Project Proposal - Literature and Technology Review
Semantic Segmentation of Google Maps data to aid Architectural Planning

James Little

MSc in Data Science
The University of Bath
2022

Semantic Segmentation of Google Maps data to aid Architectural Planning

Submitted by: James Little

Copyright

Declaration

This dissertation is submitted to the University of Bath in accordance with the requirements of the degree of Master of Science in the Department of Computer Science. No portion of the work in this dissertation has been submitted in support of an application for any other degree or qualification of this or any other university or institution of learning. Except where specifically acknowledged, it is the work of the author.

Contents

List of Figures

1. Introduction

In recent years deep learning has seen a resurgence within the computer vision research community, with GPU computing increasing its feasibility and the exponential increase in available data from which Convolutional Neural Networks (CNNs) can operate effectively. During this resurgence, the architectural industry has slowly begun to wake up from its technological slumber after failing to adapt to the changing times when other industries invested in online business growth. Although slow, this is changing with architectural startups beginning to venture into the online space to allow for a low hassle virtual approach to a deeply ingrained face to face industry. This shift in mentality was caused by the influx of novel solutions provided by advancements in machine learning and the rapid/exponential increase in the amount of generated and stored data.

In architectural businesses around the U.K., there has been a large surge in home renovation and small scale extensions, with individuals spending more time at home due to the flexible working environment brought in by the COVID-19 pandemic. With potential customers now spending more time at home than ever before and the record spike in house prices over the last three years, consumers are turning to these smaller-scale renovations to maximise space efficiency at less expense. This is where the services of a new type of architectural business model have stepped in to meet the demand created and alleviate the need for a bespoke individualised architecture service that is in-affordable to the majority and currently offered by more traditional planning routes.

This novel business model takes the more traditional in-person approach to planning and uncomplicates the process by venturing into the online space to enable low-cost, rapid early-stage design iterations such that homeowners can understand what projects are possible within a defined time frame and budget. This is enabled through a client portal that provides a hub for a customer's entire service experience and allows a fluid back and forth to facilitate changes quickly within a fully virtual design experience. This shift in how the industry is conducting its business presents an opportunity for the field of computer vision and deep learning to make an impact and aid the architectural planning process.

A drawback of this new approach is the absence of any ability to explore options early in the design process, as more traditional architectural practice consultation meetings are not feasible within the business model. Therefore, during the early stages of the customer experience, tools are needed to allow clients to explore project ideas within the scope of their properties; surroundings, boundaries and location.

This research proposes a system to aid the early stage of home renovation through the use of semantic image segmentation techniques on google satellite maps data within a data pipeline shown in Figure 1.1, to enable recommendations to users for architecture extension work. Although many semantic image segmentation techniques exist, this study shall look specifically at applying deep convolution neural networks. An assessment will be made of past techniques to understand the evolution of techniques within the field and their effectiveness against the current deep learning technologies.

Figure 1.1: Project data pipeline

1.1 Deliverables

The following section looks to define some of the primary objectives of the project. These deliverables will act as milestones accompanied by research questions that the paper will hope to answer. The success of these sections will directly influence the outcome of the project and act as the main themes within the research.

1.1.1 Semantic Segmentation Deep Learning Architecture

The main deliverable of this study is to produce a deep learning model that can semantically segment satellite imagery with the following research questions trying to be answered.

Research Question 1 - Can a deep learning model be created that can semantically segment different areas of a satellite image utilising past and novel architectures to produce a satellite image specific architecture?

Research Question 2 - Are deep learning techniques still the best way to categorise and segment visual data?

Specifically, this study proposes an Encoder-Decoder Neural Network to provide semantic segmentation for objects within a satellite image scene. This technique has seen prior use within biomedical image segmentation data such as X-ray and CT-Scan imagery, and research has not surprisingly been applied to satellite imagery, although this research still has lots to offer.

The Encoder-Decoder Neural Network architecture consists of two parts; an encoder and a decoder; the goal behind this model is not just to make a classification on the pixel level but to also provide a mechanism to project the features that were learned into the pixel space. The encoder is usually a pre-trained classification network like V.G.G./ResNet, where convolution blocks and max-pooling down-sampling are applied to encode the input image. The decoder's goal, is to project the discriminated features, this is done using up-sampling and concatenation layers, which are then followed by convolution operations.

