Data Augmentation for Improved Morphology on Mars Data

Douglas D. Agbeve
agbeve01@ads.uni-passau.de
University of Passau

Aditya V. Handrale
handra01@ads.uni-passau.de
University of Passau

Salim Fares
fares01@ads.uni-passau.de
University of Passau

Seif E. Idani
idani01@ads.uni-passau.de
University of Passau

1 INTRODUCTION

A picture, they say, is worth a thousand words. As a medium of communication, pictures have been used since the dawn of *Homo Sapiens*[1], which makes understanding the information embed in them a major part of research in the various fields of science ranging from Archaeology to the relatively more recent Computer Science.

In Computer Vision, a vital process in retrieving information in images and which form part of most systems for visual understanding is image segmentation[2]. It is the process of partitioning an image or a video frame into multiple segments[3]. Segments are fundamentally a set of pixels in a region that share a common property such as colour or texture, enabling the identification and locating of boundaries of objects in an image[4]. Digital image segmentation is applied in various domains, including but not limited to, medical (locating cancerous tumors and measuring tissue volumes)[5], facial and fingerprint recognition, self-driving vehicles (detecting pedestrian). There are and continue to be proposed, many algorithms for image segmentation. Currently, these approaches are either segregating pixels based on intensity changes (i.e. detecting discontinuity) such as edge detecting algorithms, or partitioning an image into regions with similar predefined criteria (i.e.Similarity detection) examples include thresholding approaches. The various algorithms can also be categorized based on the method used in solving the problem; namely (1) Edge detection based segmentation e.g. histogram-based[6] and gradient-based[7], (2) Thresholding approaches such as those proposed by Abd Elaziz *et al*[8], and Houssein *et al*[9], (3) Region-based segmentation techniques include region growing[10, 11], region split and merge[12, 13], (4) techniques based on Partial Differential Equation are active contour[14], C-V model[15] etc., (5) Clustering methods such as K-means clustering algorithm[16] and relatively more recent approach is the Artificial Neural Network (ANN) algorithms which solves image segmentation problems with higher accuracy compared to other approaches. Examples of Deep learning or ANN approach to image segmentation include but not limited to Convolutional Network models such as VGG16, GoogLeNet, Fast R-CNN, Faster R-CNN[17], U-NET[18], V-NET[19] and Mask-RCNN[20], Recurrent Neural Network models include ReSeg[21],Generative and Adversarial (GAN) Models e.g. [22, 23, 24], a comprehensive list of GAN-based techniques can be found at [25].

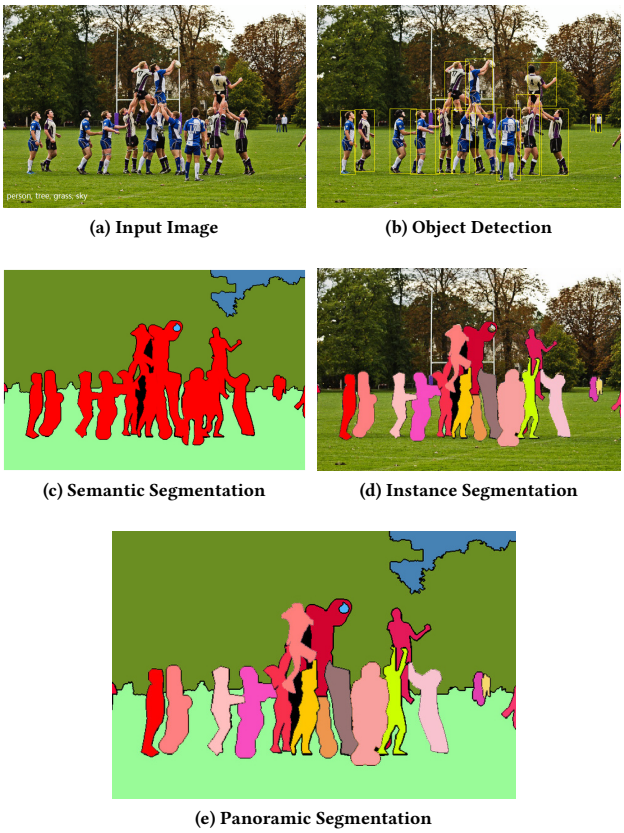
The problem of image segmentation can be coined as that of classifying pixels with semantic labels i.e. semantic segmentation (fig. 1c) or segregating pixels into individual objects in the image i.e. instance segmentation (fig 1d)[26]. In semantic segmentation, pixels are assigned to the same segment if they are of the same object type in the image. This can be thought of as classification at pixel level. Instance segmentation involves assigning all the pixels that belong to the same single object to the same segment. It is a two-step process; extracting bounding boxes around each instance of an object through object detection and then classifying pixels that corresponding to each instance in the bounding box. This technique combines object detection (fig. 1b) and segmentation. A third, relatively new, formulation is to combine both instance and semantic segmentation termed Panoramic Segmentation (fig. 1e). It involves the detection and segmentation of all objects including background and labelling different instances in an image.

