3D Reconstruction of Satellite Data - Survey

Miss. Jewela Costa
Computer Department
SRIEIT, Goa University
Shiroda, India

Miss. Stuti Tiwari
Computer Department
SRIEIT, Goa University
Shiroda, India

Mr. Krishna Bhagat
Computer Department
SRIEIT, Goa University
Shiroda, India

Mr. Akash Verlekar
Computer Department
SRIEIT, Goa University
Shiroda, India

Mr. K. M. Chaman Kumar
Computer Department
SRIEIT, Goa University
Shiroda, India

Dr. Shailendra Aswale
Computer Department
SRIEIT, Goa University
Shiroda, India

***Abstract*—3D Reconstruction has been a field of interest for multiple disciplines, and in the past decade, many researchers have devoted their studies to improve on state-of-art automated methods used for 3D Reconstruction. 3D models have their application in solving numerous visualization problems and a large number of undertakings beyond visualization. In this paper, we conduct a short survey of research in 3D Reconstruction of Satellite data, and finally, we propose a workflow that will provide a direction for future researchers in generation 3D models of Satellite Data using Deep learning techniques. The workflow includes the use of CNN for object segmentation and use of GAN for DSM or height map construction and 3D model generation**

Keywords*— *3D Reconstruction, CNN, GAN, DSM.

I. INTRODUCTION

3D City models with buildings being its prominent feature have vast use cases[1] such as Visibility Analysis, Estimation of Shadows Cast by Urban Features, Visualization for Navigation, Urban Planning, Forecasting Seismic Damage, Flooding, Change Detection, Forest Management, Archeology, etc. in which visualization plays an important role makes 3D reconstruction of city models an essential Task.

Since the increased enhancement in remote sensing techniques[21] the availability of High-quality satellite and aerial images has increased. The intense research in remote sensing using Deep learning has automated [23] the process as well as retained if not further increased the quality of the results. The state-of-the-art methods for remote sensing and photogrammetry were semi-automated and thus required manual human intervention.

The general workflow observed in many of the previous research for 3D city model generation includes 1)Data Acquisition 2) Height Map Construction 3) Object Detection 4) 3D reconstruction 5) LOD enhancement. Most of the stages of development rely heavily on algorithms and most certainly require human intervention. There also has been research in with 3D models can be generated without the construction of DSMs using shallow classifiers and predictors [11]. However, most of the 3D reconstruction pipelines heavily rely on Height maps to achieve higher accuracy [5,6,12,14,17,28].

The Data Acquisition stage usually depends on use-cases which are the following 1)Consistent with satellite data 2) Consistent with urban scenes.

The data used in the first type is for sure High spectral Satellite or Aerial Images. The priority here is to find out ways to enhance the workflow and determine how the availability of satellite images can also facilitate the generation of High-quality 3D models. Here Data acquisition is usually followed by DSM construction and then 3D city modelling.

The latter deals with generating high-quality 3D urban scenes. thus, the data along with aerial images use street photographs to get the facades of the buildings and other street-level data. Here the priority is virtual city modelling for game simulations or other such use-cases where high-quality city scenes are used. In such cases, 2D height maps are usually constructed of the target city as the second stage of the work-flow and then the further 3D model generation is done

The primary focus in both of these use cases is to make the process less reliable on humans and also be less computationally expensive. The development of deep learning techniques using this method has an advantage over the others in terms of quality and efficiency. Each of these stages in the workflow has been more or less enhanced using deep learning.

In this paper, we review each of the above-mentioned stages and how previous works in deep learning have contributed to their development. We primarily focus on remote sensing and usage of Satellite and Aerial Images. However, at the end of section 2, we mention a few of the usage of other kinds of data for 3D city modelling. We have not gone through every possible research paper instead, we try to provide an overview of the research done in each of the stages using deep learning techniques. Section II reviews previous work done in each of the stages and section III proposes a general work-flow that can be adopted for further research.

