Using A Deep Learning Object Detection Model to Identify Manufactured Units and Validate ACS Housing Counts

Camila Moreno, Research Associate
Katy Stigers, Vice President of Research

2025 ACS DATA USERS CONFERENCE

fahe
Strength in Numbers

Introduction

Due to sampling challenges in rural and lower density areas it is difficult to adequately know about the location of ***older manufactured homes*** and the potential demand for replacement of units.

Purpose:

Pilot project to develop a **methodology to count manufactured housing units in rural and other low density areas.**

In particular, our Member service areas across Appalachia

Introduction

- With construction costs and productivity being some of the greatest barriers to housing affordability, **manufactured homes are increasingly seen as a viable solution to the housing crisis**.
- Manufactured homes are an important source of housing for millions of Americans, with approx. ***7% of households living in 6.8 million manufactured homes*** *(U.S. Census Bureau)*.
- Manufactured units make up a significant proportion of the housing supply across the regions we serve e.g. rural.
- Manufactured housing does not follow the "rules" of regular subdivision – ***especially relevant as it pertains to zoning and ownership.***

fahe
Strength in Numbers

Despite the prevalence and importance of manufactured housing, there is ***currently no systematic way to count the total numbers and locations*** of these properties.

Strength in Numbers

Data, Tools and Methods

Data:

- USDA NAIP Aerial Images
- Local administrative data sets (property data)