Digital Elevation Models (DEMs) are essential in analyzing erosion and drainage, hill-slope hydrology, studying groundwater flow, watersheds, and contaminant transportation as they are important tools for parameterizing topography. DEM is representation of planetary (earth, moon, mars etc.) terrain from elevation data in 3-D image. There are two types; raster Geographic Information Systems (GIS) layer representation and the vector-based triangular irregular network format. DEMs are obtained using techniques such as Surveying, Stereo Photogrammetry, Lidar, Radar, etc.. Flight and Train simulations, GIS and Satellite navigations are some of the systems in which DEMs are used. Martian surface has an abundance of geographical features such as volcanoes, layers and gullies, with phenomena like volcanoes bringing out microbial entities that were protected from solar radiation. Moreover, glaciers and mounds that have water underneath them might contain evidence of life. Detecting and studying these morphologies can help us find extraterrestrial life on mars. The detection of mounds formed from phenomena such as volcanic eruption can be done automatically using ANN based image segmentation methods.

1.1 Research Questions

The many algorithms available for solving image segmentation problem is an indication that no single approach can solve it all accurately and in a more efficient way. An affirmation of this assertion is the numerous methodologies proposed in literature by Deep Learning researchers in recent decade. Furthermore, the choice of algorithm and by extension the need for new approaches, is principally influenced by data pertaining to the problem, its representation and the type of segmentation problem. In view of these, we formulated the following questions with regard to the data at hand.

Data Science Lab, 2021/2022, Uni Passau
2021/2022.



(a) Input Image (b) Object Detection

(c) Semantic Segmentation (d) Instance Segmentation

(e) Panoramic Segmentation

Figure 1: Variations of Segmentation [27]

- Will data augmentation have any effect on the accuracy of our selected approach (es) for application, and if yes, how much extra data relative to the training data?
- Which of the Machine Learning, more specifically Deep learning, techniques, segments the data with much more accuracy?
- Most of the Deep Learning algorithms in literature are categorize under supervised learning. Can this problem be solved with unsupervised learning techniques?

1.2 Related Work

Wuff-Jensen *et al.* proposed, to the best of our knowledge, the original idea of using GANs to generate terrain from DEMs. Their architecture was based on deep convolutional GAN (DCGAN)[28] – a type of GAN made up of a fractional-strided convolutions (generator) and a discriminator of strided convolutions.

Spick *et al.*[29] presented a method of generating height maps from digital elevation of regions of earth. The authors approach was based on Spatial GANs (SGAN)[30]. Spatial GANs remove fully connected layers from DCGAN allowing outputs to be scaled to any size and mapping features onto output as translation-invariant.

Bowles *et al.*[31] investigated augmenting training data using GAN-derived synthetic images. Progressive Growing of GANs (PGGAN)[32] network was used as a baseline architecture, and demonstrated that this can improve results across two segmentation tasks (1) - Computed Tomography (CT) images with manually delineated Cerebrospinal Fluid (CSF) labels, (2) - Fluid-Attenuated Inversion Recovery (FLAIR) images with manual White Matter Hyperintensity (WMH) segmentations.

2 PROBLEM STATEMENT

In this section, we formally present the problem of automatically detecting mounds in Mars' Arabia Terra. Specifically, classification of pixels in the digital elevation model of Mars in to segments of mounds or no-mounds using using Artificial Neural Network based techniques. We sought to predict where mounds given any DEM of Mars

More formally, given n annotated images $\{x_1, x_2, \ldots, x_n\}$ and each image, say x_i, has m_i objects that are categorize into C classes with objects labelled as y_i.