There has been an a lot of research and implementation on the topic, with a few notable examples being; U-Net, V-Net and Seg-Net. The U-Net architecture can be seen in Figure 1.2. All were trained and made with medical imagery in mind but have been utilized well for other applications. Although it is clear that a more targeted model structure would be beneficial

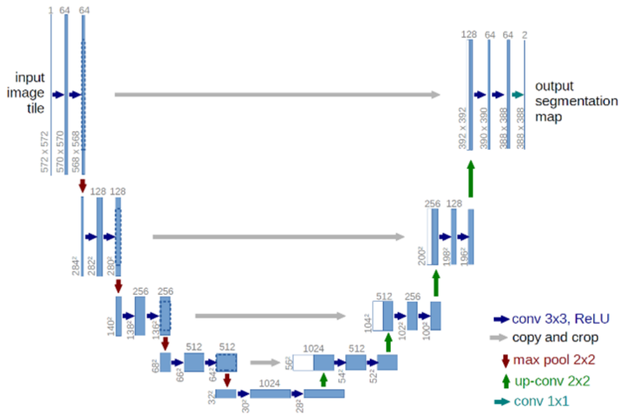

Figure 1.2: U-Net Architecture

when executed on satellite imagery, this will be an area where the paper will aim to provide insight.

1.1.2 Dataset Specification and Amendments

Research Question 3 - What amendments can be made to the dataset to aid the segmentation of further classes?

To train the deep learning architecture described above, a dataset will need to be selected to enable the semantic segmentation output for the satellite imagery. The difficulty with this selection comes with how abstracted the majority of the public datasets are, with them being used to test the accuracy of different deep learning architectures instead of solving specific problems described within this research. Therefore, this analysis will evaluate the difference between new dataset creation and how datasets can be amended to implement further segmentation classes to allow larger use cases for these datasets.

1.1.3 Use of Semantically Segmented Images for Architectural Planning

Research Question 4 - How can the segmented data be used to aid architectural design and planning?

From the segmented image output produced from the previous deliverables; the study hopes to deliver a system that allows architectural recommendations to be given based on the segmented regions in the areas around a selected property. This will take the form of an interactive U.I. that will allow the user to experiment with property extensions of different sizes within the boundaries created from the segmented regions. Furthermore, there is an abundance of publicly available property information that can and will be used to aid the final output to allow users to understand the possible renovations that can occur.

2. Literature and Technology Survey

2.1 Introduction

Image segmentation is the operation of partitioning images into segments known as image regions; segmentation aims to provide meaning to these regions by adding value through labels and spatial location. Although the research explores the more modern implementations utilising deep learning neural networks, research into image segmentation has used more classical computer vision techniques to achieve the same results for the last 35 years. To explore the history and story of image segmentation, both past and present research will be analysed to illustrate the progression within the discipline and survey its limitations.

2.1.1 Segmentation Type Variation

Image segmentation breaks down into two classes: semantic and instance. These both segment image data but have nuances that differentiate them. The semantic segmentation approach detects classes of objects within a scene by classifying individual pixels as a member of that class. Instance segmentation works by classifying instances of classes within the scene, with multiple instances of the same class being classified differently by the pixels. This difference can be seen explicitly within this imagery by TheCodingBug (2021) (Figure 2.1).