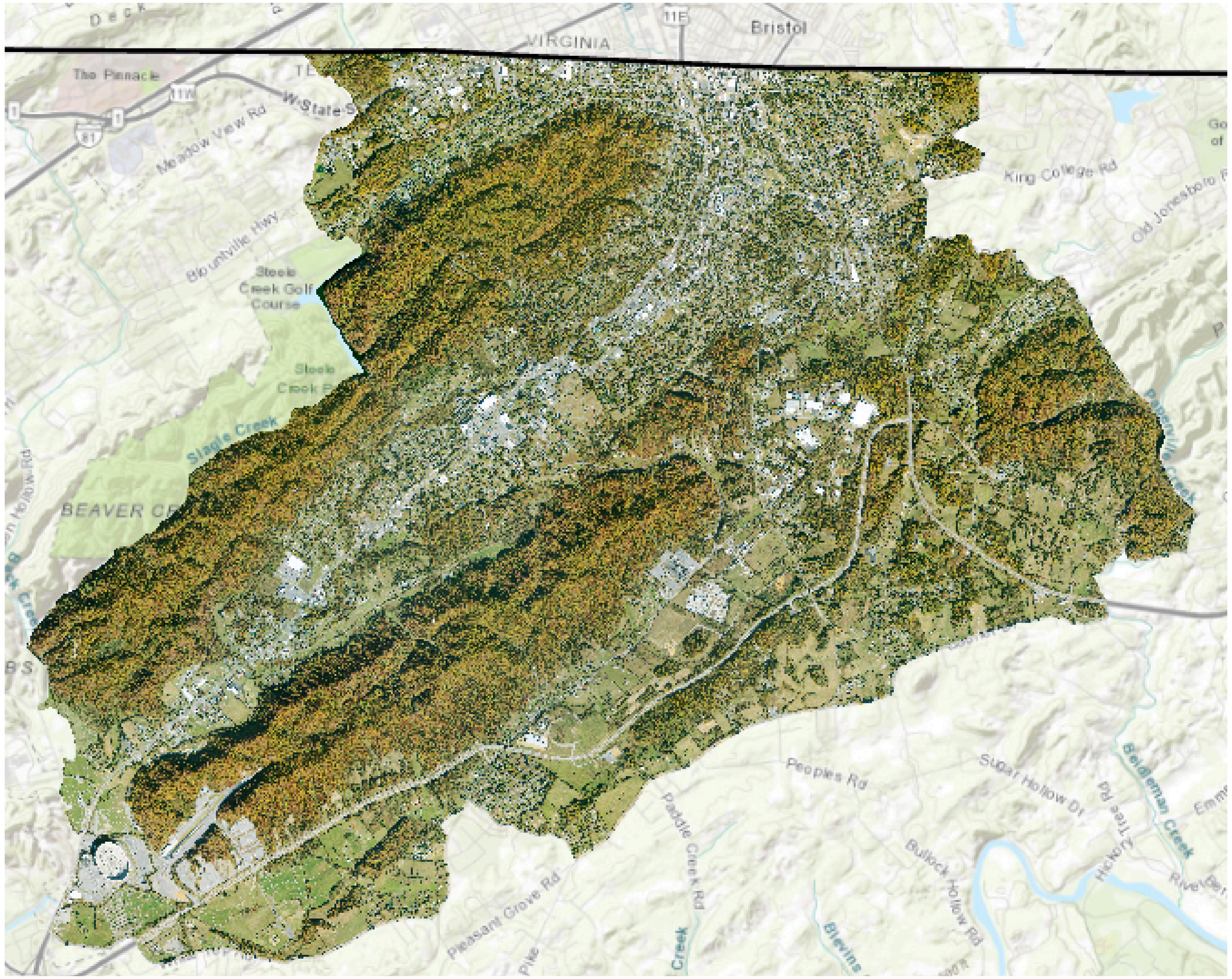
ACS Data: 5-year 2023

- Units in structure (B25024)

Tools & Methods:

- Machine learning - Object Detection
  - RCNN Mask
- Geographic Information Systems (GIS)
  - ArcGIS Pro v 3.4

