

Automatic Generation of Photorealistic Image Fillers for Privacy Enabled Urban Basemaps using Generative Adversarial Networks

Amgad Agoub *, Yevgeniya Filippovska, Valentina Schmidt, Martin Kada

Institute of Geodesy and Geoinformation Science (IGG), Technische Universität Berlin, Straße des 17. Juni 135, 10623 Berlin, Germany, firstname.lastname@tu-berlin.de

* Corresponding author

Abstract: The abundance of high-quality satellite images is salutary for many activities but raises also privacy and security concerns. Manually obfuscating areas subject to privacy issues by applying locally pixelization techniques leads to undesirable discontinuities in the visual appearance of the depicted scenes. Alternatively, automatically generated photorealistic fillers can be used to obfuscate sensitive information while preserving the original visual aspect of high-resolution aerial images.

Recent advances in the field of Deep Learning (DL) enable to synthesize high-quality image data. Particularly, generative models such as Generative Adversarial Networks (GANs) can be used to produce images that can be perceived as photorealistic even by human examiners. Additionally, Conditional Generative Adversarial Networks (cGANs) allow control over the image generation process and results. These developments give the opportunity to generate photorealistic fillers for the purpose of privacy and security in image data used within city models while preserving the quality of the original data. However, according to our knowledge, little research has been done to explore this potential. In order to close this gap, we propose a novel framework that is designed to correspond to the mentioned end goal and produces promising results.

Keywords: Geo Privacy, Deep Learning, GANs, Aerial Images, Buildings

1. Introduction

The increasing number and variety of earth observation sensors are leading to more and more publicly available data than ever before. As shown by Stagars (2016), the availability of spatial data with a high temporal frequency can have a positive correlation to the increase in prosperity of a country. Such data facilitate a variety of important applications such as effective and successful planning, crises management and situational awareness (Paderta, 2012). However, the increased availability of highly detailed public data can risk exposing private information about citizens. Many public and governmental bodies stress the importance of protecting the privacy of individuals. Siemoneit et al. (2009) and Kada et al. (2009) discuss the importance of privacy protection and the related issues in geospatial systems, especially for European countries.

Areas within images that are subject to privacy or security concerns are often obfuscated using pixelization techniques. Even though such a simple method can be easily automated, the result is arguably not pleasing. The pixelated areas lead to an interruption of the visual cognition process of these images and causes a counter-purpose effect by steering attention toward obfuscated locations. For this reason, it would be advantageous to produce patches or fillers with a photorealistic

appearance, which correspond to the surrounding environment within the original images, thus minimizing visual interruption.

It is laborious to manually obfuscate individual images by creating fillers or patches that correspond to the surrounding environment while preserving privacy or eliminating security concerns. In addition, such locations can be changed and updated depending on the activities that are held within them. This means that obfuscated parts of the images will change and privacy patches need to be re-created. The enormous quantity of satellite and aerial data captured on a daily bases and the high revisit rate raise a need for automated methods with regard to privacy protecting

Two- and especially three-dimensional (2D and 3D) urban scenes are complex. They typically contain a variety of different objects such as multiple types of buildings, roads sidewalks and natural phenomena such as trees and rivers. Therefore, synthesizing scenes with the same visual appearance in a generic and automatic way is challenging. For example, generating 2D and 3D scenes using procedurally generated graphics requires a deep understanding of the scenes' environment. The generation process requires many starting assets and constraints, which have to be implemented by experts and can differ from one to another location, the season of the year, and time of the day. For example, an image of the

same location can look very different in the middle of a sunny day in springtime than a late afternoon in winter. Many plants can have a very different appearance depending on their yearly cycles. A generic framework should have the capability to be adaptable to different conditions easily.

