Estimating property value in Ireland with satellite imagery using Airbnb data as a value indicator - Not picture perfect

Tobias Florin Oberkofler*
t.f.oberkofler@umail.leidenuniv.nl
Leiden institute for advanced computer science - Leiden University
Leiden, Zuid Holland, The Netherlands

Maxime Casara
m.x.k.casara@umail.leidenuniv.nl
Leiden institute for advanced computer science - Leiden University
Leiden, Zuid Holland, The Netherlands

ABSTRACT

We introduce a novel way to enrich data driven value estimation using satellite imagery and land-use information to enable value predictions on properties with sparse data availability. We find that using the surrounding land-use labels generated by BigEarthNet slightly increases our models accuracy in predicting Airbnb properties price. This result hints at a potential correlation between spatial attributes and price values. However, we do not find improvements by using satellite imagery for the prediction task. This holds for different scales of satellite images. Also when using external information, such as location, in combination with the satellite images our proposed neural networks cannot significantly improve over techniques that just utilize external properties. We outline several potential reasons why the proposed methodology does not work as well as might be expected to inspire more robust future work.

KEYWORDS

Urban Computing; BigEarthNet; Satellite Imagery; Property value estimation; Deep Learning; CNN

1 INTRODUCTION

Property value prediction is an important task for real estate brokerage to make good buying decisions, as well as for local authorities to monitor and analyse ongoing trends in the pricing landscape within cities or entire countries. Housing prices are often times a good indicator of wealth and financial well-being in an urban setting. Insights on pricing can often not be drawn from official brokerage data as such data is slow in nature and might often not be able to capture the fast moving dynamics of the pricing structure. To capture those fast moving dynamics we will use current Airbnb data in this study as an indicator for the actual property values of a given location. Prices are dependent on a great variety of factors which are often not immediately obvious. In this work we investigate the possibility to use visual features extracted from satellite imagery to enhance housing price prediction models. We test this approach on data for the country of Ireland, and in particular the area of Dublin, which is freely and publicly available. We first outline a brief data analysis on the Airbnb pricing of the country of Ireland and introduce the multi-spectral Sentinel-2 satellite imagery dataset [4]. We then propose a novel algorithm to enrich data driven value estimation using satellite imagery and land-use information provided by BigEarth Net dataset. We detail an experimental setting to investigate the performance of our proposed models and compare them to baseline and state-of-the-art methods. Results are critically evaluated, analysed and summarized. Concluding potential future work is outlined.

2 BACKGROUND

The original BigEarthNet imagery and land use data set used in this paper was first proposed by Sumbul et al [15] and refined in their later work [16]. The underlying idea of their original work is using deep learning to facilitate an automated remote sensing (RS) application. They present a study on the effectiveness of different state-of-the-art machine learning algorithms, in particular deep convolutional neural networks to execute the RS task and devise a new labeling strategy for land use coverage to better adjust to the specific properties of the image modality of the multi-spectral Sentinel-2 satellite image dataset. Based on the recent advances in the field of remote sensing and the extensive effort to build high quality, extremely high-resolution satellite imagery and land use maps for Europe, we utilize this data to introduce a novel factor into automated property value estimation.

Airbnb[1] is an online (vacation) rental marketplace which offers housing in over 220 countries with listings in more than 100,000 cities. Several hundred thousands of this listings are located in Europe. A lot of research has been invested into understanding the underlying pricing structures of Airbnb based on several factors of meta information [11, 17, 20]. However, to the best of the knowledge of the authors, it has never been attempted to enrich such models by satellite imaging nor land-use label information. The models proposed in these previous researches are mainly focused on exploiting (latent) information found in text descriptions, either connected to the property and its current owner, or its spatial location such as proximity to (touristic) landmarks. Lee et al [9] uncover the importance of several social features associated with room prices in the Airbnb economy such as the owner's response time and cleanliness score.

Bency et al. [1] are the first to use satellite imagery along side classical factors to predict Housing prices in London, Birmingham and Liverpool. They train a convolutional neural network classifier to label each image square to be in the top-10,top-20,...,.bottom-10 percent of most expensive houses. Following they benchmark several regression algorithms that utilize the predicted label along side other external "house-level" factors to determine a price prediction. Law et al. [7] proposed a method which utilizes Google Street View image data along side external factors to predict the price of properties within London. Both approaches are significantly different from our proposed method as 1) we do not carry out a classification

*Both authors contributed equally to this research.

