RESEARCH

Supplementary Material
Inside 50,000 living rooms: An assessment of global residential ornamentation using transfer learning

Xi Liu[1], Clio Andris[1*], Zixuan Huang[2] and Sohrab Rahimi[3]

*Correspondence: clio@psu.edu
[1]Department of Geography, The Pennsylvania State University, 302 Walker Building, 16802 University Park, PA, USA
Full list of author information is available at the end of the article

Counting decorative elements

When decorating their homes, residents display a range of few to many items. An example living room (Figure S1) was decorated with more than 10 pieces of wall art. When aggregating the decoration counts at the census tract level (for example), this particular listing would increase the average number of decorations of the whole tract by a large number of elements. Thus, when counting elements (i.e. plants, wall art, and decor), if a listing contained more than five of any these individual elements, we recorded it as five for the analysis. We capped the number at five because there are very small portion of listings (both globally and locally, Table S1, i.e. fewer than 5%, that had more than five occurrences of each decoration elements.

Table S1 Percentage of listings with more than five detected decoration elements

	Plants	Wall art	Decor
Global	0.006	0.035	0.001
Chicago, IL	0.008	0.054	0.001
Houston, TX	0.001	0.029	0.001
Los Angeles, CA	0.003	0.032	0.002
New York, NY	0.007	0.044	0.001
Philadelphia, PA	0.007	0.040	0.001
Washington, DC	0.002	0.046	0.002

Color analysis

We posit that if any color in the top 10 detected colors of an image are "vibrant", then we consider the living room to be "vibrant" as well. It is hard to know if using the top "10" colors was the most viable value, as the 10th most frequent color in an image may be very small. Yet, even with this generous criterion of 'vibrancy', only 16% of living rooms have vibrant colors. Our results provide a wide range of differences between cities: 5% of living rooms in Haiti's Port-au-Prince are considered vibrant, while 43% of living rooms in Marrakech, Morocco are considered vibrant.

What is the range of participation between the first and tenth most frequent colors in a single image? The following box plots show that, at minimum, the most frequent color (1st color) comprises at minimum 12% of the image (in this particular case, the 10th image comprises 7% of the image) with a median value of 34%. The

10th most frequent color comprises at most 16% of an image, and at minimum, nearly 0.00001% (Figure S2).

World cities level analysis: Moran's I parameterization

To test whether there is spatial clustering in ornamentation behavior among global cities, we used both Moran's I and ANOVA. The ANOVA test divides cities into pre-defined categorical regions, while Moran's I requires certain subjective decisions to define a proper kernel bandwidth (neighborhood size) for defining the spatial adjacency matrix. Since our cities are distributed unevenly around the world, the choice of how to geographically define a focal city's 'nearest neighbors' may produce different results. For example, it is compromising to select the same number of neighbors for the network of dense European cities vs. isolated Fiji–as Fiji's nearest neighbor is over 1,000 kilometers away, and at the same range, Paris has 10 nearest neighbors.

Thus, we tested two ways of defining spatial adjacency: fixed radius and K nearest neighbors (KNN). For both fixed radius and KNN methods, we used inverse distance weighting (IDW) to weight each city's set of adjacent cities. Using IDW, the influence of increasingly distant cities diminishes in step with their distance. In the paper, we presented the results using fixed radius of 4,000 km, a distance that guaranteed that each city had neighbors. In comparison, the results using KNN at $k = 15$ are shared here (Table S2). The KNN results resemble the fixed radius results, although decor and color prevalence have relatively lower Moran's I values here. In conclusion, like many modifiable areal unit problem (MAUP) issues in GIS modeling, our results are sensitive to our spatial definitions but yielded similar results when we varied the search radius.

Table S2 Patterning of interior elements across global cities (Moran's I) using KNN ($k = 15$).

Element	Moran's I	Moran's I Z-score	Moran's I P-value
Plants	0.263	5.892	<0.001
Books	0.303	6.754	<0.001
WallArt	0.293	6.542	<0.001
Decor	0.134	3.119	<0.001
Color	0.104	2.485	<0.01

Neighborhood level analysis results

For intra-city level analysis, we use linear regression (LR) and geographically-weighted regression (GWR) to test if the extent of residential ornamentation is correlated with socioeconomic properties of residents. As introduced in the paper, the socioeconomic proprieties we use include median household income, unemployment rate, percent of residents with bachelor's degrees, median house value, and racial diversity. These are the independent variables in our models. For LR, the dependent variables are aggregated to census tract; for GWR, the dependent variables are assigned to each listing.

Linear regression

The results of linear regression (Table S3) show that the socio-economic factors are not able to explain city-level variation of the extent of residential ornamentation. The R^2 for LR models in the six cities are all very low. While in some cases, the correlation coefficient is significant, it is not consistent across cities.