II. LITERATURE REVIEW

We approach this section-wise. Section 1 deals with the initial stage of data acquisition through remote sensing. Section 2 deals with techniques used for the development of Height maps Or DSMs. Section 3 inquires methods developed for large-scale Semantic labeling. Section 4 provides an overview regarding different approaches used to combine data of height maps and scene segmentation and finally section 5 focuses on methods developed for LOD enhancement and facade reconstruction.

Section 1. Remote Sensing

Satellite and Aerial photogrammetry image acquisition is standard and is available as a huge block or in the form of a few Chips or Tiles of few kilometers having multi spectral channels and high resolution.[30] demonstrates The Digital Imaging and Remote Sensing Image Image Generation (DIRSIG) model for large-scale chips generation for Training. Few overlaps are available to provide accuracy. Satellite Images are now largely available due to Earth engines and satellites like IKONOS but can be expensive. The Aerial Images are largely acquired using UAV- based photogrammetry [10].

The challenges to use these data in the AI pipeline following acquisition include :
1. High resolution.
2. Dozens of other spectral channels including RGB.
3. Geo referencing.
4. Conversion into Nadir images.
This data is available as metadata of the Satellite images in the most common GeoTiff format.

Deep learning has been used to solve these issues.

For multimodal Data fusion [21] techniques like pansharpening and Super-resolution are now performed using CNNs and Autoencoders. End to end Pansharpening was done by stacking unsampled spectral images and learning the values of the central pixel. The autoencoders predict the values through downsampled data provided to it and then upsampling it step by step. [31] reviews CNN-based approaches and Autoencoder-based approaches separately. GAN-based approach has also been proposed that uses two CNN architectures as adversary where discriminator is fully connected and feature level fusion is performed it outperformed existing approaches but still have a lot of room for improvement.

Georeferencing mostly relied on Sift and Surf feature matching earlier but [21] reviews that a CNN trained with 5 convolutional layers and 2 fully connected layers outperforms SIFT algorithm and encourages usage of the deep neural networks for tie-point matching.[23] mentions that such networks still have a need for improvement but are less computationally heavy.

The generation of true Orthophotos [29] is necessary for correct resolutions, drawing the outline in satellite images, and check the degree of the obliquity of other images. They are widely used for DSMs to correct themselves. They may be used as a point of view in stereo pairs for generating 3D objects. Stereo processing is essential for generation disparity maps that can be processed further for a 3D generation. In the next section, we review Stereo processing in detail that uses deep learning to generate DSMs.

Section 2. Height map Generation

Height maps are of different types and are usually the result of subtracting Digital Terrain Maps(DTM) from Digital Surface Maps(DSM). However, that’s not how it's generated.
To generate Digital Elevation Models e need to perform stereo matching using the satellite images. Height maps can also be viewed as GAN image to image translation problem where the input to the Generator is noise and an additional conditional vector for example city component vector obtained using semantic segmentation of urban scenes to obtain certain control over generated fake image. Input to the Discriminator is the fake generated image and Disparity map or previously generated image Heightmap to distinguish between real or fake image. The latest research employing GAN includes [28] and [19]. In [28] used this approach using conditional GAN. They represented the height map as a 2D grid where each point in the grid is a point in the 3D model and is thus an image in the Data structure. As a generator, NetMap was used and a fully connected CNN as a discriminator. In [19] authors try to create a virtual city by generating through Inverse procedural Modelling and employ DCGANs to generate 2D terrain and heightmaps resulting in high-quality 3D models.

Other research views the generation of heightmaps as a Stereo processing problem. The use of Rational Polynomial Coefficient models and Stereo pairs by using epipolar geometry to compute Disparity maps and DSM is very common [5,6,12,14,17] and is done using Semi-Global matching or Deep learning Approach. The SGM-based approach gives high accuracy and has been widely adopted in 3D reconstruction pipeline [5,6,12,14,17] Variant of SGM has been used in [12] which is a census-based method and is robust to light changes and rectification errors.

