International Cartographic Association
Association Cartographique Internationale

Automating conversion of remote sensing images to human readable map images using generative AI

Nakul Gupta [a,*]

[a] *Faculty at Management Development Institute Gurgaon, India, nakul.gupta@mdi.ac.in and Student of Online M.S. Programme at Department of Geography & Geographic Information Science, University of Illinois at Urbana-Champaign, USA nakulg3@illinois.edu*

* Corresponding author

Keywords: Generative AI, remote sensing, game cartography, pixel accuracy

Abstract:

A map could alternatively be thought of as data that is not always machine-readable. There is a growing need for real-world cartography and feature-laden maps in metaverse applications and open-world video games (Butcher et al. 2018). A precise map is required in these use cases that reflect all changes on the ground promptly. Flyover imaging planes or satellites are constantly gathering the most recent geographical data. Automating the process of creating a human-readable map from a satellite image of a given location at a specified zoom level and resolution is one way to decrease this latency. This needs an ensemble of image translation, digital image processing, and remote sensing fields and is the focus of this research study. This has applications in diverse fields ranging from on ground navigation, metaverse, architecture and open-world game development.

Game developers are playing catchup, learning well-established photogrammetry techniques from other fields, while at the same time, developing specialized tools and pipelines that allow its smooth integration into the technically challenging limits of game engines (Maximov, 2017). In remote sensing image classification, each pixel or area in the picture is given a semantic label and this process can be used for generative meaningful versions of remote sensing images. For example, human-readable versions of maps. This is also referred to as image-to-image translation. Generative AI and modelling have shown promise in enabling this pursuit of image-to-image translation in recent times (Chan et al. 2018). Low separation between spectral classes and the difficulty of learning distinguishing characteristics makes remote sensing picture categorization or semantic labelling problematic. Due to the generator's ability to replicate samples and the discriminator's ability to enhance the generator, GAN (Generative Adversarial Networks) looks promising for this job. Pix2Pix performs the best among GANs variations for image translation and image classification tasks (Wang et al. 2018). However, Pix2Pix has limitations when it comes to specifically mapping the connection between the source domain and the destination domain's rebuilt ones. Attempting to solve this limitation of Pix2Pix in the context of remote sensing image's translation to human readable images is the key tangible research pursuit of the current research. The key research question explored in this paper is – "How to perform image to image translation in the context of cartography by translating remote sensing images into human-readable images (using paired images as training data)"? We have taken a post-positivist stance during this research. The generative AI model proposed in the current research uses a satellite picture with a certain zoom level and resolution as input to create a map that can be read by humans. For the research, the open access pix2pix dataset by EECS UC Berkeley which consists of 1096 concatenated satellite images with their corresponding map images. To be more precise, the input to the generative AI model was a Source image (1200x600px) and outputs were a) classification of real/fake and b) Generated target image.

Our generative AI model is a combination of a generator and a discriminator system built on the GAN model. A mapping from the source remote sensing picture and random noise vector to the target image was learned by the generator. The discriminator differentiated between authentic and fraudulent label (classification). To be precise, we used conditional adversarial network generator. The performance of both generator and discriminator models were improved by training them concurrently in an adversarial approach in which they both attempt to deceive one another. The result was efficient translation of remote sensing images to human readable maps. The key thing here was the definition of loss functions. For this research, jupyter notebooks on accelerated GPU were used and python language as was the programming language for our research. Given the unimodal nature of the target translation function in the task of generating electronic maps, we used the average pixel translation accuracy to evaluate the performance of the model in map translation tasks. The percentage of correctly translated pixels in a remote sensing image is the accuracy of the model in a map translation task. We expected this average pixel translation accuracy to go up and it did. This augmentation in average pixel translation accuracy should not only be seen in relation to existing generative cartography models but also in terms of the

Abstracts of the International Cartographic Association, 6, 82, 2023.
31st International Cartographic Conference (ICC 2023), 13–18 August 2023, Cape Town, South Africa.
https://doi.org/10.5194/ica-abs-6-82-2023

possibilities of return-on-investment savings when maps for gaming and metaverse applications are created by hand or with machine/human in the loop.

Acknowledgements

The author would like to acknowledge the support of Prof. Chunyuan Diao, Assistant Professor, Department of Geography & Geographic Information Science, University of Illinois at Urbana-Champaign and James Oneil Smith, Graduate Student at Department of Geography & Geographic Information Science, University of Illinois at Urbana-Champaign towards this research project.

References

Arjovsky, M.; Chintala, S.; and Bottou, L. 2017. Wasserstein Generative Adversarial Networks. *Proceedings of the 34th International Conference on Machine Learning*, Sydney, Australia, PMLR 70, 2017.

Bucher, B., Schlieder, C., Cantat, F., Kavouras, M., Streilein, A., & Severo, M. (2018). Mapping places for digital natives and other generations.

Chan, C.; Ginosar, S.; Zhou, T.; and Efros, A. A. 2018. Everybody dance now. arXiv preprint arXiv:1808.07371.

Ganguli, S., Garzon, P., & Glaser, N. (2019). GeoGAN: A conditional GAN with reconstruction and style loss to generate standard layer of maps from satellite images. arXiv preprint arXiv:1902.05611.

Goodfellow, I. J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; and Bengio, Y. 2014. Generative Adversarial Networks. arXiv:1406.2661.

Gulrajani, I.; Ahmed, F.; Arjovsky, M.; Dumoulin, V.; and Courville, A. 2017. Improved Training of Wasserstein GANs. arXiv:1704.00028.

Wang, X., Yan, H., Huo, C., Yu, J., & Pant, C. (2018). Enhancing Pix2Pix for remote sensing image classification. In 2018 24th International Conference on Pattern Recognition (ICPR) (pp. 2332-2336). IEEE.

Statham, Nataska. “Use of Photogrammetry in Video Games: A Historical Overview.” Games and Culture 15, no. 3 (2018): 289–307.

Abstracts of the International Cartographic Association, 6, 82, 2023.
31st International Cartographic Conference (ICC 2023), 13–18 August 2023, Cape Town, South Africa.
https://doi.org/10.5194/ica-abs-6-82-2023 |