NASA SEES 2023: AI Land Cover Algorithm Comparison Team

Analyzing AI land cover Algorithms' classification abilities in comparison to human classification

Abstract

Our objective is to examine the accuracy of AI land cover classification algorithms for the development of a citizen science app intended to be paired with the GLOBE observer. Artificial intelligence has incredible potential for land cover classification application. It would increase efficiency in land cover monitoring, improve citizen science accessibility, and allow the public to more effectively contribute to land cover classification by removing any discrepancies caused by personal bias. Utilizing AI monitoring of land cover would help to predict droughts, natural disasters, weather patterns, ect. faster. Additionally, it would create an easier process for citizen scientists to contribute to land cover data.

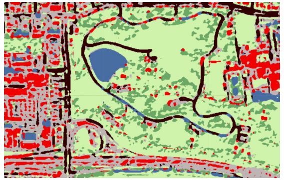
We are examining the classification abilities of PEARL an AI land cover app developed by Sajjad Anwar, Lane Goodman, Martha Morrissey, Nick Ingalls, Vincent Sarago, Vitor George, Sanjay Bhangar, Jeevan Farias, and Zhuang Fang NaN Yi. We will then utilize the Earth System Explorer's group areas of interest to compare the software's predictions to human classification.

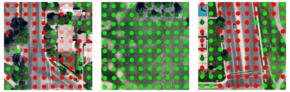
We aim to specifically identify the most prevalent mistakes the AI models make in classification, i.e. what land cover types are mislabeled the most, what areas in the United States these errors are most prevalent, and if there are patterns of error in certain areas. This data will help NASA scientists more accurately understand the biases present in AI-classified data. It will also help us understand AI's capacity in the field of land cover as is. Finally, our data will help scientists create a more accurate margin of error, and perform more accurate statistical tests

Methods

We began by defining areas of interest (AOIs) within each of our local communities. Each AOI covered a 100m by 100m area and contained 37 individual points. At each point, we used the GLOBE Observer app to collect photographs and analyze the land cover of the area. Additionally, we uploaded our grid to Collect Earth Online where we individually analyzed satellite images of the same areas. We used this data as our human classification to compare with the AI algorithms. Now that we acquired our data on how the points were classified by humans, our next task was to determine how an AI would classify the same area in order to spot discrepancies. We utilized the platform PEARL which was developed by Microsoft to inference land cover data across the united states. From there we calculated the frequency of each land cover classification per AOI vs the AI model's general percentages. From there we outlined the categories in which misclassifications were most common.

Sugar House Park, Salt Lake City, UT





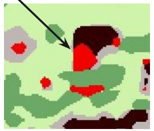
Acknowledgements

Thank you to SEES Earth System Explorer mentors; Dr. Rusanne Low, Ms. Cassie Soeffing, Mr. Peder Nelson, Dr. Erika Podest, Andrew Clark, and Marvel Hanna!

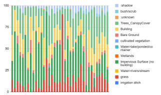
Preliminary Results

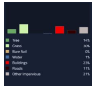
The average total difference between the human classified AOIs and the PEARL Fort Collins model classified AOIs was 5.55547%. This result is surprisingly promising! The PEARL model's primary issues lie with the tree cover and grass cover categories. It had a higher rate of difference for AOI 21, located in Salt Lake City UT, as compared to AOI 19, in Pearland TX. The primary problem with the AI's classification can most likely be attributed to challenges with more urban areas where land cover types tend to be closer and less distinct from one another.

Misclassified water as impervious surface and road



Trees between classification points





AOI 21	Human classification	PEARL AI classification
Tree cover	25.0541%	14%
Grass	14.1622%	30%
Bare Ground	3.983%	0%
Water	00.0811%	1%
Buildings	18.4865%	23%
Roads/imperious ground	32.9459%	32%

Expansion

Our research concerns an AI algorithm that was trained on data from the east coast of the United States. As a result we have determined that using this AI in southern and western regions of the United led to some misclassification in certain categories. For example any human-classified barren ground was classified as roads or other impervious by the AI. This discrepancy could be remedied if the AI was trained with data from more and varied geographic regions.
In the future for this project,

Conclusion

We predicted that the AI's abilities would be as good or better than human classification. However, we didn't have a 100% accurate method of classifying land cover data, so we examined both the AI's and human's predictions and found the human classification to be better. We noticed this via some blatantly inaccurate predictions, for instance, the AI misclassified water as a road/impervious surface in AOI 19. The AI also falls short in areas where the land cover is ambiguous, and multiple categories overlap in a small space. Our study helped us understand AI's current capacity for accurate classification in the field of land cover. It shows the potential for AI as a land cover classification tool for NASA scientists, and the areas that need to be improved on prior to its debut as a NASA tool. Utilization of data from the southwestern United States will help improve its capabilities because of its variety of vegetation types, expanding the AI's classification capabilities.

References

Collect Earth Online (n.d.). COLLECT EARTH ONLINE. Satellite Image Viewing & Interpretation System. https://app.collect.earth/home
development SEED (n.d.). PEARL Planetary Computer Land Cover Mapping. PEARL: Land Cover Mapping. https://www.landcover.io/
Esri (n.d.). ArcGIS Field Maps. ArcGIS Online. https://www.arcgis.com/apps/fieldmaps/maps
Esri (n.d.). Sentinel-2 Land Cover Explorer. ArcGIS Living Atlas of the World.
https://livingatlas.arcgis.com/landcoverexplorer/#mapCenter=-111.901%2C40.736%2k13 Mode=step time Extent=2017%2c2022 Year=2017
Esri (n.d.). ArcGIS Notebooks. Institute for Global Environmental Strategies.
https://igestrategies.maps.arcgis.com/home/notebook/notebookhome.html
Jones, T. M. (n.d.). Supervised learning models. IBM Developer. https://developer.ibm.com/articles/cc-supervised-learning-models/
(n.d.). Decision Tree Regression. Scikit Learn. https://scikit-learn.org/stable/auto_examples/tree/plot_tree_regression.html
(n.d.). RStudio IDE. RStudio. https://www.rstudio.com/categories/rstudio-ide/
R, S. E. (n.d.). Understand Random Forest Algorithms With Examples (Updated 2023). Analytics Vidhya.
https://developer.ibm.com/articles/cc-supervised-learning-models/
Sharp, T. (n.d.). An Introduction to Support Vector Regression (SVR). Medium.
https://towardsdatascience.com/an-introduction-to-support-vector-regression-svr-a3ebc1672c2

Partner