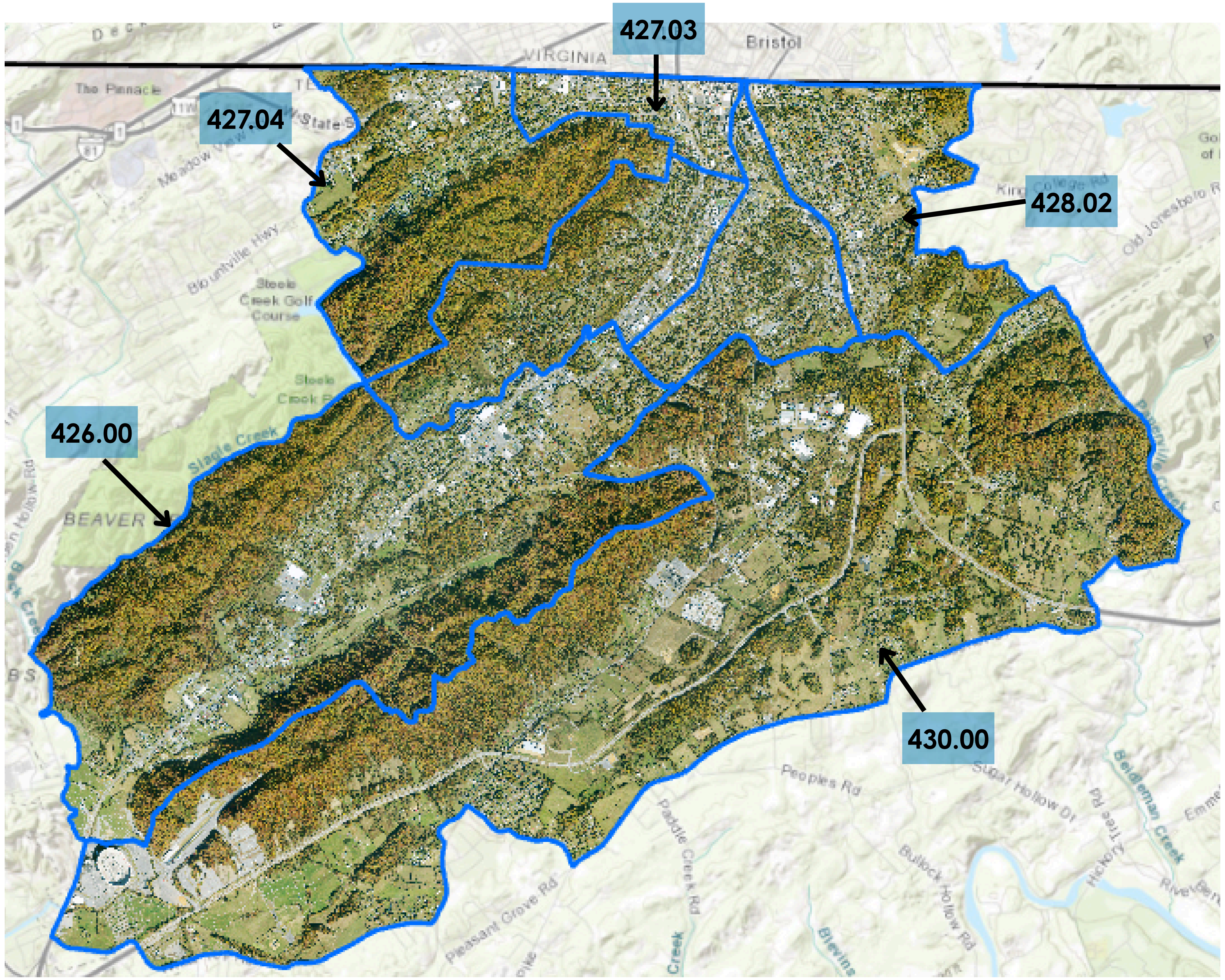
Area of Interest



NE region of Sullivan County, Tennessee

fahe
Strength in Numbers

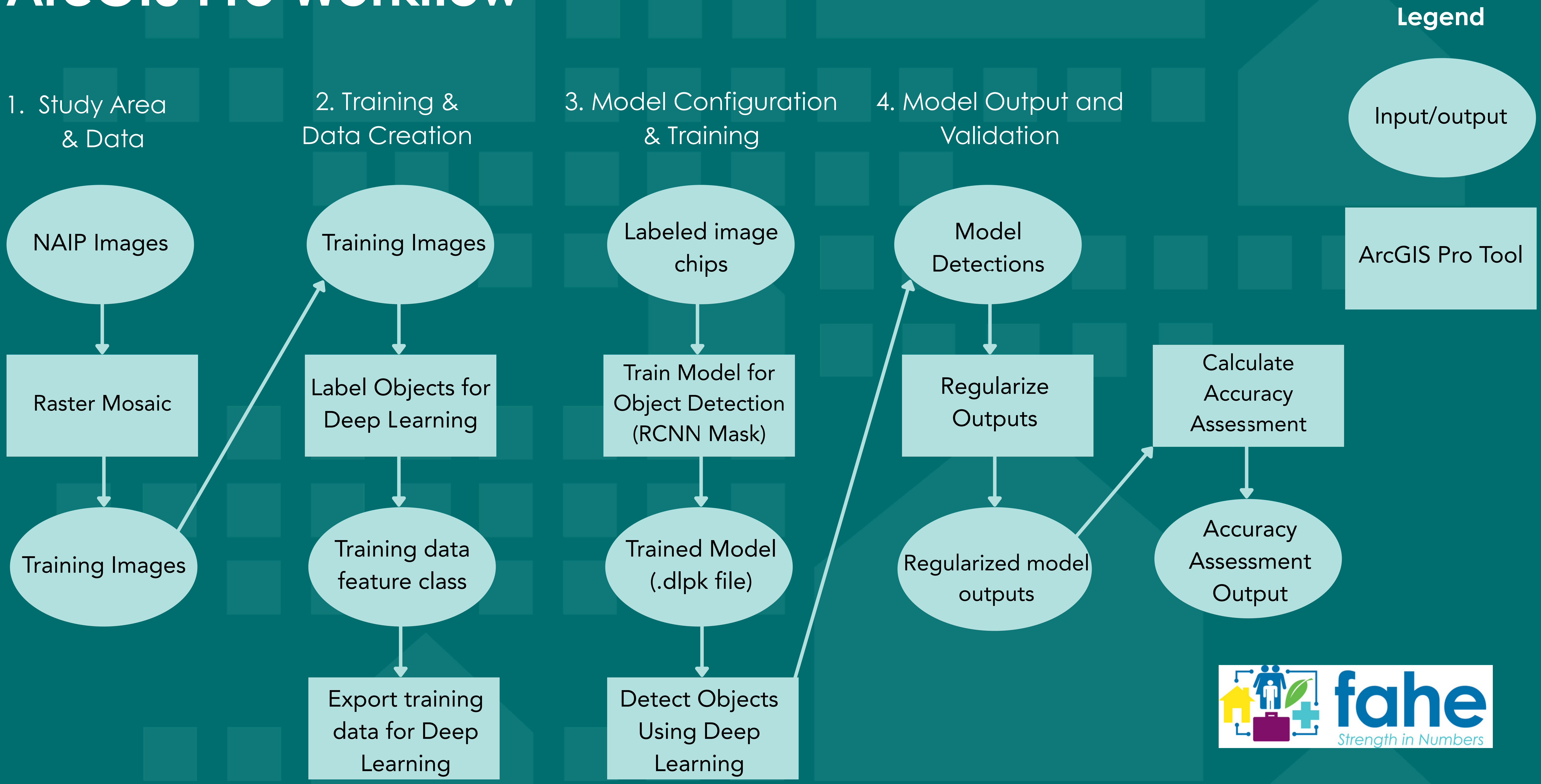
Area of Interest



Census Tracts

1. Tract 426.00
2. Tract 427.02
3. Tract 427.03
4. Tract 427.04
5. Tract 428.01
6. Tract 428.02
7. Tract 430.00

ArcGIS Pro Workflow



Training Image Samples

Manufactured Home Community

- A form of *alternative housing*
- Mixed with single and double wide units.
- More closely resembles “formal subdivision” infrastructure.
- Most units + land likely owned by residents

Mobile Home Park

- A form of *informal housing*
- Primarily single wide units. Some units are partially covered with tree canopy, creates challenges for visibility.
- Residents likely own the unit, but not the land.