Stereo Processing using Deep learning techniques have been recently gained popularity. The earliest works were reviewed in [21] where authors used MC-CNN for matching similarities between images in the initial stages of stereo matching and the rest of the stereo pipeline remained the same. However, In [37], authors have approached computation of disparity map purely by training in CNN using DispNet architecture which is Encoder-Decoder based architecture with 26 layers of contracting and expanding network parts. This approach is a starting point in considering Disparity Map computation as a deep learning issue. Recent works include [34] where the use of DenseMapNet with custom mean square loss function and replacing 2D convolutions with depth-wise separable convolutions.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 05, May-2021

DenseMapNet significantly outperformed SGM based approach making stereo matching purely data-driven rather than algorithm-driven. [39] reviewed both non-end-to-end as well as end-to-end stereo matching using Deep learning. Which includes Flownet, Cascade Residual Learning which uses DispNet of [37] , ResNet, and GC-Net. Among these Dispnet proved to be the fastest. Unsupervised deep learning has also been used for stereo matching [37]. The method uses architecture called Deep3D which focuses on minimizing pixel-wise reconstruction loss and significantly improves performance over supervised learning.

A 3D model generation without DSMs or any other kind of Elevation data has been researched in [11] where authors with through Random Forest learn the different attributes responsible for predicting the height of buildings and other city components. The authors chose 10 types of predictors based on Storey, Footprint area, and geometry of the buildings. 17 combinations were made with these predictors to train 17 models and calculated Gross error as well as individual errors in Terms of RMSE and MAE. The story information of in the available turns out to be the best predictor with. This approach, however, has lots of drawbacks like Storey information isn't readily available with satellite data, models either overestimate or underestimate the heights and the method is not foolproof.

Section 3. Scene Segmetation

After the generation of DSMs further processing is to be done to the DSMs to reconstruct the 3D objects. Scene Segmentation or large-scale semantic labeling is the third most important step in 3D model generation.

The earliest work includes [6] where the Support vector machine is used for image classification to learn the patterns in the image and later use it for 3D regeneration. [5] uses U-net architecture for Semantic Labeling of only buildings, however, the use is restricted to building rooftop extraction and polygonization lacks building masks or true textures of buildings resulting in LOD1.[17] overcomes the mask problem by using U-net architecture to extract building, and Tree classification masks in raster format and then use them for polygonization. The method generates accurate building masks and results in better quality 3D visualization I.e LOD2. In works that do not use Satellite images like [19] used Dilated Net for scene parsing to generate city component vectors so that discriminators of the GAN can distinguish from real and fake images. The use of street-level photographs required simple CNN for semantic labeling and thus a 5 layered CNN Alex-net was easily able to achieve the required accuracy for predicting city properties. Urban scenes are mostly segmented according to land use categories like buildings, roads, vegetation, water.

In [33] Volumetric 3D reconstruction algorithm partitions the area into small cells and these areas are then filled with 3D voxels and Large-scale semantic labeling is done assigning each voxel a label. If the voxel is assigned with the building label, then there is a high probability that it is occupied. However, the labels are assigned manually and a decision tree is used to boost the prediction accuracy. This approach shows the importance of scene segmentation in 3D reconstruction.
[21] reviewed usage of CNN in large-scale semantic labeling and thus can be used reliably and opens possibilities for later research. In GAN-based methods [28] used CNN having Mask-Net architecture pre-training it for semantic segmentation to learn building masks and also predict the object to which the building height belongs resulting in LOD2 obtaining MSE of just 5.1. In their works of combining stereo correspondence with semantic segmentation authors trained DenseMapNet and ResNet with 4 input channels and 4th channel being a classification label from semantic segmentation. They used ICNet for the scene parsing the baseline model.

Section 4. 3D reconstruction

Many works [12,14,29,36] approach 3D reconstruction from DSMs directly without having scene segmentation as an intermediate step. The approach is based on Alignment and fusion algorithms to reconstruct the 3D models i.e 3D point clouds are projected aligned and fused and store the 3D parametric model in a database like CityGML. The task requires precision and is computationally expensive and most importantly neglects the object information which is so readily available in the given data. The results are accurate but do not generate high-quality 3D models as reconstruction is viewed as a purely geometric problem.

