CartoGAN: Map Synthesis with Cartographic Design Based on Remote Sensing Images

CS766 Project Proposal, Spring 2021

Meiliu Wu (mwu233@wisc.edu), Yilei Hu(yhu366@wisc.edu)

Introduction

Map-related services are essential in our daily life, as we utilize maps for various scenarios/applications, e.g., daily commuting navigation, logistics distribution system, queries and visualization of geographic information, request of high-definition maps for self-driving vehicles, etc. Therefore, generating applicable maps and maintaining their latest versions are important tasks, which, unfortunately, can be a laborious and time-consuming process. Nowadays, most maps are created and updated based on the interpretation of aerial/satellite images and field surveys. On the bright side, with the rapid development of remote sensing technologies, many high spatial resolution (HSR) remote sensing images with a global coverage can be obtained frequently by sensors on aircraft/satellites. Thus, **generating map tiles automatically and aesthetically based on HSR remote sensing images** has become an emerging research direction for mapping agencies and institutions.

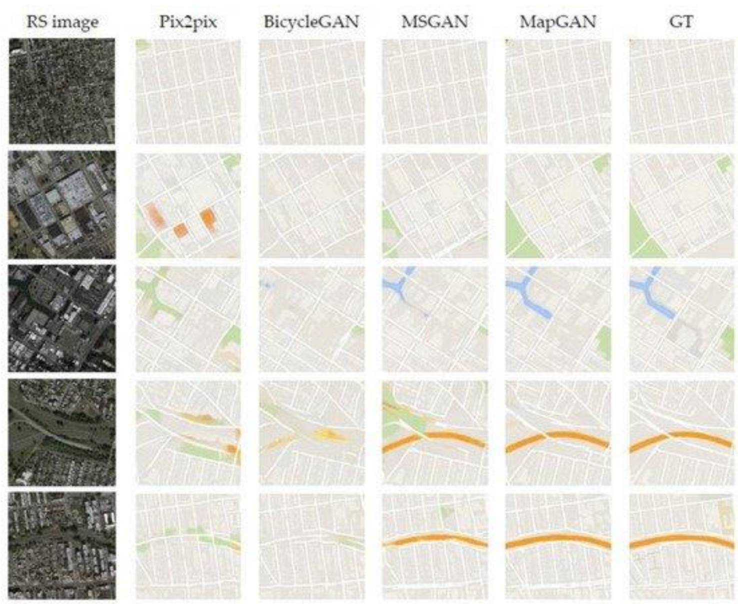

Figure 1. Comparison results of MapGAN (Li et al., 2020) and some other image translation models in the one-to-one domain map generation experiment to generate Google maps. The images from left to right are remote sensing images and the Google maps generated by MapGAN, Pix2pix, BicycleGAN, the MSGAN, and MapGAN, the real Google maps.

Literature Review

Generative Adversarial Network (GAN), a framework for estimating generative models via an adversarial process, was first proposed (Goodfellow et al., 2014) with two models being trained simultaneously: a generative model G that captures the data distribution, and a discriminative model D that estimates the probability that a sample came from the training data rather than G. The training procedure for G is to maximize the probability of D making a mistake. This framework corresponds to a minimax two-player game.

Later, researchers have experimented with their proposed GANs on the translation task using remote sensing images and Google maps. However, the purpose of previous map translation tasks was mainly to prove the feasibility of their proposed model. That is, as long as the generated images were in the style of applicable electronic maps, they were satisfied with the results, even although the quality of generated maps was far from ideal, and it was easy for people to distinguish between the real electronic maps and the generated ones. For example, Pix2pix (Isola et al., 2017) established a general framework for image translation based on a CGAN (Mirza & Osindero, 2014). However, satisfactory results cannot be achieved in specific scenarios, and the quality of the generated electronic map is poor. A breakthrough of CycleGAN (Zhu et al., 2017) is its ability to solve the problem of image translation in cases where paired training datasets cannot be obtained. When it is applied in the map translation scenario, it is still found that the resulting electronic map has many problems, such as image blurring, unclear texture, and incorrect color rendering. However, there are two recent, promising, and GAN-based studies about map synthesis and design using remote sensing images (Ganguli et al., 2019; Li et al., 2020), which are well worth exploring.