[1] http://airbnb.com

task on the satellite image but use it directly to predict the price and 2) we do not utilize Street View Images.

3 DATASET

For our work we use the BigEarthNet[2] as well as the Airbnb data from Inside Airbnb[3]. BigEarthNet was first proposed by a team from the Technical University of Berlin [15] and is considered one of the most extensive land-use and satellite imagery data set covering Europe. Inside Airbnb is an independent effort from a data journalist to make the listings on Airbnb largely available to public to enable valuable insights into usage and potential misusage of the platform.

3.1 Data insights

We use the information on Airbnb listings in Ireland sourced from Inside Airbnb to create some initial analysis of the pricing and location dynamic of Airbnb apartments within the country. Overall there are more than 27,000 listings all across the country. The average price there is at around 92€ and the 75% quantile at around 100€. To exclude outliers and create a more consistent and reliable sample, we excluded all values with a price outside the given quantile.

There are three explicit types of usage for Airbnb: "Private rooms", "Whole Apartments" and "Shared Rooms". "Private rooms" are usually intended for a single person or a pair, while "whole apartments" are mostly meant to be shared between more people making them more expensive in the overall price. As no indication on the number of indented guests is given in our dataset, which makes it impossible to account for the effect, these rooms need to be included. "Shared rooms" are shared with multiple other people making them cheaper and were therefore also excluded for our experiments. By excluding other types of rooms we hope to limit random fluctuation in the pricing of our properties that are not related to their location and quality of the surrounding environment, but rather with their difference in intended usage. After filtering out the data, 10,000 entries remain.

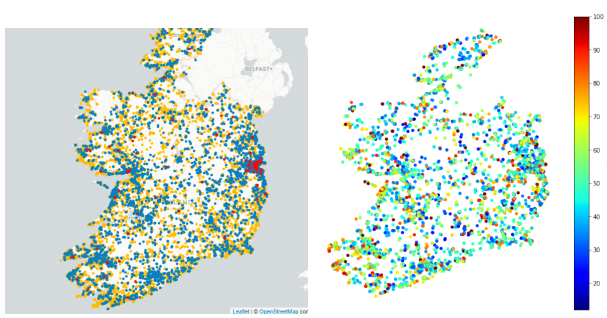

Figure 1: Left:Scatter plots of Airbnb listings in Ireland. Yellow dots indicate apartments, blue dots indicate houses and red dots indicate shared living opportunities. Right: Scatter plot of cleaned airbnb listings, including only properties within the 75% quantile and labeled as "Private rooms", in Ireland. The color intensity of a given dot indicates the price per night.

[2]http://bigearth.net
[3]http://insideairbnb.com/

The Sentinel-2 satellite images also cover Ireland with high resolution and squares of about 10 by 10 meters. Such high resolution is very rare for satellite images and allows for fine grained support to the property value estimation task. The satellite images cover all of the approximately 85000 km^2 land mass and boarding sea of the country.

Figure 2: Example of the Sentinel 2 image dataset for the city of dublin.

Using the meta-information of the Sentinel-2 data we can get the exact land-coverage of a given image in form of a Polygon. We can then iterate over all the Airbnb listings to find in which image tile a given property falls and compute the average price per tile. We will refer to the task of predicting the average price of a given tile as the average-grid price prediction task.

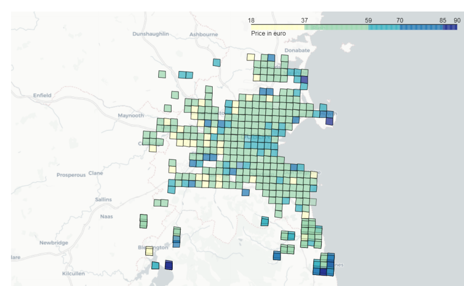

Figure 3: A choropleth map: The Image tiles as captured by the Sentinel-2 satellite, colored according to the average price of properties that fall within a certain image tile. Parts of the map with no tile do either not contain Airbnb properties or are not covered by the available satellite footage.

4 METHODOLOGY AND EXPERIMENTAL SETTINGS

This section present the methodology and experimental settings for both our approaches. The first one will take advantage of the land-use labels assigned to each patches of land by the Big EarthNet dataset. The second one will consider raw images as input for a convolutional neural network and learn its own representative features. In both approaches the spatial data will be coupled with Airbnb properties metadata, such as latitude, longitude and number of reviews to increase the price prediction accuracy.