3D reconstruction was primarily done using procedural modelling. Procedural modelling is basically generating 3D models by defining shape rules or procedural grammar of the 3D object to be built. Different deep learning techniques are now used to automatically generate these procedural grammars. In [3] Authors used a pipeline in which at each stage uses Alex-Net architecture to classify the images and extract the shape grammar. Mostly [13] used 4 classification DNN, 8 regression DNN, and 1 DNN for foundation and roof rule, for other shape rules and rotation rectification respectively from height maps generated from satellite images. DNN being a variation of Res-Net. The result had 84% accuracy.

It is, however, difficult to produce have shape rules to generate 3D models thus other approaches like [35] declarative modeling in which instead of defining procedures actual visual entities are created and molded as per user requirement.

Inverse procedural modelling [19] is another approach that first extracts the procedures and parameters from existing models and then uses it to generate 3D models from procedural modeling.

Section 5. Facade and Texture generation

Facade and texture addition is done for LOD improvement of the generated 3D model. Recent works that use deep learning include [2] in which the depth of elements made of glass was

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 05, May-2021

to be calculated. The problem with the Glass element is that the LiDaR depth data is inaccurate and so using CNN performed depth completion. Other works [9] use random forest for facade labeling into different classes like a window, wall, floor, balcony, etc., and then produce 3D shape rules for 3D reconstruction.

The satellite images mostly deal with a top view of the city and lack the street-level view which is important for facade generation. In [18] authors propose a pipeline to generate a panoramic view of the satellite images using U-Net and Bicycle-GAN. The satellite image is passed through U-net which outputs a semantically segmented image and depth. This is transformed to street-view by a geo-transformation process and then this panoramic segmented image is given as input to the BicycleGAN which translates the image back with its original facade view.

Recently high-quality 3D facades have been viewed as a GAN problem and different GAN networks try to generate the multiple features required to improve the LOD of Buildings. BIM is valuable if LOD improvement is needed. In [20] authors try to generate high-quality 3D buildings using StarGANs and introducing new loss functions like identity loss for the difference between the generated and input data and perpetual loss which minimizes loss of detail. StarGAN has the learning capability of 4 properties at a time. The method used a hybrid chain of 6 StarGANs for generating the required texture of the building and 3D-recGAN for generating 3D data.