Table S3 Linear regression results for neighborhood-level analysis in the six U.S. cities

City	Element	Median Income (log)	Unemployment Rate	Percent Bachelor Degree	Median House Value (log)	Racial Entropy	Adjusted R^2
Chicago	Plant	−0.055 (0.134)	−0.006 (0.008)	−0.002 (0.003)	0.171 (0.157)	0.197 (0.130)	0.001
Chicago	Book	0.012 (0.052)	−0.003 (0.003)	−0.002· (0.001)	0.078 (0.061)	0.033 (0.746)	0.005
Chicago	Wall Art	−0.278 (0.168)	−0.033** (0.010)	−0.0002 (0.003)	0.159 (0.196)	−0.083 (0.163)	0.041
Chicago	Decor	−0.191* (0.084)	−0.008 (0.005)	0.002 (0.002)	−0.100 (0.098)	−0.061 (0.081)	0.025
Chicago	Color	−0.018 (0.045)	0.004 (0.003)	0.001 (0.001)	0.003 (0.053)	−0.082· (0.044)	0.021
Houston	Plant	0.245 (0.152)	0.007 (0.014)	−0.002 (0.003)	−0.103 (0.123)	0.197 (0.188)	0.003
Houston	Book	0.082 (0.071)	−0.003 (0.007)	−0.001 (0.001)	0.070 (0.058)	−0.075 (0.088)	0.062
Houston	Wall Art	0.205 (0.196)	0.004 (0.019)	−0.009* (0.004)	0.331* (0.160)	−0.383 (0.244)	0.066
Houston	Decor	−0.103 (0.112)	0.0002 (0.011)	−0.004 (0.002)	0.167· (0.091)	−0.074 (0.139)	0.009
Houston	Color	0.012 (0.057)	0.005 (0.005)	0.001 (0.001)	0.009 (0.046)	−0.107 (0.071)	−0.006
Los Angeles	Plant	0.070 (0.065)	0.004 (0.006)	−0.001 (0.002)	0.067 (0.066)	−0.181* (0.087)	0.013
Los Angeles	Book	−0.058* (0.026)	0.003 (0.003)	0.002*** (0.001)	0.067** (0.026)	−0.006 (0.034)	0.070
Los Angeles	Wall Art	0.071 (0.087)	−0.0001 (0.009)	0.003 (0.002)	0.044 (0.089)	−0.193* (0.117)	0.020
Los Angeles	Decor	0.118** (0.040)	0.003 (0.004)	−0.002 (0.001)	−0.026 (0.041)	−0.062 (0.054)	0.010
Los Angeles	Color	−0.004 (0.022)	0.001 (0.002)	−0.0002 (0.001)	0.014 (0.022)	0.029 (0.029)	−0.006
New York City	Plant	0.117 (0.076)	0.005 (0.005)	−0.004* (0.002)	0.143*** (0.042)	0.247** (0.076)	0.029
New York City	Book	−0.023 (0.031)	−0.0001 (0.002)	0.001· (0.001)	0.046** (0.017)	0.057· (0.031)	0.043
New York City	Wall Art	0.012 (0.103)	0.006 (0.007)	0.002 (0.002)	0.164** (0.058)	0.164 (0.103)	0.026
New York City	Decor	−0.032 (0.043)	−0.009** (0.003)	−0.001 (0.001)	−0.017 (0.024)	−0.001 (0.043)	0.014
New York City	Color	−0.035 (0.024)	−0.002 (0.002)	0.0004 (0.0005)	0.032* (0.013)	0.002 (0.024)	0.007
Philadelphia	Plant	0.352** (0.127)	0.013 (0.011)	−0.003 (0.004)	−0.031 (0.170)	0.262· (0.154)	0.039
Philadelphia	Book	−0.021 (0.051)	−0.003 (0.005)	−0.003 (0.002)	0.084 (0.069)	−0.100 (0.063)	0.003
Philadelphia	Wall Art	0.156 (0.143)	−0.016 (0.013)	−0.002 (0.005)	−0.143 (0.192)	0.038 (0.174)	−0.010
Philadelphia	Decor	0.095 (0.069)	−0.005 (0.006)	0.0005 (0.002)	−0.151 (0.093)	−0.010 (0.084)	0.009
Philadelphia	Color	0.092* (0.041)	0.002 (0.004)	−0.002 (0.001)	0.025 (0.054)	0.012 (0.049)	0.019
Washington DC	Plant	−0.056 (0.101)	−0.022** (0.008)	−0.003 (0.002)	0.057 (0.103)	−0.177· (0.102)	0.030
Washington DC	Book	0.056 (0.044)	−0.004 (0.003)	0.0003 (0.001)	−0.011 (0.045)	0.016 (0.045)	0.126
Washington DC	Wall Art	0.221 (0.156)	−0.004 (0.012)	−0.001 (0.003)	0.086 (0.159)	−0.008 (0.158)	0.054
Washington DC	Decor	0.079 (0.070)	−0.007 (0.005)	−0.002 (0.001)	−0.058 (0.071)	−0.100 (0.071)	0.009
Washington DC	Color	0.062 (0.041)	−0.006* (0.003)	−0.001· (0.001)	−0.010 (0.042)	−0.111** (0.042)	0.047