The last decades witnessed an increased in development and use of machine learning and particularly Deep Learning techniques. This rise can be credited to the abundance of data and computational power that made it possible to build powerful mathematical models that can solve complex problems without the need of expert domain knowledge. Such models have achieved impressive results in many fields such as Natural Language Processing (Collobert et al., 2011) and sequence-to-sequence learning (Sutskever et al., 2014). Deep Learning models also achieved human competing performances in many computer vision tasks such as (Krizhevsky et al., 2012; He et al., 2016) in image classification, (Girshick et al., 2014; Ren et al., 2015) in region-based image segmentation, (Redmon et al., 2016; Liu et al., 2016) in single-pass image segmentation, and in instance segmentation (He et al., 2017). Deep Learning models have also proven to be able to produce novel data, a task which was considered until recently to be restricted to humans, an example of such models are Generative Adversarial Networks (GANs) (Goodfellow et al., 2014) that proved to be successful in synthesizing, among others, images (Isola et al., 2017) and 3D models (Liu et al., 2017).

A key factor to the success of these techniques in a certain domain is the availability of training data. Many portals around the world provide thematic and raw data publicly, for example, as described in (Land NRW (2018)) and Wichmann et al., (2018). There is an opportunity to enrich detailed city models with an automatic privacy patch generator that is highly adaptable by using Deep Learning techniques and publicly open spatial data. We explore this potential within this paper and present the following contributions:

We propose an automatic framework for a patch generator based on a Conditional Generative Adversarial Networks architecture to obfuscate images for privacy protection. In addition, we present special data augmentation and styling modules. The implementation and training process is demonstrated and evaluated on publicly available datasets. In addition, the trained framework is tested on privacy zones that are unseen by the Deep Learning model, namely privacy zones based on urban space partitioning as first proposed by Filippovska et al. (2016).

2. Related Work

2.1 Privacy Protection using GANs

Despite its recent introduction by Goodfellow et al. (2014), it has been shown that GANs have outstanding potential for the purpose of personal information protection. Beaulieu-Jones et al. (2017), e.g., use GANs to synthesize new medical records that have similar properties as the original dataset. They argue that the

generated artificial datasets can be used as a substitute and still derive conclusive medical analyses while preserving patient anonymity. Instead of using pixelization that reduces the visual quality of images, Chen et al. (2018) demonstrate the ability of GANs to generate images that preserve the facial expression of individuals while being identity invariant. Liu et al. (2018) show the potential of GANs to generate fake trajectories that preserve the statistical properties of the training datasets. To prevent the generator of a GAN to learn and memorize sensitive information that could later be extracted, Xie et al. (2018) suggest to add further layers of encryptions in a so-called differentially private GANs and inject noise into the training process.

2.2 GANs for Image Syntheses

Until recently, the generation of realistic looking images with a given data distribution was considered a challenging task. Inspired by game theory, Goodfellow et al. (2014) proved that the adversarial training of two neural networks, the generator and the discriminator, enables the generator to learn this task.

The generator attempts to output fake data points, here images, that ‘fools’ the discriminator to classify them as real. The generator’s input is sampled from a latent space, usually represented as Gaussian noise. During training, the discriminator’s input is either data from the training dataset or the generator’s output. On the other side, the discriminator’s objective is to correctly classify input as real or fake. Both networks improve their abilities through training. In the case that the discriminator and generator have unlimited capabilities they reach the Nash equilibrium (Nash, 1951) of the game. In this case, the discriminator correctly classifies the output of the generator in 50% of the cases. This equilibrium can be interpreted in a way that the generator is unable to completely fool the discriminator while the discriminator is unable to dismiss the generator’s output as fake.

The original GAN incarnations are successful in producing novel data, however, they lack mechanisms that enables to control the generator’s outcome. Mirza and Osindero (2014) introduce the concept of conditional GANs (cGAN) that allows to condition both the generator and the discriminator outputs on additional data input. The cGAN Pix2Pix (Isola et al., 2017) is able to generate from an image of edges a realistic image of the depicted object (e.g. shoes and cats); or transform a cartographic map to a satellite or street view image. Reed et al. (2016) generate images from text using cGANs, which shows the potential of this technology to map between different data domains. Zhu et al. (2017) introduce the concept of unpaired mapping between two datasets and proved the possibility to create a Deep Learning model that translates images from one source domain to another target domain without the need of paired images. Wang et al. (2017) introduce a multi-scale discriminator and generator thereby synthesizing high definition images. Zhang et al. (2017) introduce StackGAN, which progressively generates images from a lower to a higher scale. Ledig et al. (2017) successfully perform super-resolution image synthesis. Xu

and Zhao (2018) show that satellite image spoofing is possible using GANs, thereby raising the awareness against fake image generation that can affect decision making or analysis.