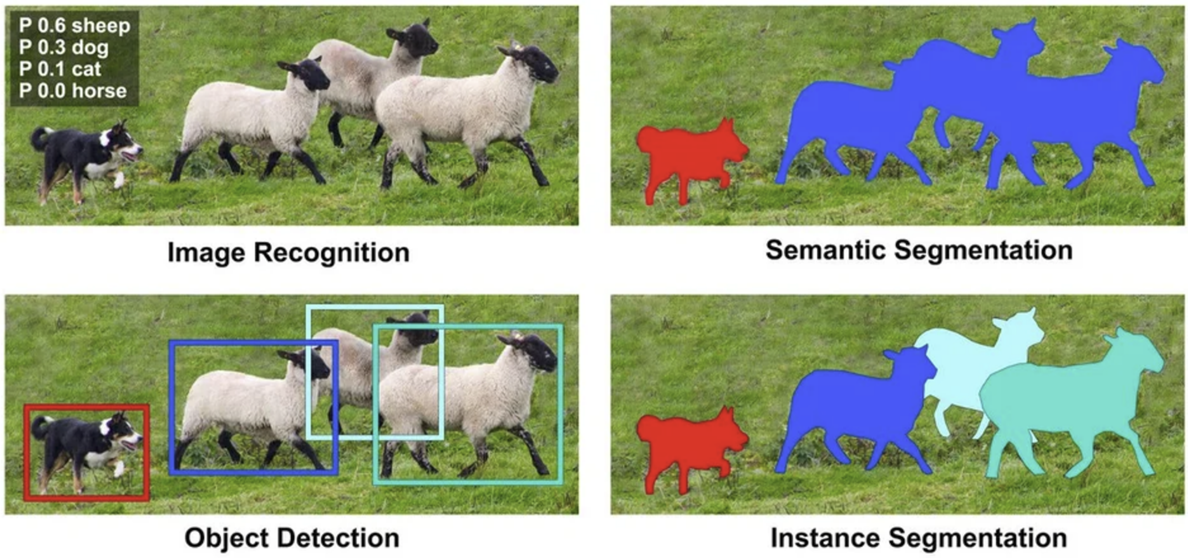

Figure 2.1: Image Segmentation Types

2.2 Segmentation Techniques

2.2.1 Classical Approaches

Thresholding

Thresholding is the simplest of the segmentation techniques available; in its most basic state, it is used to separate objects from their backgrounds within an image scene. This separation occurs by comparing each pixel value (usually pixel intensity) to a specific threshold, classifying pixels into those lower and greater than that threshold, with methodology implementations looking to identify the correct selection of the threshold value. One of the initial papers written on image segmentation was by Otsu, N (1979); although it provided a simple solution, this paper became the foundation for all thresholding research that succeeded it. As thresholding image segmentation entered the 21st-century, research began to slow; apart from a notable exception, Al-Amri, S.S. (2010) implemented several published techniques to compare their segmentation results when used on complex satellite imagery and contrast the effectiveness of each implementation.

Region Growing

Region-growing image segmentation works on the assumption that pixels that share spatial proximity will also share similar values. The main process of region growing works to cluster pixels that represent homogeneous areas in the image, and this is done by selecting random pixels and regions that are grown pixel by pixel outwards according to similarity properties (e.g. intensity). This approach works well in complex imagery due to its adaptability to overcome noise, ambiguous boundaries, and partial occlusion.

Region-growing image segmentation began with a paper written by Adams, R. and Bischof, L. (1994); this paper presented region growth in its most basic form whilst presenting the algorithm's properties. Research then turned to build upon this foundation with two papers highlighting the shift in study. Espindola, G.M. et al. (2006) looked to solve the problem of parameter selection within the algorithm to overcome the need for user-defined control parameters that directly affect segmentation accuracy. Shih, F.Y. and Cheng, S. (2005) chose to solve the issue of automatic segmentation of colour images. Later research did build on this, seen in research conducted by Preetha, M.M.S.J. et al. (2012). Although, this paper did mark the end of research into the area with researchers looking to utilise newer technologies.

2.2.2 Machine Learning Approaches

Clustering Methods

Clustering is a technique used as a un/semi-supervised learning implementation within machine learning. The goal for clustering is to group similar data points based on the similarity between properties they have; in more straightforward cases, these properties can be spatial (Euclidean Distance) but with more complex implementations utilising Chi-squared and Jenson-Shannon. When applied to image segmentation, three distinct algorithms are predominantly used: K-means, adaptive K-means and fuzzy c-means.

Burney, S.A. et al. (2014) proposed a methodology based on the K-means algorithm in conjunction with a novel subtractive clustering algorithm that generates the cluster's centroid based on potential values instead of exclusively using currently used clustered data points.

Zheng, X. et al. (2018) proposed an adaptive k-means solution that built upon the classic k-means algorithm by generating accurate segmentation results that avoid the input defined value of K. The paper also explains the preprocessing stages used to make the prepare the image; using a transform from to the L.A.B. colour space and then setting a luminance threshold to reduce the effect of bright light. When used in conjunction with other techniques, these papers illustrated the ability of clustering to provide a complete solution when training data is absent.

Neural Networks and Deep Learning

Fully Convolutional Networks

F.C.N.s are based on C.N.N.s (Convolutional Neural Networks), representing the most widely used deep learning technique for computer vision tasks. F.C.N.s utilise the same structure as C.N.N.s, excluding dense/fully connected layers, with them being substituted for fully convoluted layers.

One of the initial deep learning models for semantic segmentation was proposed by Long, J. et al. (2015). The authors expand on existing C.N.N structures, AlexNets, V.G.G. net and GoogLeNet, to implement an F.C.N. that took arbitrarily sized inputs and produced segmented outputs of the corresponding size. The paper was the first to define and implement a skip architecture that helped aid accurate and detailed segmentation. The skip architecture combined coarse, high layer information with fine, low layer information; this led to the prediction of finer details that also retained some higher-level semantic information about the image. F.C.N.s have been used in a wide range of segmentation use cases such as; automatic road extraction (Buslaev, A et al., 2018), satellite segmentation (Wurm M, 2019), brain tumour identification segmentation (G. Wang et al., 2017) and livestock detection (Han, L. et al., 2019.). A drawback experienced when using F.C.N.s is their ability to neglect higher-level contextual information; the skip architecture attempts to solve this but does not do enough to nullify the problem. In response to F.C.N.s inability to incorporate higher-level semantic information, the research looked to implement graphical models to aid the process with; Markov Random Fields (Liu, Z. , 2017) and Conditional Random Fields (Chen LC, et al. 2014) being used.