4.1 Label based approach

4.1.1 Methodology. The Big EarthNet dataset contains land-use labels for several territories throughout the European continent. It is a valuable resource we hope to make use of. In the following figure, we show that the Airbnb dataset and the Big EarthNet dataset can be superposed using latitude and longitude:

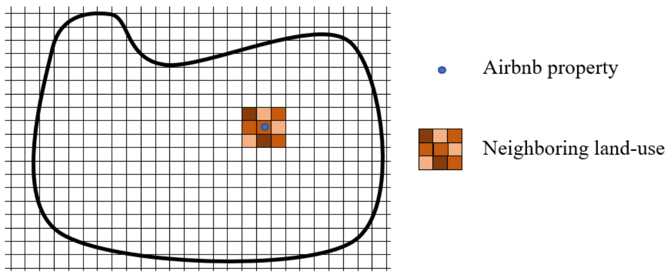

Figure 4: Enrichment of Airbnb dataset with land-use labels from Big EarthNet dataset, joined on geographical attributes

By merging both dataset on their geographical attributes, we can associate neighbouring spatial information to Airbnb properties based on their position. We are going to use these neighbouring features together with Airbnb metadata to train a property price predictor with supervised learning.

We limit the area of interest to Ireland. Unfortunately, there are no subset of BigEarthNet dataset available. To avoid working with the whole data, we use Earth Engine by Google which allows us to explore the area of interest only.

4.1.2 Experimental settings. We are here free to chose whichever neighbourhood model we want. In the following figure we present some examples of neighbourhood models:

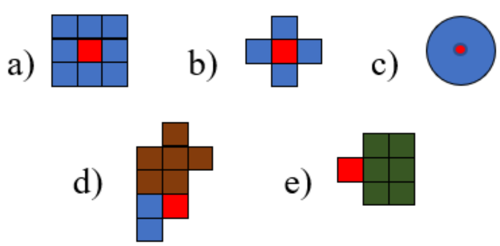

Figure 5: Neighbourhood models: a) Queen model, b) Rook model, c) Equidistant model, d) Contiguous identically labeled lands, e) Neighboring specific type of land (e.g., forest, water...)

Due to the lack of documentation for tile-by-tile exploration with Earth Engine, which made it impracticable to manually construct our proposed neighborhood models, we are going to test our models with the Queen model. Our approach is to draw and vectorize the most common labels surrounding a property and append them to the property metadata. The metadata is then fed to a predictor that is trained with supervised learning to predict the price of the property. We chose not to consider all properties in the dataset. For this task, we filtered them with the following features: cost for a night less than 100€, room type is "Private rooms" and the property lies within a delimited area around Dublin (100*200km polygon, outside of which we found out Earth Engine doesn't have access to land-use labels). Below is an example of Airbnb properties and their metadata:

latitude	longitude	neighbourhood	number_of_reviews	commonlabel
53.33072	-6.31305	Ballyfermot-Drimnagh LEA-5	14	[Discontinuous urban fabric, Green urban areas
53.33558	-6.37986	Ballyfermot-Drimnagh LEA-5	1	[Discontinuous urban fabric, Industrial or com
53.33953	-6.35568	Ballyfermot-Drimnagh LEA-5	1	[Discontinuous urban fabric, Industrial or com
53.33704	-6.38110	Ballyfermot-Drimnagh LEA-5	52	[Discontinuous urban fabric, Industrial or com
53.33970	-6.35577	Ballyfermot-Drimnagh LEA-5	3	[Discontinuous urban fabric, Industrial or com
...	...	...	...	
52.98128	-6.04111	Wicklow LEA-6	3	[Pastures, Sea and ocean, Non-irrigated arable
53.29584	-6.15439	Blackrock LEA-6	57	[Discontinuous urban fabric, Sea and ocean, Co
53.10829	-6.04060	Greystones LEA-6	37	[Sea and ocean, Non-irrigated arable land, Pas
53.36250	-6.17793	Clontarf LEA-6	10	[Discontinuous urban fabric, Intertidal flats,
53.15367	-6.06782	Greystones LEA-6	9	[Sea and ocean, Non-irrigated arable land, Pas

Figure 6: Airbnb properties and their metadata. In this example, each property is enriched with the surrounding 3 most common land-use labels

In total, 4678 properties satisfied these conditions. For our regression task, the 3 benchmark predictors we are going to use are a Decision tree model, a Random forest model and a Polynomial regression model. We design our training set and test set to represent respectively 90% and 10% of the dataset.