III. PROPOSED METHODOLOGY

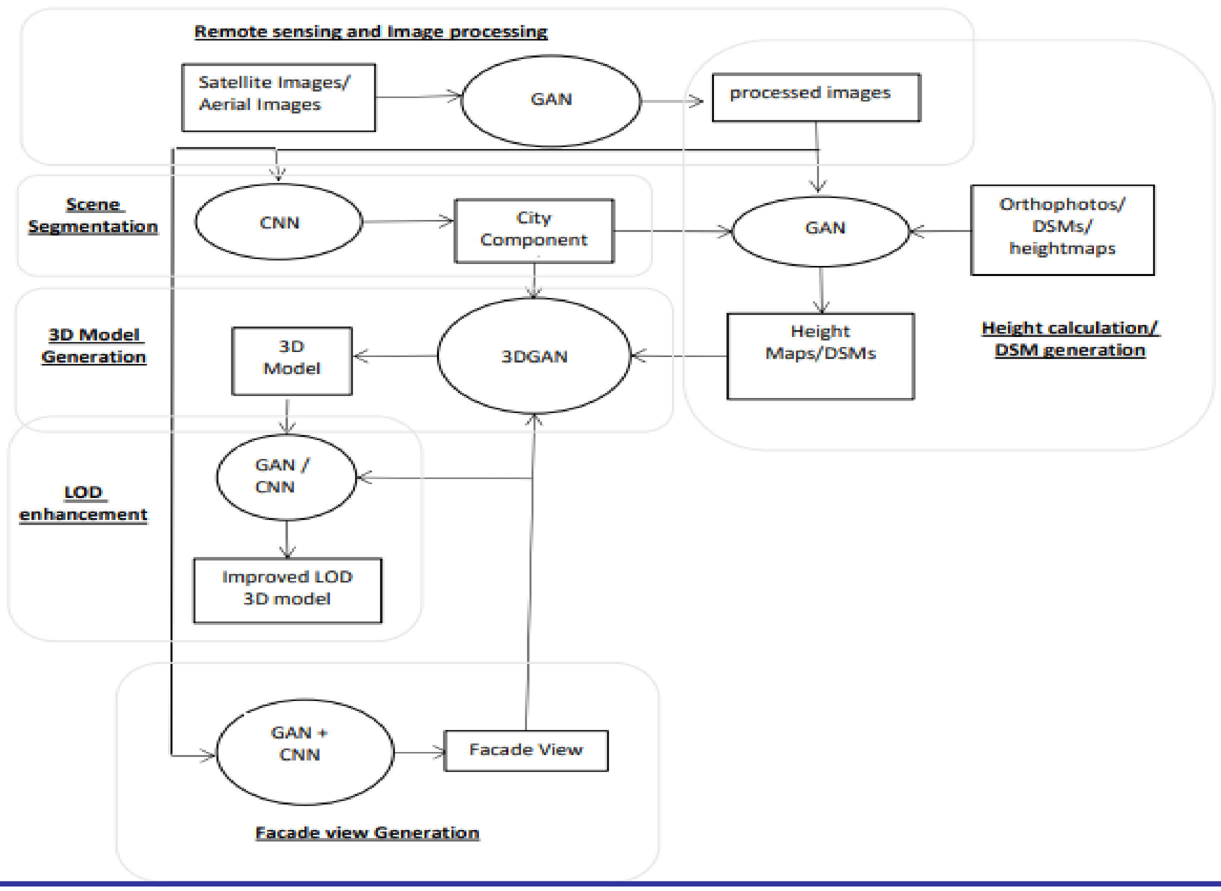
We propose the methodology by viewing the generation of 3D models as a GAN problem. Therefore we start by giving an overview of GANs and then propose a workflow that might help in future research.

Generative Adversarial Networks (GANs)

GANs have come a long way after been proposed by Goodfellow in 2014. The [27] basic idea of GAN is that Generator and discriminator play a game of outwitting one another. In this game, we support the generator to fool the discriminator to produce such a fake result that it persuades the discriminator that it's a real result or at least confuses it about its authenticity. The generator for example is provided with the input as random noise to generate an image of a particular object and it initially does a very lousy job of doing it. The discriminator however is trained with real images of hat particular object and thus can easily classify the generated image as fake or real. Thus generator has to learn more to do its job of generating better fake image and continue learning until it fools its adversary, the discriminator in classifying the fake image as real.

GANs are unsupervised learning models and so the generation cannot be controlled once started and often they have low resolution and noisy output. These generators earlier used o be RBM or VAE but to add a little control to

Figure 1. Work-flow



the learning process many different variations have been proposed since it was first introduced. cGAN, StyleGAN, wGAN, fittingGAN, DCGAN, BEGAN, DRAGAN, cycleGAN, pix2pixGAN, etc. Each of them has a different loss function and handles randomness uniquely. Different applications include Data Enhancement, Domain Transfer, Image Restoration, etc. Recently they have been also useful in 3D data generation [24,26]. We can also transfer the style of the image from one to another which is demonstrated in [22].Gan-based map generation task was demonstrated in[25].GAN usage is still young and still remains lot more applications to be explored in this domain