fahe
Strength in Numbers

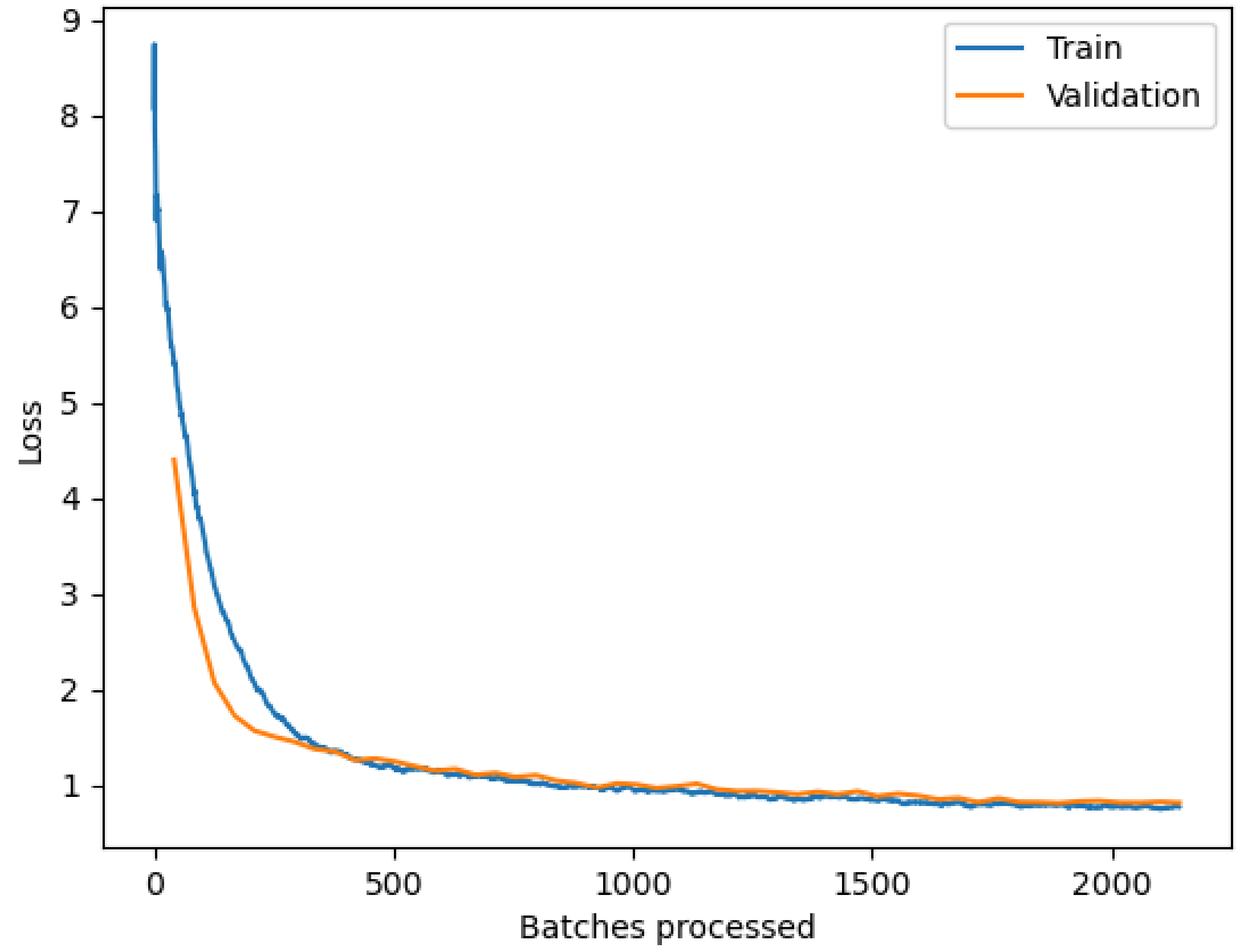
An estimated 80% of households in MHPs own their homes, however ***only 14% own the land*** (HAC 2011). It is more common for contemporary MHP’s to be ***cooperatively owned by residents or housing nonprofits*** (Sullivan, 2018).