2.3 Privacy Enabled City Models

Within the scope of the VarCity project, Martinovic et al. (2015) suggest using avatars that follow the average statistical behavior of individuals to simulate pedestrians and cars in detailed city data in order to preserve privacy when sharing videos. Filippovska et al. (2016) propose an automatic framework for generating a 2D and 3D space partitioning that can be used to introduce privacy zones in city layouts without ruining the visual quality of ground and façade image maps. Transitioning zones are applied to gradually change the images from photorealistic to abstract representation. Filippovska et al. (2016) produce pleasing results when manually generating a background base map whose colors match those of the corresponding aerial or satellite images. However, this manual step is a bottleneck that limits automation. The goal of the proposed methodology is to train cGANs to generate fake images for these privacy zones to obfuscate privacy relevant areas.

3. Methodology

3.1 Framework

As depicted in Figure 1, we propose a privacy framework that automatically replaces image parts that are subject to privacy or security concerns with generated (fake image patch) fillers. The visual aspects of these fillers should be similar to the ones of the original aerial or satellite images. In other words, the newly generated images follow the statistical properties of the training data distribution.

In addition, to increase the flexibility of the approach, the framework should further offer a mechanism to control the spatial location, extent, and particularly the content of the fillers. Hence, the framework takes the spatial extent of the privacy zones as input as well as a thematic representation of the phenomena that exists within.

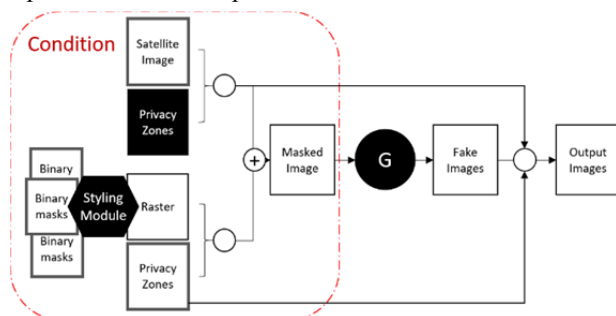


Figure 1. The proposed framework. Original inputs are indicated by the bold grey boundaries. The condition generation process is highlighted by the red dashed line. G is the generator network. Image's pixels that correspond to false values in the privacy masks are retained.

At the core of the framework is the generator, which is a neural network that stochastically outputs images that overlay privacy regions with fake fillers. The weights of the generator are learned in advance through an adversarial training process. This process is explained in the following subsection 3.2. In order to sample pixel

values for the fake fillers, the generator is conditioned for each inference on a different combined image that contains both no-privacy zones (surrounding context areas) as well as privacy zones (thematic representation).

A privacy pixel mask is used to define within the satellite image the areas to be removed. Because the generator network generates a completely new image, the inverted mask is then used to overlay the original image to the non-privacy zones. This procedure is important as non-privacy zones should have no change in the image and thus the pixel values should be kept.

3.1.1 Thematic data

Due to its simplicity, several binary masks are used to represent such phenomena as roads and rooftops. A road mask, e.g., is shown in Figure 2(b). The accuracy of these masks results from the quality of the input thematic data. Missing or erroneous data is expected to have a negative impact on the quality of the final images that the generator network is able to generate.



Figure 2. (a) Input image as orthophoto © Land NRW (2018) with its associated (b) binary road mask. White pixels represent where the value of the mask is equal to one.

Although building roof structures can easily be recognized within the satellite imagery, their associated footprints do not necessarily correlate with their complex roof shapes. For example, a building with a simple rectangular shaped footprint can feature a roof structure of complex shape with many superstructures. This complexity within a roof structure has a huge impact on the illumination of the roof itself where the individual surfaces have different light intensities if they are shaded or not. Shadows, therefore, impact the pixels that represent both the building roofs and their surrounding environment. Therefore, we argue that it is necessary to include the edges of roof structures to condition the generator. An exemplary representation of roof edges is shown in Figure 3(b).