Generative Adversarial Networks

While CNN's are a natural and popular fit for computer vision segmentation tasks, other networks are utilised to produce results similar to or greater than their CNN counterparts; one such group is the Generative Adversarial Networks (GANs).

Luc, P et al. (2016), proposed one of the initial pieces of research to understand the cross-compatibility of GAN to semantic segmentation. The research proposed training both a convolutional semantic segmentation network and a discriminative network that distinguishes between the ground and synthesised images. This methodology explicitly created probability maps for each class in each image then the discriminator distinguished between the maps of ground truth values and the image map.

Souly, N et al. (2017) looked to build upon this foundational research but instead with a semi-supervised GAN that utilised a generator network to act as a discriminator to provide extra training samples and enable increased accuracy within the multi-class output mask. Iqbal, T et al. (2018) looked to apply an amended architecture to medical imagery, focusing on utilising smaller datasets known to overfit on other deep learning techniques. Although most GAN-based image segmentation applications look to segment medical data, Abdollahi A et al. (2021) proposed a GAN approach to semantic segment satellite data. The approach used a modified U-Net (encoder/decoder) to provide the generative section and then utilised a standard discriminator and edge-preserving filters to obtain a strong classification rate.

Encoder-Decoder (Spatial Pyramid Pooling) Based Models

The models that represent the most cutting-edge solutions within the field of image segment are known as Encoder-Decoder models. These models are split into the encoding and decoding fragments that provide the architecture's backbone. The encoder works to take an arbitrarily sized input, and then using convolution layers, it encodes the input into a fixed-length 'internal' representation. The decoder then works to deconvolute the encoding to create a pixel-wise segmentation mask.

Noh H et al. (2015) and Badrinarayanan V et al. (2017) published earlier papers on Encoder-Decoder's usage within semantic segmentation. Both sought to implement the VGG16 CNN for the encoding segment and very similar Decoders, with the latter building on the previous by implementing up-sampling to prevent the need for the 'learning portion' of the training.

The primary usage case for the Encoder-Decoder architectures has been medical imagery segmentation. Ronneberger P et al. (2015), proposed the first widely recognised architecture known as U-Net to segment biological microscopic images. The architecture proposes a contracting (encode) and expansive (decode) path to enable the segmentation transformation. The contracting path consists of convolutions, R.E.L.U. and max pooling operations to simplify input imagery. The expansive then consists of up-convolutions mixed with R.E.L.U. to map each component vector to the number of classes present. This novel architecture allowed for high accuracy on small training datasets, a rare feature for implementations of this type. Milletari F et al. (2016) published a paper introducing V-Net the following year, which built off the same foundations as U-Net (Ronneberger P et al., 2015) but with notable changes in architecture and performance. The work proposes introducing a novel objective function based on the Dice coefficient to optimise the model more efficiently during training. Like previous research, the ability to source training data was slow and produced little in terms of quantity; therefore, non-linear transformations and histogram matching were utilised to increase the size of the dataset. Although the above has been highlighted, there are hundreds of implementing architectures, all with their use cases and accuracy with various data (Kim JU et al., 2017) (Chen, L C et al., 2018).

2.3 Technology Survey

The section identifies current software implementations that look to either aid architecture planning or identify property features using satellite data. Due to the vast range of implementations currently available, two particular software implementations are analysed as they represent the most cutting edge technology within this sector that directly apply to this research.

SearchLand (SearchLand Ltd, 2020) is a web app that uses private and publicly available data from the U.K. land registry and other governmental and private data services to aid architectural decision making by enabling more informed property and land decisions. This is then used in cooperation with the Google Maps API to provide an interactive U.I. that layers information regarding a property's land boundary (per the U.K. land registry), planning application information, price paid analysis and planning constraint information (Figure 2.2). Although the application boasts an 80 percent U.K. area coverage, there are still areas where data is missing, and therefore, analysis is only partially available. The system also falls short by not utilising the data curated to make strategic recommendations for its clients, which would move this product into the consumer market as it is only currently marketed towards more industrial applications.