4.2 Image based approach

4.2.1 Methodology. Utilizing the high resolution satellite images of BigEarthNet we hope to be able to infer and exploit even more implicit characteristics of a given location, as deciding factors in the determination of the pricing structure. The underlying idea is that based on recent success of convolutional neural networks for remote sensing, it might be possible to infer characteristics that make a good neighborhood by itself without the need of expensive labeling by human workers. Law et al. [7] have demonstrated the possibility of using images to enrich property value estimation by using Google Street View data in combination with known neighborhood characteristics to create pricing models of London. However they find the neighborhood characteristics to be significantly more correlated with the pricing than any other attribute they could infer from their image based neural models. A clear distinction between our work and the work of Law et al. is that we strive to use actual satellite images instead of Street View data which might be able to capture a more comprehensive picture of the surrounding environment.

This approach is computationally far more complex and expensive than the previous one. Due to computational restrictions posed by the unavailability of the usual server infrastructure and the according resort to using the openly accessible but restricted cloud computing environment Google-Colab[4], we had to restructure our

[4]https://colab.research.google.com/

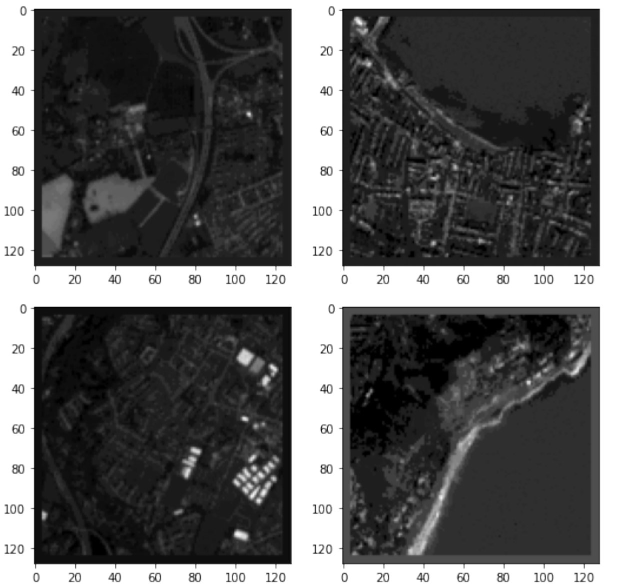

Figure 7: Example of the extracted and converted sentinel-2 images for the surrounding area of Dublin. This includes more Urban and industrial areas as well as parts of the countryside or the coastline etc.

data to fit on the provided 12gb GPU ram. Instead of utilizing all spectral information of the satellite images, we only selected the "B4","B3" and "B2" (red, green, blue) channel and merged this information into a single gray-scale image. We used an experimental approach to determine the weighing factors between the 3 main channels to create the gray-scale image(with r:0.29,g:0.58,b:0.11). A comparison of multiple more advanced approaches for gray-scale image creation can be found in the work of Cadik et al. [3].

Based on the network architecture outlined by Sumbul et al. [15], we construct a 3 layer convolutional neural network with the first channel using 1 channel input and 32 channel output, a kernel-size and stride of 2 and padding of 0. The second and the third layer use (32,32,2,2,0) and (32,64,2,2,0) respectively. Between each convolutional layer we apply a max pooling with a (2,2) and (4,4) kernel-size and stride respectively to effectively reduce the matrix dimensions along the original images axis. Furthermore do we apply a Batch-Norm operation [6]. We use the processed one-channel gray-scale image of 128×128 as input to the network. The output of the convolutional part then takes the shape of $o = 64 \times 2 \times 2$.

Following the conventional part, we use a fully connected neural network on the output o with 2 fully connected linear Layers of dimension (256,100) and (100,1) with a ReLU [5, 13] non-linearity between the layers. The output of the network is taken as the price estimation for a given square.

We use the mean-square-error (MSE) loss criterion given by

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{1}$$

and a stochastic gradient descent (SGD) [14] optimizer to compute the learning updates. Learning rate and momentum are determined by using a Bayesian optimization [2] for 100 runs (see section 4.2.2). The optimal found parameters turned out to be: lr=0.00047 and momentum=0.57. We use a batch size of 8 to balance information propagation and computational restrictions on graphical memory usage.