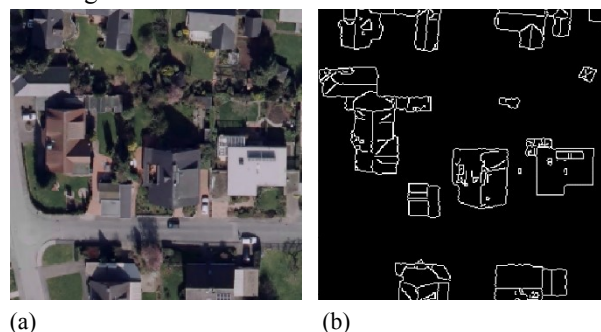


Figure 3. Input image as orthophoto © Land NRW (2018) with its associated (b) binary building mask. Only rooftops segment edges are shown. The white pixels correspond to mask values of one.

3.1.2 Styling module

Data points/images at test time must come from the same data distribution as the training data to ensure a good generalization over the features learned by the generator. This means that the visual appearance of the image on which the framework is conditioned on should be similar to those of the images used for training. In other words, the cartographic representation of spatial data, which are subject to obfuscation, needs to follow the representation used for the training samples.

To endow the framework with capabilities to process different datasets, we design a styling module that converts binary masks to a target cartographic representation by manually selecting suitable colours.

3.1.3 Masked Images

The number of channels of the input data, on which the generator is conditioned, increases the number of parameters in the model and thus the overall computations. Therefore, we reduce the number of input channels in the generator. Instead of having each binary mask as a separate channel, the satellite and thematic raster data are merged. A privacy zone binary mask is applied to the thematic representation to eliminate the surrounding areas. An inverted mask is generated and applied to the satellite image to eliminate those pixels that correspond to the private zones. Both images are then merged and the resulting masked image is then used as a conditional input.

3.2 Training and Data Augmentation Module

The generator is trained in a standard conditional adversarial training process as shown in Figure 4 after which the generator is embedded in the framework.

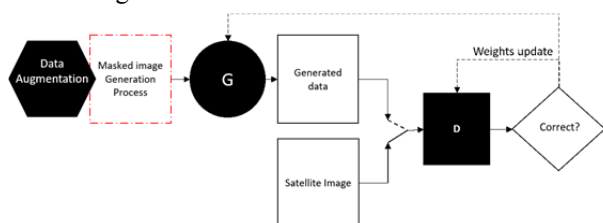


Figure 4. Training process. The dashed red block represents an aggregation of the mask generation process. G and D are the generator and discriminator networks respectively.

It is usually laborious to acquire a sufficient amount of labeled data to successfully train a Deep Learning model. Data augmentation is, therefore, performed to increase data quantity, e.g., by applying different transformations on the training images such as rotation, translation, skewing, and changing the color intensity among other techniques. The aforementioned techniques lead to spatial displacements of the objects within the image. However, since the generated images need to match the appearance of the original training dataset, such changes are undesirable.

To overcome this obstacle, we have implemented a specialized data augmentation module. The module augments training data by randomly producing different privacy zone masks during each training iteration while alternating between three data privacy masking options:

covering the entire image, randomly moving window, and building buffers.

The first option is shown in Figure 5, masking a privacy zone that covers the entire satellite image. We argue that this option will allow the generator to learn a mapping between the thematic representation and the satellite images.



Figure 5. (a) A condition where the privacy area covers the complete span of the input image with its associated (b) ground truth orthophoto © Land NRW (2018)

In the second option, shown in Figure 6, a random privacy window mask with random dimensions and location is created. The masked image will contain a combination of the thematic representation at the pixels that are covered by the mask window and the original satellite pixels at all other locations. We argue that this will increase the amount of data that enables the generator to learn to perform an identity mapping at non-private zones.



Figure 6. (a) A condition where the privacy area corresponds to a randomly initiated window with its associated (b) ground truth orthophoto © Land NRW (2018)

The last possible option, shown in Figure 7, is that the module will generate privacy zones that correspond to the enlarged building footprints as the result of a typical (GIS) buffer operation. Due to the different building footprints, this option gives the generator enough data to learn how to map different shapes of privacy masks.