1. *GeoGAN*

GeoGAN translates satellite images to the corresponding electronic map image using three main model architectures: (i) a conditional GAN which compresses the images down to a learned embedding, (ii) a generator which is trained as a normalizing flow (RealNVP) model, and (iii) a conditional GAN where the generator translates via a series of convolutions to the electronic map layers and the discriminator input is the concatenation of the real/generated maps and the satellite images.

In order to improve the results, GeoGAN is the first model to incorporate a reconstruction loss (for pixel-wise accuracy) and a style loss (to reduce high frequency artifacts) in addition to the GAN loss (a feature-wise learnt similarity metric or content loss) for the task of generating the electronic map from a satellite image. However, one challenge of this study was the coarse resolution of the dataset which upper-bounds the quality of the results of the GeoGAN model. Additionally, the GeoGAN model is often confused by water against vegetation due to the lack of usage of more bands such as the infrared band.

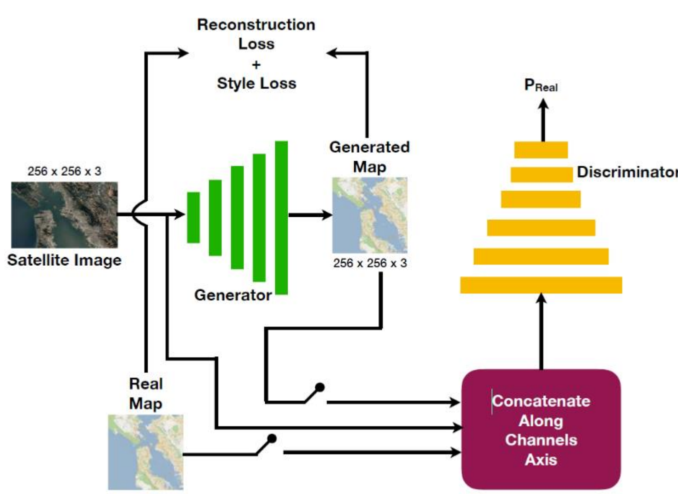

Figure 2. A schematic of the GeoGAN model without encoding.

2. Map Generative Adversarial Networks (MapGAN)

The other model, called **Map Generative Adversarial Networks (MapGAN)**, was proposed for generating multitype electronic maps accurately and quickly based on both remote sensing images and render matrices. MapGAN improves the generator architecture of Pix2pixHD and adds a classifier to construct map classification loss so as to improve the model's feature recognition and multitype map generation capabilities.

However, in order to construct the render matrices, multiple types of electronic maps are needed. Additionally, if an electronic map with many features needs to refer to third-party database information for color rendering, multiple render matrices need to be created, but considering that most of the elements of the render matrix in the model input are 0, there exists the problem of memory resource waste to some extent, which would possibly slow down the training speed of the model.