Workflow

1. Figure 1 gives the overall workflow in diagrammatic representation.
2. 1.The work-flow we propose considers that satellite imagery or aerial image preprocessing can be performed end-to-end using CNN by adding extra layers and fine-tuning it for specific use-cases however since GANs are specific for the image to image translation it might be better suited for operations like pansharpening.
3. 2.The process of height map generation is also been explored as a GAN problem however recent stereo processing use end-to-end CNNs and GANs provide a lot of room for improvement.
4. 3.Image segmentation has now become extremely important for 3D modeling and we already have extremely well-trained CNNs that have high performance for such tasks.
5. 4.3D data can then be learned for height maps and the segmented image and be stored as a parametric model.
6. The LOD enhancement can be achieved if we have a street-level view. For it, we need a facade view of the satellite image that was already carried out using GANs in [18]. Thus GANs can provide a way to improve LODs of already existing 3D models

IV. CONCLUSION

In this paper, we review past and current practices used at various stages of 3D reconstruction of satellite and aerial images. Since the emergence of various deep learning techniques the performance at individual stages and also as a whole has increased. We reviewed various deep learning architectures, GAN and CNN in particular that have outperformed algorithms like SGM for stereo processing and 3D modeling and manual scene parsing. We also review the importance of scene segmentation for 3D modeling and that semantic information can be used to further enhance the 3D modeling process and 3D modeling should not be viewed only as a geometric task. The use of deep learning is still young and since the development of GAN lot of semi-automated or CAD-dependent tasks can be automated. We also demonstrated the basics of GAN and finally provided a workflow using GANs and CNN 3D reconstruction that can be further explored. We provided a review regarding each of the stages and finally proposed our own workflow that can be helpful for future research..

REFERENCES

[1] Biljecki, F.; Stoter, J.; Ledoux, H.; Zlatanova, S.; Çöltekin, A. Applications of 3D City Models: State of the Art Review. ISPRS Int. J. Geo-Inf. 2015, 4, 2842-2889. (*references).*

[2] Xu H., Chen CY., Delmas P.J., Gee T.E., van der Mark W. (2019) Multimodal 3D Facade Reconstruction Using 3D LiDAR and Images. In: Lee C., Su Z., Sugimoto A. (eds) Image and Video Technology. PSIVT 2019. Lecture Notes in Computer Science, vol 11854. Springer, Cham. https://doi.org/10.1007/978-3-030-34879-3_22

[3] Gen Nishida, Adrien Bousseau, Daniel G. Aliaga. Procedural Modeling of a Building from a Single Image. Computer Graphics Forum, Wiley, 2018, Eurographics, 37 (2), ⟨10.1111/cgf.13372⟩. ⟨hal-01810207⟩.

[4] E. Zheng, K. Wang, E. Dunn and J. Frahm, "Minimal Solvers for 3D Geometry from Satellite Imagery," 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, 2015, pp. 738-746, doi: 10.1109/ICCV.2015.91.

[5] Sebastien Tripodi, Liuyun Duan, Veronique Poujade, Frederic Trastour, Jean-Philippe Bauchet, et al.. Operational Pipeline for Large-scale 3D Reconstruction of Buildings from Satellite Images. IGARSS 2020 - IEEE International Geoscience and Remote Sensing Symposium, Sep 2020, Big Island/Virtuel, United States. ffhal-02966821f.

[6] Pamungkas I.R., Suwardi I.S. (2015) 3D-Building Reconstruction Approach Using Semi-global Matching Classified. In: Intan R., Chi CH., Palit H., Santoso L. (eds) Intelligence in the Era of Big Data. ICSIIT 2015. Communications in Computer and Information Science, vol 516. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-46742-8_35.

[7] Filip Biljecki, Hugo Ledoux, Jantien Stoter, An improved LOD specification for 3D building models, Computers, Environment and Urban Systems,Volume 59, 2016, Pages 25-37, ISSN 0198-9715, https://doi.org/10.1016/j.compenvurbsys.2016.04.005.

[8] Yichuan Deng, Jack C.P. Cheng, Chimay Anumba, Mapping between BIM and 3D GIS in different levels of detail using schema mediation and instance comparison, Automation in Construction, Volume 67, 2016, Pages 1-21, ISSN 0926-5805,https://doi.org/10.1016/j.autcon.2016.03.006.