$$y^i = \{(c_1^i, v_1^i), (c_2^i, v_2^i, \ldots, (c_{m_i}^i, v_{m_i}^i))\} \tag{1}$$

where $c_{m_i}^i \in C$ and $v_{m_i}^i$ is the object's pixel mask.
The goal is to learn the values of some parameterized (θ) function f such that:

$$y^i = f(x_i, \theta_i) \tag{2}$$

And be able to predict object (mound) locations $\hat{y}$ for all new (unseen) x_i
A loss function to optimize the prediction would be:

$$J(\theta) = \frac{1}{n} \sum_{i=1}^{n} l(\hat{y}^i, x_i, y^i; \theta) + \frac{\lambda}{2} \|\theta\|_2^2 \tag{3}$$

Due to the relatively small number of training samples, we proposed, as a starting point of our implementation, inspired by the work done by Souly *et al.*[23], an architecture with a generator to provide extra training samples and classifier as a discriminator in a Generative Adversarial Network. We evaluate the performance of each approach against the following metrics *Pixels Accuracy* i.e. ratio of properly classified pixels to total number of pixels, *Mean Pixel Accuracy (MPA)* is the average of the ratio of properly classified pixels to the number of pixels in a class and *Mean Intersection over Union (mIoU)* is the ratio of the intersection between the ground truth and the predicted segmentation map to their union, averaged over all classes.

3 DATA ACQUISITION & PRE-PROCESSING

3.1 Data acquisition

In our paper, we are using data from HiRise (High Resolution Imaging Science Experiment) camera, onboard the Mars Reconnaissance Orbiter (MRO) spacecraft. We have the DEM (Digital elevation model) of the 'Firsoff' crater, which is an impact crater on mars. Two high resolution images of a specific area on the ground are taken from different camera angles, and then these stereo images are combined together to obtain a DTM (Digital Terrain Model). The photos were taken at an altitude of 272 kilometers above ground. The scale of the image is very high in resolution at 0.27 meter per pixel. This data has been made publicly available by NASA/JPL/

University of Arizona. We decided to utilize a DEM format, since it is a three dimensional digital representation of a terrain with X,Y and Z coordinates. The Z coordinate (Elevation data) will help us find natural morphologies on the surface like mounds, craters and channels. We can use DEMs to generate additional features by performing slope and aspect analysis. The use of DEMs also allows us to build 3D models of the surface. Moreover the distance between the sample points (spatial resolution) and the vertical resolution are very high in the HiRise DEM.

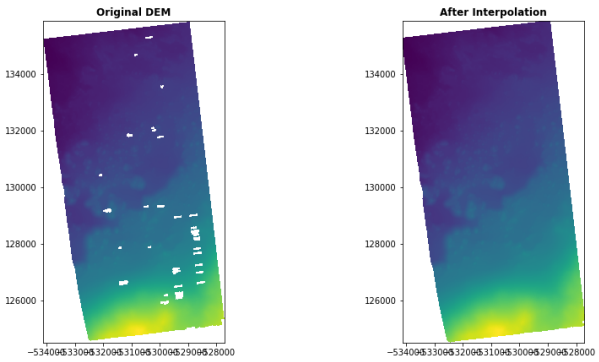
3.2 Data preprocessing

The first step in preprocessing was to fill in the No Data values. A no data value is present in the dataset when there is no depth reading available at that coordinate. This usually represents things such as backgrounds and borders, but also occurs when data is not available due to technical difficulties etc.

There are a few methods that can be used to treat the NoData values. We chose to perform automatic interpolation using the 'fillnodata' method from the python package 'rasterio'[1]. Manual interpolation is also possible, but requires expertise in domain. Sometimes interpolation can also cause geometric patterns and artifacts in the image which end up adding noise to our dataset.

The rasterio interpolation we used has two tunable parameters, Max_search_distance and smoothing_iterations. Max_search_distance corresponds to the maximum number of pixels to search in all directions to find values to interpolate from (using inverse distance weighting). The smoothing_iterations refers to the number of 3x3 average filters (passes to run on interpolated pixels) are applied to smooth out artifacts. We used a value of 35 for the max_search_distance and a value of 0 for the smoothing iterations.

After interpolating the missing values, we obtain a proper DEM with no missing values. The DEM file is now ready for the tiling process.