fahe
Strength in Numbers

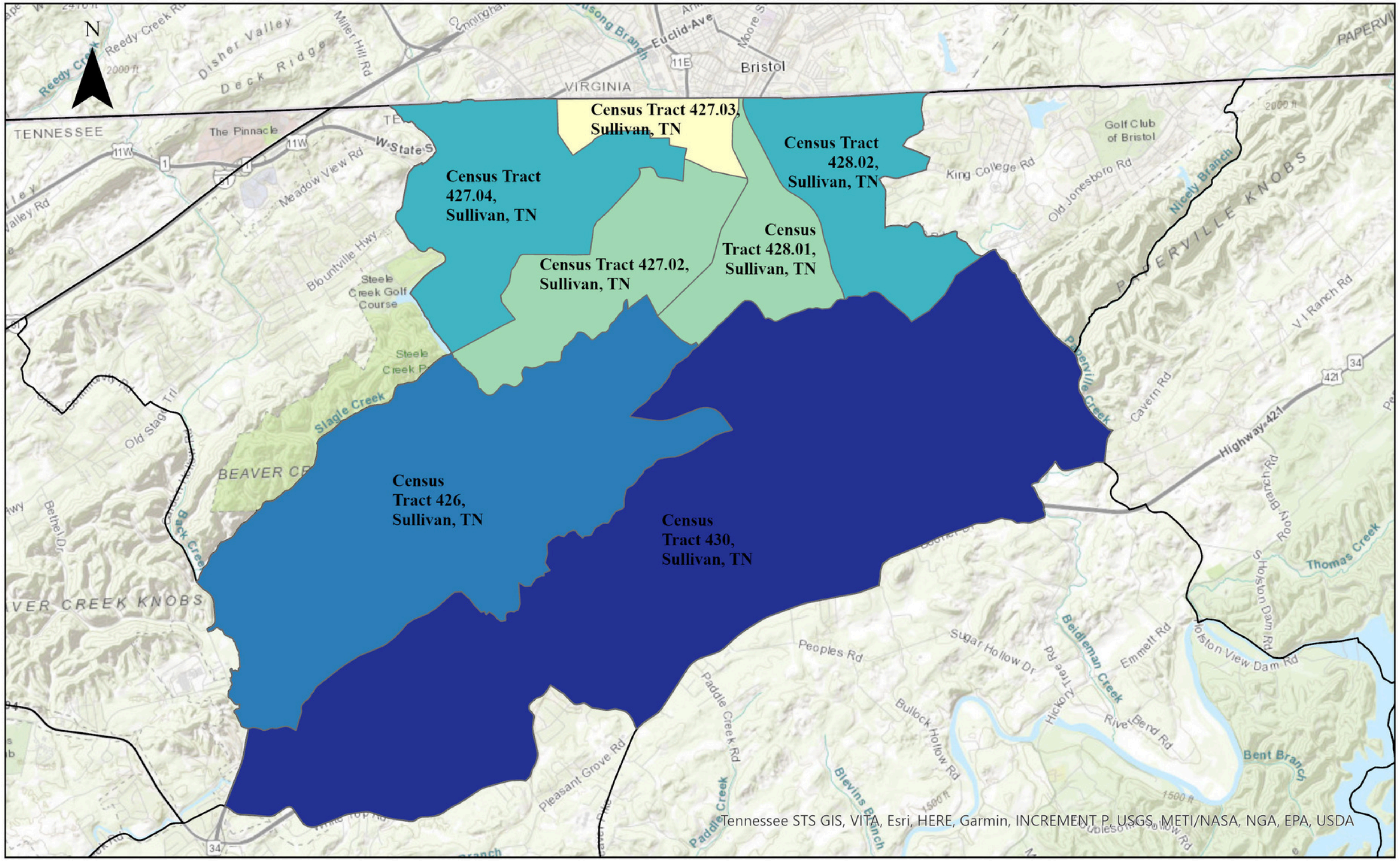
Interpreting Model Performance

Training vs. validation loss

- Both training loss and validation curves drop quickly and converge
- Lower validation loss = model is effectively learning patterns
- Reaches optimal "generalization" or learning performance between 500-1000 batches
- **AP score = 81.9%**