For our second approach, we merge this architecture together with an additional 2-layer fully connected neural network with a ReLU non-linearity to incorporate known characteristics such as location and average number of reviews for a given location in addition to the satellite image. The remaining part of the network is identical to the former described architecture. The optimal learning rate and momentum was also determined by using Bayesian optimization: lr=0.0001 and momentum=0.89

Our third approach uses a concatenation of neighbouring images instead of a single satellite image as network input. To stay within the computational restrictions, a scaling needed to be applied to the images. A concatenation of 9, 128×128 images would indeed have led to an exploding computational and memory requirement. To facilitate this, we stored a second set of the images scaled down by a factor of ten to 12×12 pixel. Again we use a color to gray-scale mapping to obtain 1 channel image.

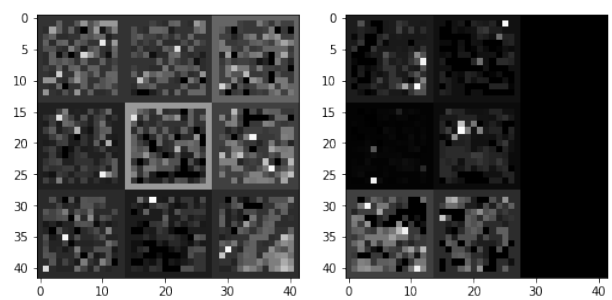

Figure 8: Scaled down and merged neighborhood images derived from the sentinel-2 satellite images. Each 12×12 pixel section is the down-scaled version of a given satellite image. Some parts of our included area are not covered by the satellite images. This parts are left black in the merged image.

This merged neighborhood images are then fed to the network as input. We use a similar architecture than the two cases before, with an adaption from three to two convolutional layers and a change of the convolutional kernel and stride sizes to account for the changed input image dimensions. The convolutional layers in this case are described by (input channels: 1, output channels: 32, kernel: [2,2], stride: [2,2], padding: [0,0]) and (32,32,[2,2],[2,2],[0,0]) for the first and second layer respectively. Again we use a (2,2) Max-Pooling and 2d-BatchNormalization operation after each convolutional layer.

The output of the convolutional part is then merged with a fully connected network to compute a final network prediction.

We use a stochastic gradient descent optimizer and the mean-squared-error loss criterion for computation of the step wise updates. Overall we run the training procedure for 500 epochs. A batch size of 8 is used to split the data into computational feasible mini-batches and avoid exceeding (graphical-)RAM usage. The optimal parameters for learning rate and momentum are set to (0.0009,0.71).

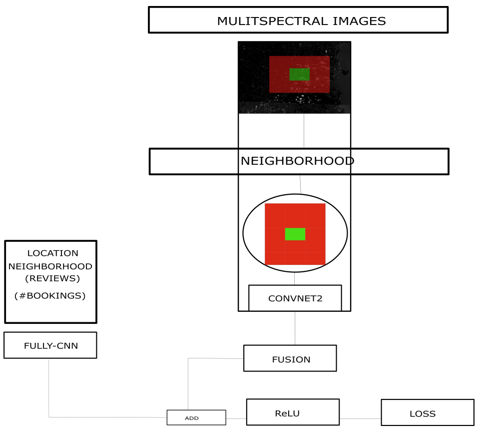

Figure 9: Network model using raw images as input and Airbnb metadata to predict property prices. The convolutional part of the architecture takes orientation by the model proposed by Sumbul et al. [15] while the overall network is inspired by the work of Zhang et al. [19] who proposed a network to combine the outcomes of a series of convolutional networks with external attributes in a fully-connected way.

4.2.2 Hyperparameter optimization. We use sequential model based optimization (SMBO) as outlined in the work of Mockus et al.[12] and for our implementation Bergstra et al. [2][5]. Utilizing the expected improvement for a given parameter setting x on our evaluation function f compared to the best found setting $\mathbf{x}^{+}$, as driving exploration force:

$$\mathrm{EI}(\mathbf{x}) = \mathbb{E} \max \left(f(\mathbf{x}) - f\left(\mathbf{x}^{+}\right), 0\right) \tag{2}$$

We set the maximum number of evaluations to 100 and pose no further restriction on the optimization process. The explored parameter space for the stochastic gradient descent optimizer in all cases is $lr : [0.000001, 0.001]$ and $momentum : [0.1, 0.9]$. After having initially experimented with the Bayesian optimization method of Liaw et al. [10] we settled for the method of Bergstra et al. due to the better fit with the available computational architecture and simplicity in the implementation.