Geographically-weighted regression

The GWR results showed no significant improvement over LR for the prediction of the presence of any object within a specified geographic range. Using a using a bandwidth of 85 observations and 281.8 as the effective number of parameters, the adjusted R^2 values range from 17.2% to 7.8% (Table S4). Raw R^2 values are at maximum near 60% for plants and wall art (only), for a few observations. In comparison, we tested the price as a dependent variable using the same independent variables, and this regression reaches an unadjusted R^2 value of 87% in Manhattan and Brooklyn in New York City (although in its entirety, this GWR was also statistically insignificant). We show an example of what the highest R^2 values look like in geographic space using plants in New York and Los Angeles (Figure S3).

GWR Parameters

Regarding the residual sum of squares (RSS), color performs better than books; and décor performs better than plants and wall art. We cannot comparise the RSS across the different objects (Table S4) as the RSS is naturally higher for dependent variables with higher actual values. Specifically, plants, wall art and décor ranged from 0-5, meaning that the residuals would naturally be larger than variables than books and color, which ranged from 0-1. Furthermore the RSS for variables with the same range will be sensitive to the distribution of those variables (i.e. bimodal or uniformly distributed). Sigma is defined as the square root of the normalized RSS (RSS divided by the degrees of freedom), and scales with the RSS. Since the dependent variables (i.e. object types) are different, the Akaike Information Criterion (AIC) can also not be compared across the different objects, but is reported in case the experiment is re-conducted under different circumstances (i.e. divided into different regions).

Table S4 GWR Parameters for different objects

Object	Plant	Books	Wall Art	Decor	Color
Residual Sum of Squares	968	170	1972	412	137
Sigma	0.93	0.39	1.33	0.61	0.35
Akaike Information Criterion	3991	1559	4984	2797	1258
Adjusted R2	0.172	0.153	0.125	0.091	0.078

Author details

[1]Department of Geography, The Pennsylvania State University, 302 Walker Building, 16802 University Park, PA, USA. [2]Department of Geography, University of Utah, 84112 Salt Lake City, UT, USA. [3]Department of Architecture, The Pennsylvania State University, 16802 University Park, PA, USA.

References

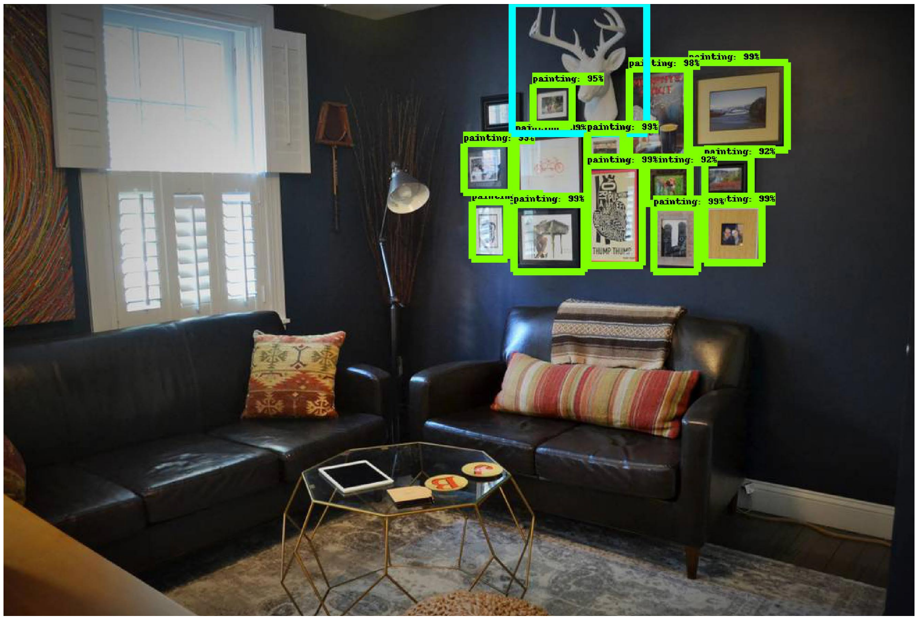

Figure S1 Example of outliers. Some listings may have an extreme number of decoration elements, which would influence the overall numbers when aggregated to higher levels (e.g. census tract).