Number of Mobile Homes in NE Tennessee Census Tracts



Mobile Home Unit Count

- >16
- 16-20
- 21-88
- 89-149
- <500

Sullivan County Census Tracts

0 0.75 1.5 3 Miles

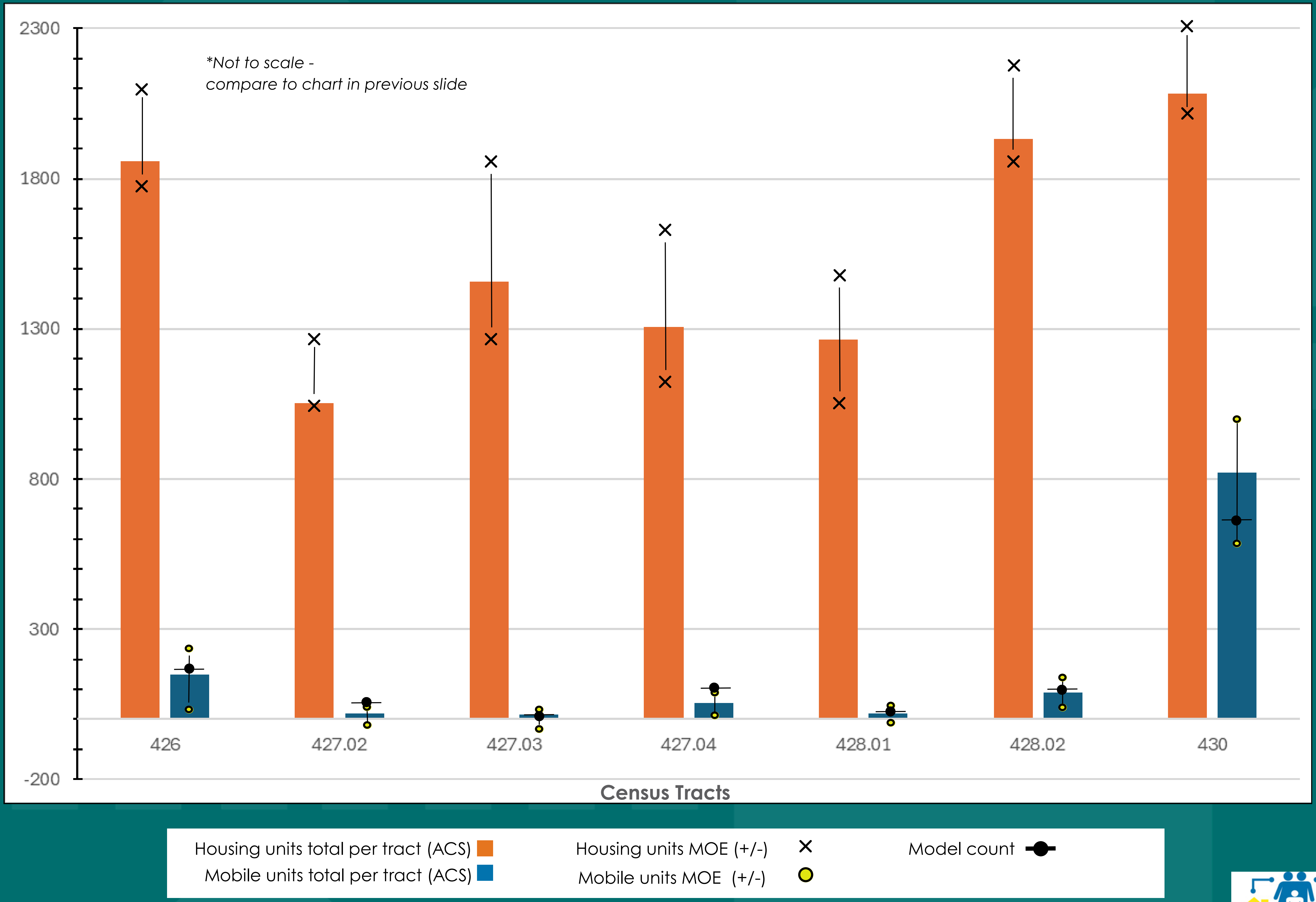
fahe
Strength in Numbers

Model Output Compared to ACS 2023 5-YR Units in Structure Data Table

TRACT	TOTAL HOUSING UNITS	MARGIN OF ERROR (+/-)	MOBILE HOMES	MARGIN OF ERROR (+/-)	MODEL COUNT	% MODEL ACCURACY
426.00	1857	155	149	84	196	62.8%
427.02	1053	127	18	21	64	84.4%
427.03	1456	304	15	26	11	0%
427.04	1305	243	55	36	104	76%
428.01	1264	177	20	21	43	0%
428.02	1932	142	88	57	114	35.1%
430.00	2083	215	821	191	621	91.8%

fahe
Strength in Numbers

Visualizing the Model Counts Compared to ACS



Interpreting Model Performance

- **Average Precision** (AP) is a common metric for object detection tasks. It evaluates the quality of the model's bounding box predictions and classifications across different confidence thresholds.