Figure 7. (a) A condition where the privacy area corresponds to building footprint buffers with its associated (b) ground truth orthophoto © Land NRW (2018)

4. Evaluation

For the evaluation of our proposed methodology, we have chosen the city of Everswinkel in North Rhine-

Westphalia, Germany. Reasons are the on the one side the availability of suitable public data and on the other side the variety of urban patterns that exists within the city as shown in Figure 8. Especially that most of the buildings are freestanding and feature a similar roof appearance served our intended purpose. Training and evaluation locations were selected within the north and the south part of the city respectively.

The urban area featured in the training data Land NRW (2018), includes neighbourhood with an industrial appearance at the western and eastern sides of the city. The middle northern part of the city contains residential buildings in a structured repetitive pattern while the city center has a relatively non-repetitive pattern. Prominently green and rural locations are observed in the northern and north-western part of the city. The test dataset includes images that display a similar variety of urban patterns. However, the training dataset has a higher frequency of red roof structures.



Figure 8. Training area on the northern part of the city and testing on the southern part. The red cells represent the span of training and testing samples © Land NRW (2018).

The implementation of the generator network, introduced in the next subsection, takes input images with dimensions of 256 by 256 pixels. In order to generate meaningful samples at such a low resolution, both training, and testing locations are overlaid and divided with grids 20 by 20 meters. Data samples were generated according to these cells yielding 220 and 20 samples for training and testing respectively.

4.1 Deep Learning Model and Training

Pix2Pix network (Isola et al., 2017) was chosen for this evaluation due to its established success in generating a variety of photorealistic images. Due to the small number of training samples transfer learning and fine-tuning techniques were applied during the training process.

The initial weights of both the generator and the discriminator networks were learned during an adversarial training on images that are scraped from Google Maps by Isola et al. (2017) then fine-tuned on the training dataset. The styling module is adjusted to

produce privacy masks that correspond to the cartographic representation of Google Maps.

During fine-tuning, the data augmentation module was calibrated to generate complete, random window and buffer privacy zones masks with a probability of 40%, 20%, 40% respectively. These numbers were empirically calibrated.

Roof segments are not included in the training dataset. Therefore, we automatically calculated an approximation of these segments. A building footprint mask was used to select rooftop pixels, then the Canny edge detector (Canny, 1986) to calculate corresponding edges. The results are merged back into the building footprint mask and represented with the same color as the building outline. Results are shown in .

4.2 Results and Discussion

After training the generator, it was used to generate a large number of examples for evaluating its potential for the intended purpose. The generator's output is first examined, then results of the proposed framework are shown.

4.2.1 Generator Results

The quality of the generator output is heterogeneous. For example, Figure 9 shows that the generator generated fake images with arguably visually pleasing qualities.

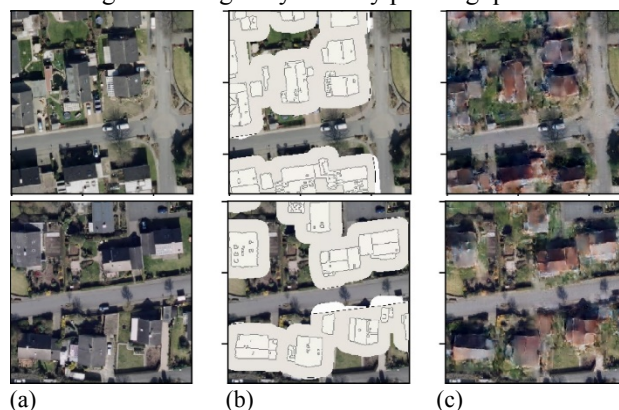


Figure 9. (a). Ground truth image © Land NRW (2018). (b) Generator's input. (c) Fake images produced by the generator.

The results show that the generator is able to produce recognizable parcels and houses. The generator is also able to predict a reasonable light intensity difference that reflects the geometry of rooftops and plausible shadows. The generator is also able to correctly generate continues roads and green patches that visually match their surroundings. The generator was fine-tuned on a training dataset which contained predominantly red rooftops and its output is affected by this setup and therefore more likely to produce roof with a red texture.

