)	-2.7875*** (0.3301)	-0.4647*** (0.0619)
(A) S.I.C.	0.0119 (0.5532)	-0.2390*** (0.0782)	-0.6666** (0.3329)	-0.0481 (0.0624)
Chip Size	0.0297*** (0.0108)	-0.0005 (0.0015)	-0.0061 (0.0065)	-0.0080*** (0.0012)
W	-0.9325* (0.4945)	-0.1376** (0.0699)	-0.9556*** (0.2991)	-0.1785*** (0.0560)
(S)Urbanity	4.6650*** (1.0696)	0.6574*** (0.1512)	0.1156 (0.6460)	-0.1258 (0.1211)
(S)Dense urban neighbourhoods	1.7796* (1.0695)	0.5094*** (0.1512)	0.7480 (0.6487)	0.1609 (0.1216)
(S)Dense residential neighbourhoods	-0.8545 (1.0958)	0.0672 (0.1550)	-0.4636 (0.6647)	-0.0920 (0.1246)
(S)Connected residential neighbourhoods	-0.3656 (1.1018)	0.1543 (0.1558)	-0.4388 (0.6647)	-0.1447 (0.1246)
(S)Gridded residential quarters	-0.2000 (1.0744)	0.1009 (0.1519)	-0.6203 (0.6517)	-0.2111* (0.1221)
(S)Disconnected suburbia	-0.9752 (1.1213)	-0.1719 (0.1586)	-1.0303 (0.6684)	-0.3358*** (0.1252)
(S)Open sprawl	1.8342* (1.0604)	0.1734 (0.1499)	2.1575*** (0.6432)	0.3576*** (0.1205)
(S)Warehouse park land	0.5496 (1.0694)	0.2123 (0.1512)	1.2245* (0.6487)	0.3054** (0.1216)
(S)Urban buffer	-0.0558 (1.0521)	-0.0931 (0.1488)	2.7027*** (0.6382)	0.5164*** (0.1196)
(S)Countryside agriculture	-1.3759 (1.0521)	-0.2511* (0.1488)	0.6623 (0.6382)	0.0670 (0.1196)
(S)Wild countryside	-2.0183* (1.0521)	-0.5065*** (0.1488)	-0.5918 (0.6382)	-0.1635 (0.1196)
R^2	0.1589	0.1954	0.2118	0.2660
R^2 Adj.	0.1368	0.1743	0.1913	0.2468
N.	665	665	670	670

Table 4: Regression outputs explaining (log of) differences in the spatial pattern between observed and predicted values, as measured by the Join Counts statistic. The Join Counts for each signature were computed using two types of spatial weights: one based on a distance threshold of 1Km (W_thr), and another one built as a the union of nearest neighbor and queen contiguity matrices (W_union). Explanatory variables with a preceding (M), (A) and (S) correspond to binary variables for the type of model (with histogram-based boosted classifier, or HGBC, as the baseline), architecture (with baseline image classification, or BIC, as the baseline) and spatial signature (with Accessible suburbia as the baseline), respectively. Standard errors in parenthesis. Coefficients significant at the 1%, 5%, 10% level are noted with ***, **, and *, respectively.

Conclusions

- Space matters for the spatial signatures
- There's value in combining NNs & other ML
- A *bit* closer to frequent Spatial Signatures

Urban Grammar

“Learning from Deep Learning”

Lessons from using computer vision to identify (urban) form and function in open data satellite imagery

#AAG2023

Dani Arribas-Bel

@darribas

Martin Fleischmann

@martinfleis



The
Alan Turing
Institute



Geographic
Data Science
Lab