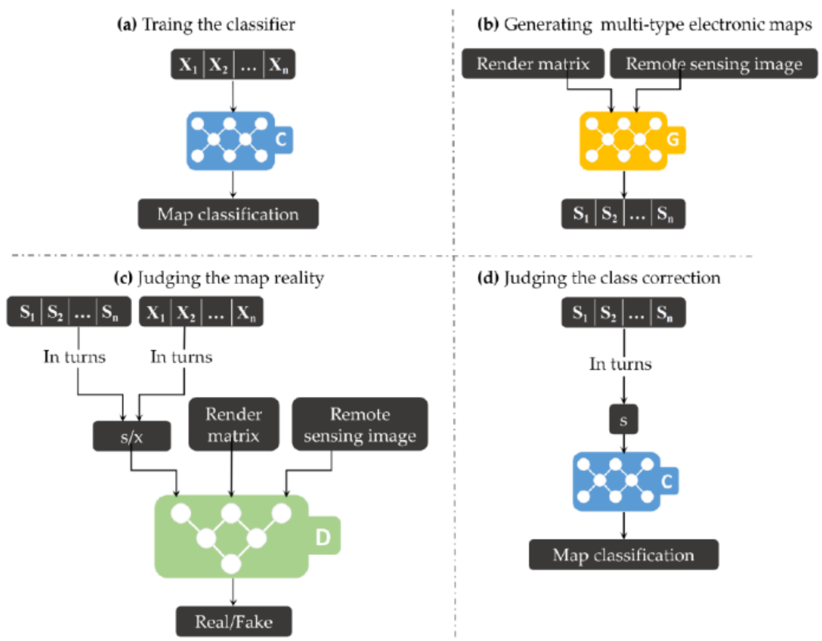

Figure 3. The overall architecture of Map Generative Adversarial Networks (MapGAN), consisting of three modules: a discriminator, D, a generator, G, and a classifier, C. (a) The authors trained classifier C to have map classification capabilities using n types of real electronic maps. (b) G outputs the generated n type electronic maps based on remote sensing images and the render matrix, (c) D tries to distinguish between real and generated maps corresponding to remote sensing images and the render matrix. (d) C tries to judge the category of each generated map.

Map Synthesis with Cartographic Design (CartoGAN)

In contrast to the former methods, this project aims to propose a CartoGAN model for generating electronic maps that can be in multiple types and more realistic and aesthetic. In order to accomplish this goal, the steps below are required:

1. Collect multiple types of electronic maps, e.g., Google Maps, OpenStreetMap, Baidu Maps, etc.
2. Search for usable HSR remote sensing images corresponded to those electronic maps, and use multiple bands (e.g., infrared) in order to improve the model's ability in feature recognition.

3. Construct a classifier similar with (Li et al., 2020) to help the model learn the differences among multiple types of electronic maps so that the model can determine whether the generated electronic map belongs to the correct type during the training process.
4. Incorporate the concept of “render matrix” and think about its structure with less memory used.
5. Build effective architectures of the generator and the discriminator.
6. Consider suitable loss functions for the model, e.g., reconstruction loss (for pixel-wise accuracy), a style loss (to reduce high frequency artifacts), and the GAN loss (a feature-wise learnt similarity metric or content loss).
7. Determine the evaluation metrics of the model, e.g., Kernel Maximum Mean Discrepancy (Kernel MMD), Fréchet Inception Distance (FID), Mode Score, Inception Score, Pixel-Level Translation Accuracy, etc.
8. Compare the results from our model with the ones from other state-of-the-art GANs.

Tentative Schedule

Tasks	**Expected due date**
Complete research references	Feb 27
Search for suitable datasets	March 6
Implement our CartoGAN model by referencing the existing methods	March 20
Draft the mid-term report	March 24
Try to improve the accuracy and aesthetics of our results	April 10
Compare our model with other state-of-the-art models	April 17
Prepare for the final presentation	April 23
Finish refining the webpage	May 5

References

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*.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial networks. *Communications of the ACM*, *63*(11), 139–144. https://doi.org/10.1145/3422622

Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, *2017-Janua*, 5967–5976. https://doi.org/10.1109/CVPR.2017.632

Li, J., Chen, Z., Zhao, X., & Shao, L. (2020). MAPGAN: An intelligent generation model for network tile maps. *Sensors (Switzerland)*, *20*(11). https://doi.org/10.3390/s20113119

Mirza, M., & Osindero, S. (2014). *Conditional Generative Adversarial Nets*. 1–7. http://arxiv.org/abs/1411.1784

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. *ArXiv*, 2223–2232.