Iteration and Implementation of Area Classification AI For Enhancing Civil Development Predictions

By: David Ajao, Elle Bates, Naisha Bhandari, and Jackson Choi   Mentors: Roayba Adhi and Oseremen Ojiefoh

C S R Center for Space Research

NASA SEES

NASA

Abstract

The U.S. is a diverse nation, both ethnically and economically containing exceptional differences in socioeconomic and urbanization levels throughout the nation. In order to maximize quality of life, especially within fading communities, the prediction of developmental and civil data is crucial for urban planners as it provides the basis and context for the problems within a community that should be addressed. Our goal was to be able to generate a tool that will be able to utilize citizen science pictures to supplement the existing geospatial data often used in civil development. Though pictures from citizen science are able to provide extra context on their own, the tool created would involve automated classification, generating more information at a faster pace.. BRAILS (Building and Infrastructure Recognition using AI at Large-Scale) is a Python library which utilizes deep learning and computer vision techniques. The BRAILS modules, Occupancy Classifier and Year Built Classification, fine-tune a model that outputs a score (1-100) for urbanization level. The CNN model, supplemented by the BRAILS library, increased accuracy by accommodating additional urban features predicted by the BRAILS modules. Our CNN model was trained on a dataset of 50 high quality representations of each urbanization level (urban, suburban, rural) from our SEES2024 GLOBE Observer database using Tensorflow, reaching exceptional accuracy in determining urbanization level. Currently, the model only expands on the urbanization information given from geospatial imaging however in the near future the model will be expanded by incorporating national databases like SEDAC and the US Census to provide even more civil data that can give more context to the imaging, helping civil developers further determine the best course of action.

Background information

Urban development: the planning and executing of a city that will increase social and economic growth. (Gov.si) Area classification is looking at an area and giving it descriptors. This can include: how much water is in the area, how many trees, or in our case, whether the photo is of an urban, suburban, or rural area. Building and Infrastructure Recognition using AI at Large-Scale. BRAILS takes images and satellite data and outputs building features. Refer to Figure 1

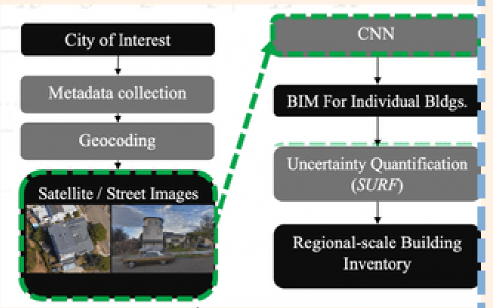

Figure 1

Our goal was to train an AI model that would be able to predict urbanization trends based on citizen science photos, satellite photos, and census data

Methods

1. Manually searched for and labeled 50 high quality representations of each urbanization level from the SEES2024 Globe Observer photos
2. Trained a CNN using TensorFlow to accurately determine urbanization level
3. Manually labeled and handpicked which photos were fit for BRAILS (had a clear building in them)
4. Supplemented CNN with the fine-tuned BRAILS Occupancy Classifier and Year Built Classification
5. Creating a function that output a score from 1-100 representing urbanization level
6. Currently working on supplementing time series imagery from Landsat-8 as well as socioeconomic data from SEDAC and US Census to plot trends and predict future numbers

Figure 2

Area Classification Model

OUTPUT \ TARGET	Urban	Suburban	Rural	SUM
Urban	35 32.41%	1 0.93%	0 0.00%	36 97.22% 2.78%
Suburban	4 3.70%	29 26.85%	3 2.78%	36 80.56% 19.44%
Rural	2 1.85%	1 0.93%	33 30.56%	36 91.67% 8.33%
SUM	41 85.37% 14.63%	31 93.55% 6.45%	36 91.67% 8.33%	97 / 108 89.81% 10.19%

Results

Urban: 97%

Suburban: 81%

Rural: 97%

Overall: 90%

Refer to figure 2 for the matrix confusion and figure 3 for the data table of area classification

Figure 3

Class Name	Precision	1-Precision	Recall	1-Recall	f1-score
Urban	0.9722	0.0278	0.8537	0.1463	0.9091
Suburban	0.8056	0.1944	0.9355	0.0645	0.8657
Rural	0.9167	0.0833	0.9167	0.0833	0.9167
Accuracy	0.8981				
Misclassification Rate	0.1019				
Macro-F1	0.8971				
Weighted-F1	0.8992				

Conclusions

Overall, the AI model we created using Tensorflow and BRAILS was very accurate for classifying the photos into the categories of urban, suburban, and rural. We can see that it was more accurate for the Urban and rural at 97% accuracy, but less accurate for Suburban at only 81%. A reason for this comes with the fact that suburban areas are literally, “sub-urban”, meaning that it is supposed to be in between rural and urban and the line differentiating the three blurs. The Next steps for us are to also add in satellite data and photos from the landsat-8 and census data, to further train the AI. We would also set stricter parameters for suburban images to increase the accuracy. The product we have created can be utilized by urban planners and designers, to create short and long term projects based on the needs of the city based on the data this AI model provides.

Sources Used

Gov.si definition of urban development: https://www.gov.si/en/policies/environment-and-spatial-planning/prostor-2/urban-development/

BRAILS user guide: https://nheri-simcenter.github.io/BRAILS-Documentation/common/user_manual/userGuide.html

GLOBE API - GLOBE.gov. (n.d.). https://www.globe.gov/ru/globe-data/globe-api

Google Colab. (n.d.). https://research.google.com/colaboratory/faq.html

Paper about using Arcgis and R to predict urban development: https://www.esri.com/arcgis-blog/products/arcgis-pro/analytics/a-deep-dive-into-predicting-urban-growth-using-arcgis-and-r/

NHERI SimCenter for BRAILS: https://simcenter.designsafe-ci.org/backend-components/brails/

An Image-Based Machine Learning Method for Urban Features Prediction With Three-Dimensional Building Information: https://www.researchgate.net/publication/369791313_An_Image-Based_Machine_Learning_Method_for_Urban_Features_Prediction_with_Three-Dimensional_Building_Information

https://www.kaggle.com/datasets/dansbecker/urban-and-rural-photos

SEDAC: https://sedac.ciesin.columbia.edu/

Acknowledgements

Thank you to all of our amazing mentors: Andrew Clark, Dr. Russanne Low, Peder Nelson, Cassie Soeffing, as well as the SEES program as a whole

The photos utilized in our project were found in the GLOBE Observer app.