(a) Interpolation

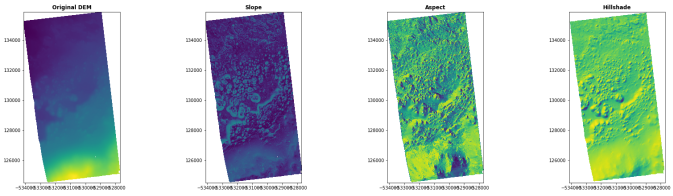
Figure 2: Filling NoData Values Using Interpolation

[1]https://rasterio.readthedocs.io/en/latest/api/rasterio.fill.html

The DEM image we are using has a resolution of 6418 x 11339, and is a geoTIFF file. Each pixel corresponds to the depth measurement at a specific area on the ground. The maximum depth value is -1595.222, maximum value is -3050.866 and the No Data value is represented by -32767. The DEM has 75% valid values and a mean of -2594.830. A standard deviation value of 336.037 was observed.

3.3 Feature engineering

Adding channels or features to our data will help the segmentation neural network detect morphologies better. We can make use of additional features such as hillshades, slope and aspect analysis etc. We decided to generate hillshade, slope and aspect from our DEM file as additional features using 'gdal.DEMProcessing' [2].



(a) Feature Extraction

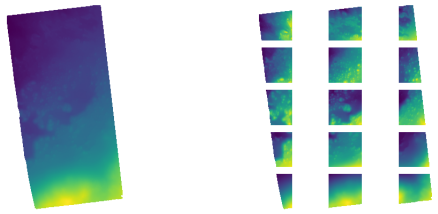
Figure 3: Variations of Features

Since the dataset is very small and consists of only one image, we decided to tile the image into XDIV x YDIV parts. Tiling will allow us to have multiple training samples, which in turn will improve the performance of our neural network. All of the morphologies to be detected have a polynomial, amoeba-like shape. There arise a few problems when using tiling, ex. A tile border can cut through a morphology, giving it a linear edge where there is none. Since we will also be tiling the testing data set, this should not be a problem.

Tiling Pseudocode

- Find the top-left corner of the image which has the lowest x-coordinate and the highest y-coordinate using gdal.GetGeoTransform. (xmin, ymax) = (gdal.GetGeoTransform()[0],gdal.GetGeoTransform()[3])
- Get the size of the raster image using which is the length of the coordination axis multiplied by the respected pixel size:
  - xlen = xres * gdal.RasterXSize; xres = gdal.GetGeoTransform()[1]
  - ylen = yres * gdal.RasterYSize; yres = gdal.GetGeoTransform()[5]
- Define the number of divisions on each axis. (xdiv, ydiv)
- The size of your tile would be:
  - xsize = xlen/xdiv
  - ysize = ylen/ydiv
- Define the strides(steps) of tiling
  - xsteps = [xmin + xsize * i for i in range(xdiv+1)]
  - ysteps = [ymax - ysize * i for i in range(ydiv+1)]
- Finally using gdal.Warp for each step in the steps vectors, pass:
  - The wanted name of the output.
  - The input DEM.
  - outputBounds=(xsteps[i], ysteps[i], xsteps[i+1], ysteps[i+1]

[2]https://gdal.org/python/osgeo.gdal-module.html#DEMProcessing

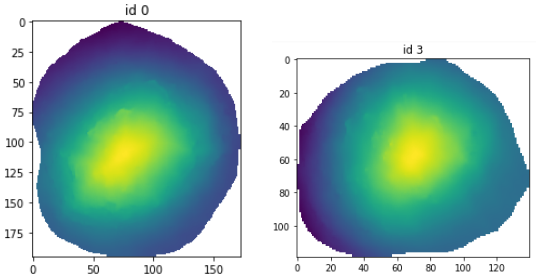


(a) Original DEM (b) Tiles

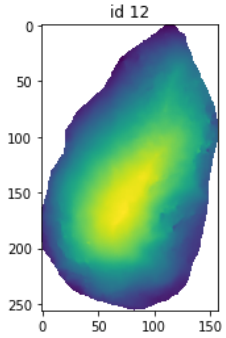
Figure 4: Tiling

3.4 Annotation

To annotate our dataset we need to extract the mounds from our DEM using the shape file that describe the mounds. We use the geopandas.read_file()[3] to read the shape file. We reproject the labels coordinate system to that of the original DEM which is stored in the 'meta' field of the DEM data.



(a) Mound 1 (b) Mound 2

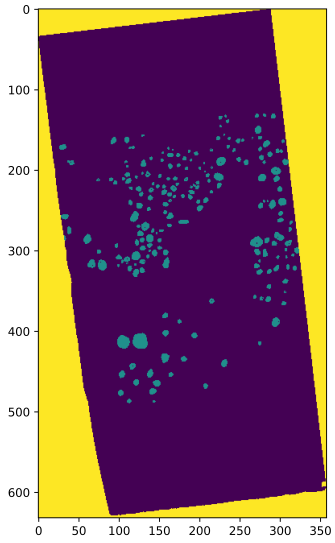


(c) Mound 3

Figure 5: Mounds

We mask the labels using rasterio.mask.mask()[4], which takes the dataset, the geometry of the mounds (stored in the shape file) and the noDataValue of the DEM.