Figure 10 shows that the generator produced rooftops instances with an interrupted representation at the borders of the privacy zone mask. We argue that this due to the fact that constraints such as identifying object instances are unenclosed in the model.



Figure 10. (a). Ground truth image © Land NRW (2018). (b) Generator's input. (c) Fake images produced by the generator.

The top example in Figure 11 shows that green areas resembling trees were generated on top roads. This area is a non-privacy zone. Original pixels are included in the generator's condition. A simple identity multiplication would have produced the original pixels. In the second example, the generator outputs a dark patch and ignored the condition. These examples demonstrate the stochastic nature of the generation process.

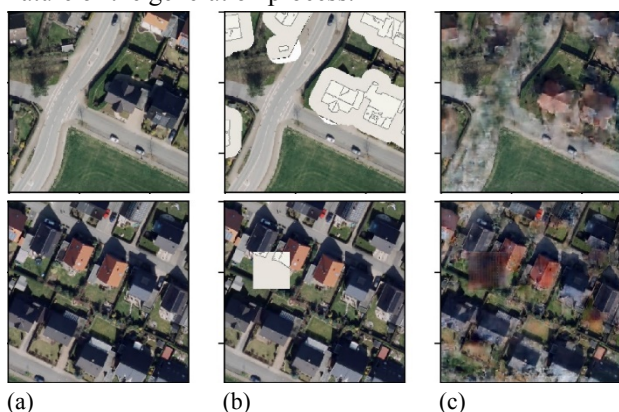


Figure 11. (a). Ground truth image © Land NRW (2018). (b) Generator's input. (c) Fake images produced by the generator.

4.2.2 Interactive Alteration of Images

An interactive session is conducted to demonstrate the framework's potential in generating altered data in real time. The privacy mask is interactively altered and then passed through the generator to make predictions after each change. Figure 12 shows the starting privacy mask that is produced from the data augmentation module.

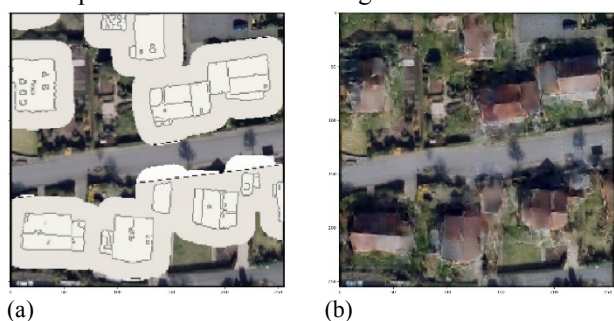


Figure 12. (a) Privacy labels generated by the data augmentation module. (b) Fake images produced by the generator.

In Figure 13 the mask is altered to reflect a privacy zone that covers only buildings which belong to the northern block. The remainder of the original pixels is included as non-privacy zones. Roof predictions are all red even though the output of the generator was conditioned on a non-privacy zone that contains buildings with grey rooftops.



Figure 13. (a) Privacy labels generated interactively. (b) Fake images produced by the generator.

Figure 14 shows the deletion of all buildings in the southern block from the original mask. The lack of building footprints in this area gives a cartographic representation of a rural or agricultural area. This is reflected in the generator prediction as it filled the privacy zone with dense green plants.



Figure 14. (a) Privacy labels generated interactively. (b) Fake images produced by the generator.

Figure 15 shows the addition of multiple buildings in the southern block. The generator's prediction reflects that by generating the corresponding fake houses in the given locations.

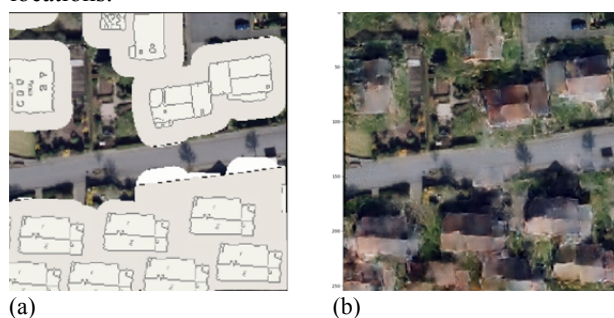


Figure 15. (a) Privacy labels generated interactively. (b) Fake images produced by the generator.