[9] A. Martinović, J. Knopp, H. Riemenschneider and L. Van Gool, "3D all the way: Semantic segmentation of urban scenes from start to end in 3D," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015, pp. 4456-4465, doi: 10.1109/CVPR.2015.7299075.

[10] O. Esrafilian and D. Gesbert, "3D City Map Reconstruction from UAV-Based Radio Measurements," GLOBECOM 2017 - 2017 IEEE Global Communications Conference, Singapore, 2017, pp. 1-6, doi: 10.1109/GLOCOM.2017.8254657.

[11] Filip Biljecki, Hugo Ledoux, Jantien Stoter, Generating 3D city models without elevation data, Computers, Environment and Urban Systems, Volume 64, 2017, Pages 1-18, ISSN 0198-9715, https://doi.org/10.1016/j.compenvurbsys.2017.01.001.

[12] G. Facciolo, C. De Franchis and E. Meinhardt-Llopis, "Automatic 3D Reconstruction from Multi-date Satellite Images," 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, 2017, pp. 1542-1551, doi: 10.1109/CVPRW.2017.198.

[13] Zeng H., Wu J., Furukawa Y. (2018) Neural Procedural Reconstruction for Residential Buildings. In: Ferrari V., Hebert M., Sminchisescu C., Weiss Y. (eds) Computer Vision – ECCV 2018. ECCV 2018. Lecture Notes in Computer Science, vol 11207. Springer, Cham. https://doi.org/10.1007/978-3-030-01219-9_45.

[14] Ewelina Rupnik, Marc Pierrot-Deseilligny, Arthur Delorme. 3D reconstruction from multi-view VHRsatellite images in MicMac. ISPRS Journal of Photogrammetry and Remote Sensing, Elsevier, 2018, ff10.1016/j.isprsjprs.2018.03.016ff. ffhal-02369304f.

[15] Bshouty, Eliana & Shafir, Alexander & Dalyot, Sagi. (2019). Towards the generation of 3D OpenStreetMap building models from single contributed photographs. Computers, Environment and Urban Systems. 79. 101421. 10.1016/j.compenvurbsys.2019.101421.

[16] Döllner, J. Geospatial Artificial Intelligence: Potentials of Machine Learning for 3D Point Clouds and Geospatial Digital

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 05, May-2021

Twins. PFG 88, 15–24 (2020). https://doi.org/10.1007/s41064-020-00102-3.

[17] Tripodi, S. et al. “AUTOMATED CHAIN FOR LARGE-SCALE 3D RECONSTRUCTION OF URBAN SCENES FROM SATELLITE IMAGES.” ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (2019): 243-250.

[18] X. Lu, Z. Li, Z. Cui, M. R. Oswald, M. Pollefeys and R. Qin, "Geometry-Aware Satellite- to-Ground Image Synthesis for Urban Areas," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020, pp. 856-864, doi: 10.1109/CVPR42600.2020.00094.

[19] Kim, S., Kim, D. & Choi, S. CityCraft: 3D virtual city creation from a single image. Vis Comput 36, 911–924 (2020). https://doi.org/10.1007/s00371-019-01701-x.

[20] Du, Z., Shen, H., Li, X. et al. 3D building fabrication with geometry and texture coordination via hybrid GAN. J Ambient Intell Human Comput (2020). https://doi.org/10.1007/s12652-020-02488-9.

[21] .X. X. Zhu et al., "Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources," in IEEE Geoscience and Remote Sensing Magazine, vol. 5, no. 4, pp. 8-36, Dec. 2017, doi: 10.1109/MGRS.2017.2762307.

[22] Zhang, H. 3D Model Generation on Architectural Plan and Section Training through Machine Learning. Technologies 2019, 7, 82.

[23] Qin, Rongjun & Gruen, Armin. (2020). The role of machine intelligence in photogrammetric 3D modeling – an overview and perspectives. International Journal of Digital Earth. 1-17. 10.1080/17538947.2020.1805037.