- Our initial model AP: **67.3%**, which indicates that the model learned to identify manufactured housing units to a ***moderate degree.*** This means that on average, the model was successful 2 out of 3 predictions. Our model is making some incorrect detections (**false positives**) and/or failing to detect some actual units (**false negatives**).

fahe
Strength in Numbers

Understanding the Patterns in the Model's Mistakes

- **False Positives:** What is the model incorrectly identifying as manufactured housing? Are there common features confusing it?
- **False Negatives:** What manufactured housing units is the model missing? Are they too small, occluded, or have unusual appearances?
- **Localization Errors:** Are the bounding boxes and masks accurate for correctly classified objects?

fahe
Strength in Numbers

Model Improvments

Overfitting: when a model learns the training data too well and cannot make accurate predictions on new, unseen data

How to overcome this:

- Improve training sample dataset
    - increase number of samples
    - Use multiple images to increase diversity
- Attempt running the model with a variety of parameters to test differences in outputs
- Collection of ground truth points for validation of accuracy

Challenges and Limitations

Limitations

- Computational power
- Human labor constraints

Challenges

- Aerial imagery
  - processing
  - spatial resolution
- Nature of the study area
  - Tree cover
  - Changes in slope
  - Shadows

Next Steps

- ***Our purpose is to help practitioners identify areas of need and support planning by local developers.*** It can also help inform future policy and program development.

- ***Expand on other pre-trained building footprint models.*** *We would like to successfully replicate this approach where local property data may not be available to help validate administrative sources.*

- *Recruiting and training volunteers and staff to help build our database of observations to use for validating model training (ArcGIS Survey123)*

- ***Moving beyond Appalachia and identifying any systematic biases that might exist in ACS sampling*** (i.e. are some areas consistently under or over counted in ACS relative to the observations - why might that be?)
    - local zoning
    - Hard to reach places
    - Urban v rural

fahe
Strength in Numbers

References

- Council, H. A. (2011). Rural housing research note preserving affordable manufactured home communities in rural America: A case study. Washington DC: Housing Assistance Council.
- Durst, N. J., Sullivan, E., Huang, H., & Park, H. (2021). Building footprint-derived landscape metrics for the identification of informal subdivisions and manufactured home communities: A pilot application in Hidalgo County, Texas. Land Use Policy, 101, 105158.
- Jones, M., Koebel, C. T., McCoy, A. P., Shanholtz, S. A., & Moeller, J. (2016). Mobile and manufactured homes in Central Appalachia and Alabama: age, condition and need for replacement.
- Jones, M., Eades, D., & Choi, S. (2023). Housing Needs and Trends in Central Appalachia and Appalachian Alabama.
- Mccue, C. R. (2024). A Review of Barriers to Greater Use of Manufactured Housing for Entry-Level Homeownership.
- National Mobile Home Park Database, https://www.arcgis.com/home/item.html?id=4cdbccc5c538452aa91ceee277c460f9
- Property Assessment Data, https://assessment.cot.tn.gov/TPAD
- Sullivan, E. (2018). Manufactured insecurity: Mobile home parks and Americans’ tenuous right to place. Univ of California Press.
- U.S. Census Bureau, U.S. Department of Commerce. "Tenure." American Community Survey, ACS 5-Year Estimates Detailed Tables, Table B25003, 2023
- U.S. Census Bureau, U.S. Department of Commerce. "Units in Structure." American Community Survey, ACS 5-Year Estimates Detailed Tables, Table B25024, 2023
- U.S. Geological Survey, NAIP GEOTIFF, 2023, Accessed March 2025, National Agriculture Imagery Program (NAIP) Digital Object Identifier (DOI) number: /10.5066/F7QN651G

Thank You

fahe

Strength in Numbers