4.2.3 Framework Results

The framework last processing step is to replace pixels from the generator's results at a non-privacy location with the original image pixel values. Figure 16 shows results that are heterogeneous in quality due to the generator's stochastic nature.

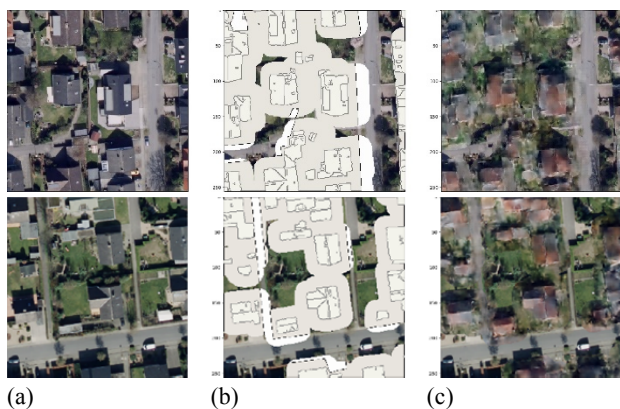


Figure 16. (a) Ground truth image © Land NRW (2018). (b) Generator's input. (c) Fake images generated by the framework.

4.2.4 Results on Space Partitioning Privacy Zones

Space partitioning privacy zones are not included as a part of the generator's training data. However, as shown in Figure 17 the generator was able to produce arguably plausible results with such boundaries.



Figure 17. (a) Ground truth image © Land NRW (2018). (b) Privacy labels generated with space partitioning methods (Filippovska et al., 2016). (c) Fake images generated by the framework.

5. Conclusion and Outlook

In this research, a novel and automatic framework for privacy patches generation is introduced. The framework can generate privacy patches that are conditioned and harmonized to the surrounding images. In addition, we designed a special data augmentation module that can increase the amount of training data and a styling module that facilitate framework adaptation on different datasets. Our evaluation demonstrates that the framework can generate arguably promising results when considering the aesthetic quality of the resulting images.

In the evaluation, the detailed roof segments were generated using image processing techniques as presented in (Canny, 1986). Acquiring an actual 3D detailed segmentation of the roof may improve the quality of the results. Further research is needed to test this hypothesis.

In our evaluation, images were taken in relatively uniform light conditions and time of the year. A generator which is trained on such a dataset will only produce images with a similar visual aspect as the training dataset. Further evaluation is needed to estimate the generator's performance and the possibility of conditioning in different locations and times of the day and year. Furthermore, according to the intended use, user cognition tests can evaluate the quality of the framework results.

The dataset used in training at inference time had a unified scale. Results predictions were also limited to cell boundaries. The image generation is a stochastic process and there is no guarantee to the continuity of the same feature instance. A progressive generation of images in a hierarchical manner might help overcome this obstacle. Further research is needed to investigate this assumption.

The network used in our evaluation is limited in that it produces low-resolution images. Multiple state-of-the-art cGAN can produce high definition results, e.g., (Zhang et al., 2017). Further research is needed to evaluate this potential.

Remote sensing images are an important information source. Many amateurs and professionals depend on such data to acquire situational awareness, to perform spatial analysis and draw conclusions that can affect decision making. Generating fake data without informing users can be misleading and even dangerous. With the advancement of image generation technology, it would be more difficult to recognize fake data and therefore there is a responsibility to inform users through cartographic techniques about the authenticity of the data.

Within a relatively short time, Generative Adversarial Networks have achieved many remarkable milestones. It is a promising technique that can be used in solving many cartographic and remote sensing problems such as privacy enablement for city models.

6. Acknowledgment

We are grateful to the state of North Rhine-Westphalia, Germany, for making the data used in this study publicly available. The source and license of the orthophotos used in this research are: Land NRW (2018) Datenlizenz Deutschland - Namensnennung - Version 2.0 (www.govdata.de/dl-de/by-2-0).

This work was supported by the German Research Foundation (DFG) [grant number KA 4027/1-1].