4.2.3 Experimental setup. We use the available Polygon information that describes the exact location and dimension of a given satellite image to find the average price and number of ratings of private Airbnb properties within a given polygon square. After having computed the average price per square and excluded squares for which no pricing information is available, as no properties fall within their borders, we benchmark several algorithms for the price prediction task. The task of a given algorithm is to estimate the average price pr of a square w_i as closely as possible.

$$\hat{pr} = f(w_i) \tag{3}$$

Each square w_i has an associated satellite image s_i, external properties p such as location and average number of reviews and a neighborhood n_i. The different classifiers used this associated metadata to make their price prediction. For the neural models, we either feed just the associated satellite image of a given square, the image and the available external properties or the image and the next neighboring images as input and obtain a single real number as output. we compare the results of these neural models to a Random Forest regressor, a second order Polynomial regression predictor and the average prediction. These models take as input the available external information (location and average number of reviews) and compute a single real number as output.

For the neural models we split the available data into a train, development and a test set, taking 70%,20% and 10% of the data accordingly. For the classical models we apply 5-fold cross validation to validate our results.

5 RESULTS

5.1 Label based results

The baseline results were estimated with 3 different benchmark models: a Decision Tree, a polynomial regressor and a random forest. For each model we used the following meta data: neighbourhood (as defined by Airbnb, usually the name of the county), latitude and longitude of the property and its number of reviews. We return the best RMSE out of 500 runs for each models, shuffling at random the dataset at every run. The results are shown in the following table:

Model	RMSE	R2-score
Decision tree	22.63	-0.56
Constant model (Price = Avg-price)	17.87	-0.01
Polynomial regression	17.74	0.043
Random forest	**16.36**	**0.17**

Table 1: Baseline results

The better model is the Random forest with a small difference of 1.38€ on prediction accuracy compared to the Polynomial regressor. Compared to the constant model, it is 1.51€ more accurate. For each property, we are next going to keep the same metadata and append the new metadata which describe its 3 most common neighboring land-use labels in a square of side 1.5km. The results for each estimators are shown in the table below:

Model	RMSE	R2-score	Baseline Comparison
Decision tree	19.48	-0.22	▲
Polynomial regression	15.70	0.15	▲
Random forest	16.83	0.137	▼

Table 2: Prediction with new metadata: 3 most common land-use labels. Only the polynomial regressor took advantage of the new metadata to increase its accuracy by 11.5%

We notice a slight increase in accuracy for our Polynomial regression model. It now beats the constant model by 2.17€, compared to

[5]https://github.com/hyperopt/hyperopt

0.13€ when not using the new spatial metadata. However, the Random forest model lost 0.67€ in accuracy. We can change our spatial data collection settings to see how they influence our polynomial model:

Model	r(km)	n° of top labels	RMSE	R2-score
Polynomial regr.	2.5	1	16.09	0.09
		3	15.70	0.15
		5	15.97	0.14

Table 3: Different parameter settings: r is the side length of a square centered on the property and the top labels are the most common land-use labels (in number) found in the square. There is no significant change in accuracy when choosing different neighborhood parameters

Changing our metada creation approaches did not improve the accuracy for the polynomial model.

5.2 Image-based results

We find that overall when predicting the average price for a fixed grid structure the results are, unsurprisingly, better than when trying to predict individual properties. Comparing the outlined neural architectures, each with the best found hyperparameter setting for a given model, we see that the CNN which utilizes the neighbourhood concept and the network which utilizes images along side with external information slightly outperform the architecture which always relies on a single input image. However none of the models yields the significant improvement over the constant average-price prediction model. All proposed methods are clearly outmatched by the Random Forest Regressor model which is able to achieve a root-mean-squared-error(RMSE) of 6.18 alongside an R2-score of 0.78.

Model	RMSE	MAE	R2
CNN-Singel_Square	13.14	10.15	-0.02
CNN-Multiple_Squares	11.50	9.16	0.13
CNN-Image+External	11.88	9.01	0.001
Random Forest	**6.18**	**3.63**	**0.78**
Polynomial Regression(d=2)	11.66	9.32	0.32
Constant Model(average price)	12.55	9.67	0

Table 4: Performance of various ML algorithms for the average-grid price prediction task.