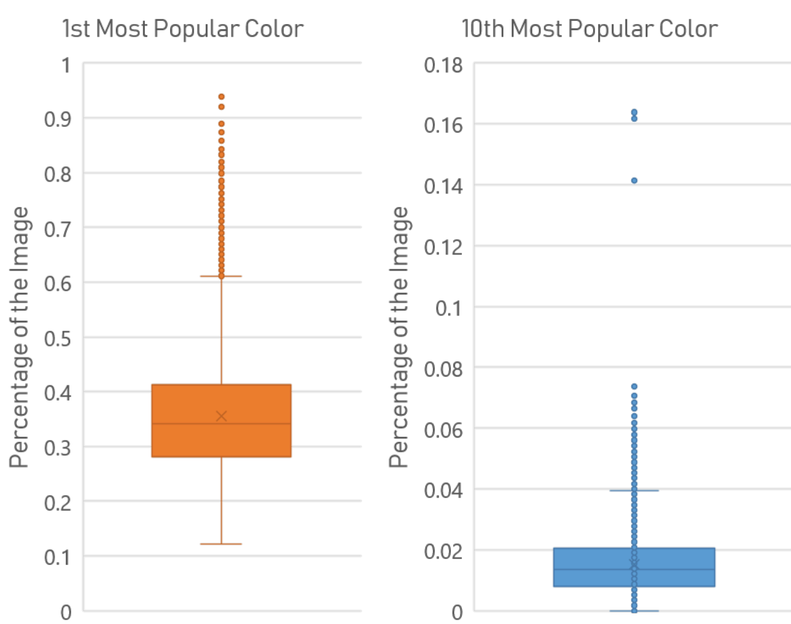

Figure S2 First and tenth most popular color frequencies. The most popular (frequent) color detected in an image can range from twelve percent to almost the entire image. The tenth most popular color detected comprises at most sixteen percent of the image, but is often quite small in composition.

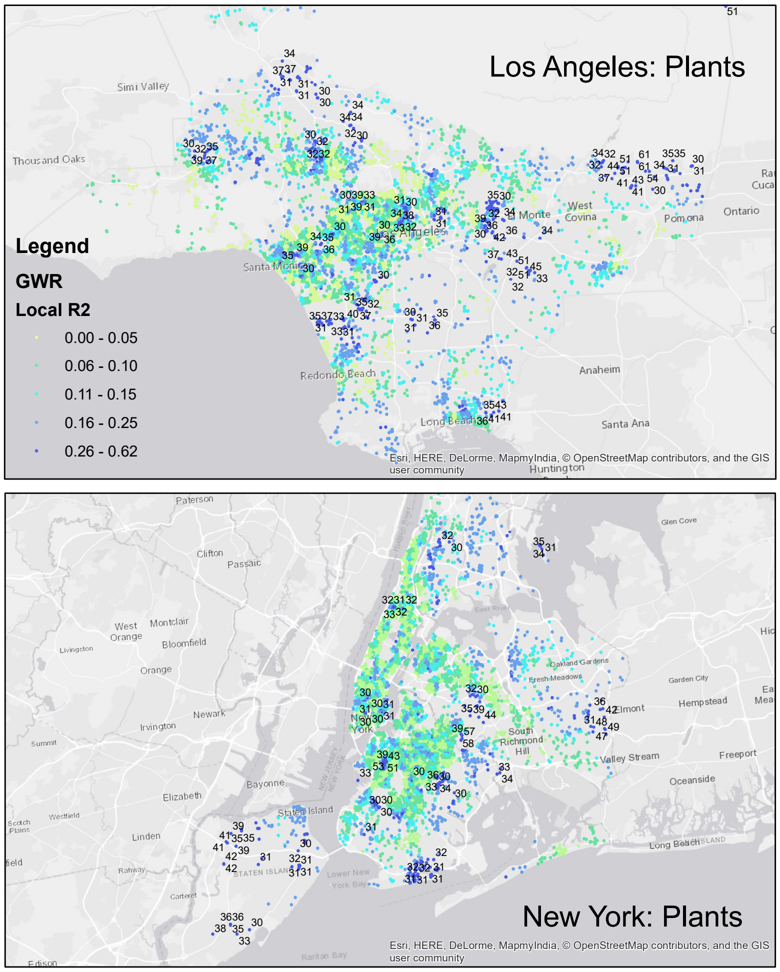

Figure S3 GWR examples of plants in Los Angeles and New York City. This image shows an example of GWR regression results for predicting the number of plants for each focal listing (with an 85-neighbor kernel bandwidth) in Los Angeles, California and New York City, where R^2 values reached 60%. R^2 values over 30% are labeled.