7. References

- Beaulieu-Jones, B. K., Wu, Z. S., Williams, C., & Greene, C. S. (2017). Privacy-preserving generative deep neural networks support clinical data sharing. *bioRxiv*, 159756.
- Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, (6), 679-698.
- Chen, J., Konrad, J., & Ishwar, P. (2018). VGAN-Based Image Representation Learning for Privacy-Preserving Facial Expression Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 1570-1579).
- Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., & Kuksa, P. (2011). Natural language processing (almost) from scratch. *Journal of Machine Learning Research*, 12(Aug), (pp. 2493-2537).
- Filippovska, Y., Wichmann, A., & Kada, M. (2016). Space Partitioning for Privacy Enabled 3D City Models. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 42.

- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 580-587).
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems* (pp. 2672-2680).
- He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017, October). Mask R-CNN. *Computer Vision (ICCV), 2017 IEEE International Conference on* (pp. 2980-2988).
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-Image Translation with Conditional Adversarial Networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 5967-5976).
- Kada, M., Peter, M., Fritsch, D., Siemoneit, O., & Hubig, C. (2009, November). Privacy-enabling abstraction and obfuscation techniques for 3D city models. *Proceedings of the 2nd SIGSPATIAL ACM GIS 2009 International Workshop on Security and Privacy in GIS and LBS* (pp. 53-57).
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems* (pp. 1097-1105).
- LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), (pp. 2278-2324).
- Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., & Shi, W. (2017). Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. *CVPR* (Vol. 2, No. 3, p. 4).
- Liu, J., Yu, F., & Funkhouser, T. (2017). Interactive 3D Modeling with a Generative Adversarial Network. *2017 International Conference on 3D Vision (3DV)* (pp. 126-134).
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single shot multibox detector. *European conference on computer vision* (pp. 21-37). Springer, Cham.
- Liu, X., Chen, H., & Andris, C (2018). trajGANs: Using generative adversarial networks for geo-privacy protection of trajectory data (Vision paper).
- Martinovic, A., Knopp, J., Riemenschneider, H., & Van Gool, L. (2015). 3d all the way: Semantic segmentation of urban scenes from start to end in 3d. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4456-4465).
- Mirza, M., & Osindero, S. (2014). Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*.
- Paderta, D. (2012). Open Data - Raumbezogene Daten. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-364743>
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).
- Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., & Lee, H. (2016). Generative Adversarial Text to Image Synthesis. *33rd International Conference on Machine Learning* (pp. 1060-1069).
- Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems* (pp. 91-99).
- Siemoneit, O., Hubig, C., Kada, M., Peter, M., & Fritsch, D. (2009). Google Street View and Privacy. *Proceedings of the Fifth Asia-Pacific Computing and Philosophy Conference*.
- Stagars, M. (2016). Open Data in Southeast Asia: Towards Economic Prosperity, Government Transparency, and Citizen Participation in the ASEAN. Springer.
- Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. *Advances in neural information processing systems* (pp. 3104-3112).
- Wang, T. C., Liu, M. Y., Zhu, J. Y., Tao, A., Kautz, J., & Catanzaro, B. (2017). High-resolution image synthesis and semantic manipulation with conditional gans. *arXiv preprint arXiv:1711.11585*.
- Wichmann, A., Agoub, A., & Kada, M. (2018). ROOFN3D: Deep Learning Training Data for 3D Building Reconstruction. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 42(2).
- Xie, L., Lin, K., Wang, S., Wang, F., & Zhou, J. (2018). Differentially Private Generative Adversarial Network. *arXiv preprint*.
- Xu, C., & Zhao, B. (2018). Satellite Image Spoofing: Creating Remote Sensing Dataset with Generative Adversarial Networks (Short Paper). *LIPICs-Leibniz International Proceedings in Informatics* (Vol. 114). Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.
- Zhang, H., Xu, T., & Li, H. (2017). StackGAN: Text to Photo-Realistic Image Synthesis with Stacked Generative Adversarial Networks. *2017 IEEE International Conference on Computer Vision (ICCV)* (pp. 5908-5916).
- Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. *IEEE International Conference on Computer Vision*.