6 DISCUSSION

We found that hyperparamter optimisation had little effect to improve over the default for all proposed network architectures apart from the merged image-external properties network where the learning rate turned out to make a significant difference, however the found optimal was still close to the common default so that optimization did not yield drastic improvements. We hypothesis that the outlined 3 layer convolutional neural network is simply not expressive enough to capture the required nuances of the value estimation task to yield good results. Hence with or without parameter tuning the model might simple not be able to capture the required characteristics of the task.

In such case it is also possible that the selected satellite images are not representative enough of the underlying property pricing mechanism of a given location. This is suggested by the fact that including the price predictions of the convolutional networks often times made the prediction worse than when just relying on basic attributes such as location and the average number of obtained reviews. Furthermore is or reduction procedure for the down-scaling of the satellite images for the neighborhood based model not ideal as it is almost impossible to identify urban or rural structures in the down-sized images. Even though we oriented on several popular computer vision datasets such as MNIST[8] and Fashion-MNIST[18] when choosing the image dimensions, it might be fruitful further research approach to run the network on a larger scale neighborhood image if according computational infrastructure is available.

Furthermore is the rather small number of available data samples which provide data for Airbnb pricing as well as being covered by the Sentinel-2 image data rather small. In total there are about 1000 tiles which fulfill this criteria. This is arguable to small a dataset to train expressive deep convolution neural networks on. By using a more shallow network with only 3 convolutional layers we tried to balance the limited data availability with the possible expressiveness of the network. However we do expect significant improvements in stability and robustness of the networks when training on a larger set of images. Whether or not also increases in performance will be seen in such a scenario outlines an interesting path for future research.

7 CONCLUSIONS AND FUTURE WORK

The aim of our project was to take advantage of the Sentinel-2 and BigEarth Net spatial database to predict Airbnb price properties. Our hypothesis was that the price would be correlated with the surrounding land. For our label and image based approaches, while our best models increased slightly its accuracy by taking into account this spatial relation, the improvement is not sufficient in effect size to give evidence underlining the hypothesis.

Predicting such a volatile target value as housing prices by inferring from spatial information is a challenging task. It indeed feels that housing pricing remains just as much an art as a science. We were fortunate to have access to open geographical data to conduct our project. And it is fair to believe that the very reason the data is open source is because it is incredibly difficult to make insightful - and profitable - findings from it. More highly resolved satellite images for network training might furthermore contribute a lot to better predictive performance. We strongly advise the focus of future research to be involved in the question of including more "house-level" attributes such as number of bedrooms, window orientations, etc. as well as surrounding landmark features, proximity to schools, shops, job opportunities... We found that our proposed methods were not expressive enough to infer this information themselves from the available satellite images, hence future work should investigate priors that might ease the networks extraction of such features. Such a process could first train a network on landmark detection in the Sentinel-2 images and then proceed by cutting off the final classification and substituting it with a fully connected layer which could facilitate the price prediction task. Similarly one might want to solve an easier task such as classifying a given image tile to correspond to the top-50 or bottom-50 percent of prices and then just use this information as enhancement to already existing predictors instead of trying to find a complete solution in the latent space of the satellite images.

One might want to expand the data collection from Ireland to utilize all Airbnb locations, using a commercial service like Google Earth to obtain satellite imagery of those locations. Such an expansion could be easily facilitated with our existing processing and prediction routines. However, it is expected to introduce a significant computational overhead. Another attempt might be to use official pricing data issued by local authorities instead of the Airbnb data to create a more robust basis for the underlying pricing structure which might benefit all prediction algorithms.

8 CODE AND DATA AVAILABILITY

In the spirit of open and reproducible science we invite the reader to take a look at our three main code repositories: https://shorturl.at/pvBR0, https://shorturl.at/hCGKX and https://shorturl.at/ntIOT

Airbnb-Ireland data, as of the latest update in october 2020, is accessible via http://data.insideairbnb.com/ireland/2020-10-16/visualisations/listings.csv further datasets for multiple other cities can be found on http://insideairbnb.com/get-the-data.html.

The BigEarthNet data can either be downloaded completely, directly at the website of the provider http://bigearth.net/downloads/BigEarthNet-v1.0.tar.gz or for restricted areas via the Google Earth-Engine interface https://developers.google.com/earth-engine/datasets/catalog/TUBerlin_BigEarthNet_v1.