[24] https://arxiv.org/abs/1610.07584v2

[25] https://arxiv.org/abs/1902.05611v2

[26] TY - BOOK AU - Zhu, Jing AU - Xie, Jin AU - Fang, Yi PY - 2018/02/10 SP - T1 - Learning Adversarial 3D Model Generation with 2D Image Enhancer ER - TY - JOUR AU - Jin, Lianchao AU - Tan, Fuxiao AU - Jiang, Shengming PY - 2020/08/01 SP - 1 EP - 17

[27] T1 - Generative Adversarial Network Technologies and Applications in Computer Vision VL - 2020 DO - 10.1155/2020/1459107 JO - Computational Intelligence and Neuroscience ER -

[28] Gorbatsevich, V. & Kulgildin, B. & Melnichenko, M. & Vygolov, Oleg & Vizilter, Yu. (2020). SEMI-AUTOMATIC CITYSCAPE 3D MODEL RESTORATION USING GENERATIVE ADVERSARIAL NETWORK. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XLIII-B2-2020. 415-420. 10.5194/isprs-archives-XLIII-B2-2020-415-2020.

[29] Amhar, Fahmi & Jansa, Josef & Ries, Christine. (1998). The generation of true orthophotos using a 3D building model in conjunction with a conventional DTM. International Archives of Photogrammetry and Remote Sensing. 32.

[30] Han, Sanghui & Farfard, Alex & Kerekes, John & Gartley, Michael & Ientilucci, Emmett & Savakis, Andreas & Law, Charles & Parhan, Jason & Turek, Matt & Fieldhouse, Keith & Rovito, Todd. (2017). Efficient generation of image chips for training deep learning networks. 1020203. 10.1117/12.2261702.

[31] Tsagkatakis, Grigorios; Aidini, Anastasia; Fotiadou, Konstantina; Giannopoulos, Michalis; Pentari, Anastasia; Tsakalides, Panagiotis. 2019. "Survey of Deep-Learning Approaches for Remote Sensing Observation Enhancement" Sensors 19, no. 18: 3929. https://doi.org/10.3390/s19183929

[32] Nex, Francesco & Macay Moreira, José & Agugiaro, Giorgio & Remondino, Fabio & Lim, N.J.. (2013). From DSM to 3D building models: A quantitative evaluation. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XL-1/W1. 10.5194/isprsarchives-XL-1-W1-213-2013.

[33] Häne, Christian & Zach, Christopher & Cohen, Andrea & Angst, Roland & Pollefeys, Marc. (2013). Joint 3D Scene Reconstruction and Class Segmentation. Proceedings / CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 97-104. 10.1109/CVPR.2013.20.

[34] Bosch, Marc & Foster, Kevin & Christie, Gordon & Wang, Sean & Hager, Gregory & Brown, Myron. (2018). Semantic Stereo for Incidental Satellite Images.

[35] Freiknecht, J.; Effelsberg, W. A Survey on the Procedural Generation of Virtual Worlds. Multimodal Technol. Interact. 2017, 1, 27.https://doi.org/10.3390/mti104002.

[36] Kocaman, Sultan & Zhang, L & Gruen, Armin & Poli, D. (2006). 3D city modeling from high-resolution satellite images. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 36.

[37] Mayer, Nikolaus & Ilg, Eddy & Hausser, Philip & Fischer, Philipp & Cremers, Daniel & Dosovitskiy, Alexey & Brox, Thomas. (2016). A Large Dataset to Train Convolutional Networks for Disparity, Optical Flow, and Scene Flow Estimation. 4040-4048. 10.1109/CVPR.2016.438.

[38] Hamid, Mohd Saad & Abd Manap, Nurulfajar & Hamzah, Rostam & Kadmin, ahmad fauzan. (2020). Stereo Matching Algorithm based on Deep Learning: A Survey. Journal of King Saud University - Computer and Information Sciences. 10.1016/j.jksuci.2020.08.011.