[3]https://geopandas.org/en/stable/docs/reference/api/geopandas.read_file.html
[4]https://rasterio.readthedocs.io/en/latest/api/rasterio.mask.html



(a) Annotated Image

Figure 6: Annotated Image

The annotations would be:

- '0' For non-mounds points.
- '1' For mounds points.
- '2' For the invalid data.

REFERENCES

[1] Yuval Noah Harari. 2015. *Sapiens: A Brief History of Humankind*. Harper, 195 Broadway New York, NY 10007 USA.

[2] David A. Forsyth and Jean Ponce. 2012. *Computer Vision - A Modern Approach, Second Edition*. Pitman, Hoboken, New Jersey, 1–91. ISBN: 978-0-273-76414-4.

[3] Ying Tan. 2016. Chapter 11 - applications. In *Gpu-Based Parallel Implementation of Swarm Intelligence Algorithms*. Ying Tan, editor. Morgan Kaufmann, San Francisco, CA, 167–177. ISBN: 978-0-12-809362-7. DOI: https://doi.org/10.1016/B978-0-12-809362-7.50011-X.

[4] Rajeshwar Dass and Swapna Devi. 2012. Image segmentation techniques. *International Journal of Electronics & Communication Technology*, 3, 1. ISSN: 2230-7109 (Online).

[5] Dzung L. Pham, Chenyang Xu, and Jerry L. Prince. 2000. Current methods in medical image segmentation. *Annual Review of Biomedical Engineering*, 2, 1, 315–337. PMID: 11701515. DOI: 10.1146/annurev.bioeng.2.1.315. https://doi.org/10.1146/annurev.bioeng.2.1.315.

[6] Ming Zeng, Youfu Li, Qinghao Meng, Ting Yang, and Jian Liu. 2012. Improving histogram-based image contrast enhancement using gray-level information histogram with application to x-ray images. *Optik*, 123, 6, 511–520. ISSN: 0030-4026. DOI: https://doi.org/10.1016/j.ijleo.2011.05.017.

[7] Jamil A. M. Saif, Mahgoub H. Hammad, and Ibrahim A. A. Alqubati. 2016. Gradient based image edge detection. *IACSIT International Journal of Engineering and Technology*, 8, 3. DOI: 10.7763/IJET.2016.V8.876.

[8] Mohamed Abd Elaziz, Siddhartha Bhattacharyya, and Songfeng Lu. 2019. Swarm selection method for multilevel thresholding image segmentation. *Expert Systems with Applications*, 138, 112818. ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2019.07.035.

[9] Essam H. Houssein, Marwa M. Emam, and Abdelmgeid A. Ali. 2021. An efficient multilevel thresholding segmentation method for thermography breast cancer imaging based on improved chimp optimization algorithm. *Expert Systems with Applications*, 185, 115651. ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2021.115651.

[10] Zhuang Cheng and Jianfeng Wang. 2020. Improved region growing method for image segmentation of three-phase materials. *Powder Technology*, 368, 80–89. ISSN: 0032-5910. DOI: https://doi.org/10.1016/j.powtec.2020.04.032.

[11] Nagaraj Jothiaruna, Joseph K. Abraham Sundar, and Balasubramanian Karthikeyan. 2019. A segmentation method for disease spot images incorporating chrominance in comprehensive color feature and region growing. *Computers and Electronics in Agriculture*, 165, 104934. ISSN: 0168-1699. DOI: https://doi.org/10.1016/j.compag.2019.104934.

[12] Marie Lachaize, Sylvie Le Hégarat-Mascle, Emanuel Aldea, Aude Maitrot, and Roger Reynaud. 2018. Evidential split-and-merge: application to object-based image analysis. *International Journal of Approximate Reasoning*, 103, 303–319. ISSN: 0888-613X. DOI: https://doi.org/10.1016/j.ijar.2018.10.008.

[13] Lifeng Liu and Stan Sclaroff. 2004. Deformable model-guided region split and merge of image regions. *Image and Vision Computing*, 22, 4, 343–354. ISSN: 0262-8856. DOI: https://doi.org/10.1016/j.imavis.2003.11.006.