The preprocessed dataframes such as the square based price grid and pretrained neural models are stored in the authors private Cloud Storage and can be shared on request.

We strongly encourage the reader to have a look at this sources and reproduce or find inspiration by our experiments.

ACKNOWLEDGMENTS

We thank Mitra and Lau for the initial inputs and support in the creation of this project.

REFERENCES

[1] Archith J Bency, Swati Rallapalli, Raghu K Ganti, Mudhakar Srivatsa, and BS Manjunath. 2017. Beyond spatial auto-regressive models: Predicting housing prices with satellite imagery. In *2017 IEEE winter conference on applications of computer vision (WACV)*. IEEE, 320–329.

[2] James Bergstra, Dan Yamins, and David D Cox. 2013. Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms. In *Proceedings of the 12th Python in science conference*, Vol. 13. Citeseer, 20.

[3] M Ĉadík. 2008. Perceptual evaluation of color-to-grayscale image conversions. In *Computer Graphics Forum*, Vol. 27. Wiley Online Library, 1745–1754.

[4] Matthias Drusch, Umberto Del Bello, Sébastien Carlier, Olivier Colin, Veronica Fernandez, Ferran Gascon, Bianca Hoersch, Claudia Isola, Paolo Laberinti, Philippe Martimort, et al. 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote sensing of Environment* 120 (2012), 25–36.

[5] Richard HR Hahnloser, Rahul Sarpeshkar, Misha A Mahowald, Rodney J Douglas, and H Sebastian Seung. 2000. Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. *Nature* 405, 6789 (2000), 947–951.

[6] Sergey Ioffe and Christian Szegedy. 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167* (2015).

[7] Stephen Law, Brooks Paige, and Chris Russell. 2019. Take a look around: using street view and satellite images to estimate house prices. *ACM Transactions on Intelligent Systems and Technology (TIST)* 10, 5 (2019), 1–19.

[8] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. 1998. Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 11 (1998), 2278–2324.

[9] Donghun Lee, Woochang Hyun, Jeongwoo Ryu, Woo Jung Lee, Wonjong Rhee, and Bongwon Suh. 2015. An analysis of social features associated with room sales of Airbnb. In *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing*. 219–222.

[10] Richard Liaw, Eric Liang, Robert Nishihara, Philipp Moritz, Joseph E Gonzalez, and Ion Stoica. 2018. Tune: A Research Platform for Distributed Model Selection and Training. *arXiv preprint arXiv:1807.05118* (2018).

[11] Yuanhang Luo, Xuanyu Zhou, and Yulian Zhou. 2019. Predicting Airbnb Listing Price Across Different Cities. (2019).

[12] Jonas Mockus, Vytautas Tiesis, and Antanas Zilinskas. 1978. The application of Bayesian methods for seeking the extremum. *Towards global optimization* 2, 117-129 (1978), 2.

[13] Vinod Nair and Geoffrey E Hinton. 2010. Rectified linear units improve restricted boltzmann machines. In *ICML*.

[14] Herbert Robbins and Sutton Monro. 1951. A stochastic approximation method. *The annals of mathematical statistics* (1951), 400–407.

[15] Gencer Sumbul, Marcela Charfuelan, Begüm Demir, and Volker Markl. 2019. Bigearthnet: A large-scale benchmark archive for remote sensing image understanding. In *IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, 5901–5904.

[16] Gencer Sumbul, Jian Kang, Tristan Kreuziger, Filipe Marcelino, Hugo Costa, Pedro Benevides, Mario Caetano, and Begüm Demir. 2020. BigEarthNet Deep Learning Models with A New Class-Nomenclature for Remote Sensing Image Understanding. *arXiv preprint arXiv:2001.06372* (2020).

[17] Emily Tang and Kunal Sangani. 2015. Neighborhood and price prediction for San Francisco Airbnb listings.

[18] Han Xiao, Kashif Rasul, and Roland Vollgraf. 2017. Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms. *arXiv preprint arXiv:1708.07747* (2017).

[19] Junbo Zhang, Yu Zheng, and Dekang Qi. 2016. Deep spatio-temporal residual networks for citywide crowd flows prediction. *arXiv preprint arXiv:1610.00081* (2016).

[20] Ang Zhu, Rong Li, and Zehao Xie. 2020. Machine Learning Prediction of New York Airbnb Prices. In *2020 Third International Conference on Artificial Intelligence for Industries (AI4I)*. IEEE, 1–5.