[14] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. 1988. Snakes: active contour models. *nternational Journal of Computer Vision*, 1, 321–331. DOI: https://doi.org/10.1007/BF00133570.

[15] Xianghai Wang, Yu Wan, Rui Li, Jinling Wang, and Lingling Fang. 2016. A multi-object image segmentation c–v model based on region division and gradient guide. *Journal of Visual Communication and Image Representation*, 39, 100–106. ISSN: 1047-3203. DOI: https://doi.org/10.1016/j.jvcir.2016.05.011.

[16] Man Yan, Jianyong Cai, Jiexing Gao, and Lili Luo. 2012. K-means cluster algorithm based on color image enhancement for cell segmentation. In *2012 5th International Conference on BioMedical Engineering and Informatics*, 295–299. DOI: 10.1109/BMEI.2012.6513157.

[17] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. 2016. Faster r-cnn: towards real-time object detection with region proposal networks. (2016). arXiv: 1506.01497 [cs.CV].

[18] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: convolutional networks for biomedical image segmentation. (2015). arXiv: 1505.04597 [cs.CV].

[19] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. 2016. V-net: fully convolutional neural networks for volumetric medical image segmentation. (2016). arXiv: 1606.04797 [cs.CV].

[20] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. 2018. Mask r-cnn. (2018). arXiv: 1703.06870 [cs.CV].

[21] Francesco Visin, Marco Ciccone, Adriana Romero, Kyle Kastner, Kyunghyun Cho, Yoshua Bengio, Matteo Matteucci, and Aaron Courville. 2016. Reseg: a recurrent neural network-based model for semantic segmentation. (2016). arXiv: 1511.07053 [cs.CV].

[22] Pauline Luc, Camille Couprie, Soumith Chintala, and Jakob Verbeek. 2016. Semantic segmentation using adversarial networks. (2016). arXiv: 1611.08408 [cs.CV].

[23] Nasim Souly, Concetto Spampinato, and Mubarak Shah. 2017. Semi supervised semantic segmentation using generative adversarial network. In *2017 IEEE International Conference on Computer Vision (ICCV)*, 5689–5697. DOI: 10.1109/ICCV.2017.606.

[24] Wei-Chih Hung, Yi-Hsuan Tsai, Yan-Ting Liou, Yen-Yu Lin, and Ming-Hsuan Yang. 2018. Adversarial learning for semi-supervised semantic segmentation. (2018). arXiv: 1802.07934 [cs.CV].

[25] [n. d.] https://github.com/hindupuravinash/the-gan-zoo. accessed: 18-11-2021.

[26] Suman Paneru and Idris Jeelani. 2021. Computer vision applications in construction: current state, opportunities & challenges. *Automation in Construction*, 132, 103940. ISSN: 0926-5805. DOI: https://doi.org/10.1016/j.autcon.2021.103940.

[27] BigdataAILab. 2021. What is semantic segmentation, instance segmentation, panoramic segmentation? (April 2021). https : / / becominghuman . ai / what - is - semantic - segmentation-instance-segmentation-panoramic-segmentation-3bbb03856c12. accessed: 18-11-2021.

[28] Alec Radford, Luke Metz, and Soumith Chintala. 2016. Unsupervised representation learning with deep convolutional generative adversarial networks. (2016). arXiv: 1511.06434 [cs.LG].

[29] Ryan J. Spick, Peter Cowling, and James Alfred Walker. 2019. Procedural generation using spatial gans for region-specific learning of elevation data. In *2019 IEEE Conference on Games (CoG)*, 1–8. DOI: 10.1109/CIG.2019.8848120.

[30] Nikolay Jetchev, Urs Bergmann, and Roland Vollgraf. 2017. Texture synthesis with spatial generative adversarial networks. (2017). arXiv: 1611.08207 [cs.CV].

[31] Christopher Bowles, Liang Chen, Ricardo Guerrero, Paul Bentley, Roger Gunn, Alexander Hammers, David Alexander Dickie, Maria Valdés Hernández, Joanna Wardlaw, and Daniel Rueckert. 2018. Gan augmentation: augmenting training data using generative adversarial networks. (2018). arXiv: 1810.10863 [cs.CV].

[32] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. 2018. Progressive growing of gans for improved quality, stability, and variation. (2018). arXiv: 1710.10196 [cs.NE].

A PHASE CONTRIBUTION

- Phase One Douglas D. Agbeve
- Phase Two Salim Fares