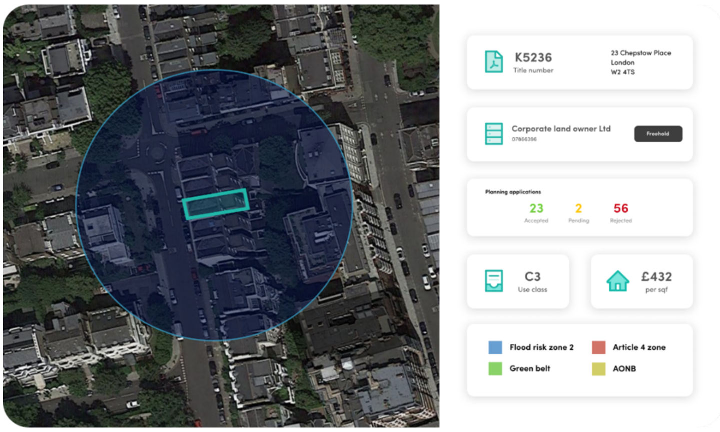

Figure 2.2: SearchLand – Property Boundary Identification

Cover (Cover Technologies Inc, 2021) provides a service matching the proposed system with notable differences. Instead of providing a tool to aid the architectural planning of homeowners, they provide an all-in-one solution to utilise backyard space. Consequently, the system they use to automate property eligibility for the renovation process is unique. This system uses the 'map' layer within the Google Maps API to identify the property boundary and buildings for the desired property (Figure 2.3); this data is then processed to make renovation recommendations within the constraints of the property and its surroundings. However, the property boundaries provided by google maps are not available in any other location worldwide, effectively rendering this solution unusable in other locations. Another drawback includes the house building blueprint adopted within Los Angeles. Within this area, analysis of property boundaries is simple due to the grid-based planning system adopted in parts of the U.S, which will inevitably cause problems when the system is applied to the more complicated property boundaries in other areas of the United States and Europe.

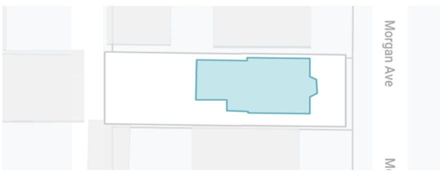

Figure 2.3: Cover – Identification Process

2.4 Conclusion

To conclude, the literature review analysed classical vision techniques and the more 'state of the art' machine learning implementations to identify critical research points and implementational knowledge to build upon the initially proposed system. This analysis enabled a full picture of how the discipline has developed and evolved from its conception in the late 1970s (Otsu. N, 1979). From this analysis, it is clear that the more classical approaches influence the more contemporary research. However, as can be seen from the analysis, the practicality of the older techniques ended with the newer research proposed by the machine learning community, with thresholding and region-growing looking archaic in comparison.

On analysis of the machine learning approaches, clustering presented ideas that within the context of the proposed problem will struggle to aid the process as the unsupervised aspect of the technique would cause too much variability within segmented output imagery, causing problems in the latter stages of development. F.C.N.s then offered a supervised technique for the problem faced, but the inability to capture higher-level contextual information nullified its usefulness but offered insight into the ability of deep learning to outperform the status quo. The evaluation of GANs came to a similar conclusion, with the technique providing good segmentation results but providing functionality not needed for this study. Additionally, the lack of intrinsic metric evaluation for this architecture meant iterative improvements in model accuracy are harder to evaluate, with other methodologies offering increased insight. The Encoder-Decoder architecture offered the best solution for the proposed problem. Much of its functionality and the academic record indicate its strength in semantic segmentation, plus the ability to handle variable-length inputs will allow for flexibility with the proposed software. For these reasons, the proposed solution will utilise fundamentals outlined within the U-Net architecture (Figure 1.2) (Ronneberger P et al., 2015) to create a satellite imagery semantic segmentation model.

The two technologies described within the technology review represent the current best of what is possible within the field of automated architecture to improve the architectural experience for all involved. SearchLand provided a solution to providing information regarding property boundaries and information relating to the property's past. However, these insights fall short when land registry data is not present or incomplete, leaving gaps in the coverage of the solution. Cover applied a system like the one proposed by this paper, but its implementation is currently limited to the Los Angeles area due to data limitations that have stopped the company from expanding its reach commercially. The method proposed within this paper will build upon the foundations of these applications and look to implement deep learning techniques to see if we can alleviate some of their issues or offer an alternate way to solve this multifaceted problem.

3. Appendix

3.1 Project Timeline

The project should run from the 10th of March until the 1st of September. Figure 3.1 breakdowns how progress will be made during the time period for both the larger deliverables and smaller milestones.

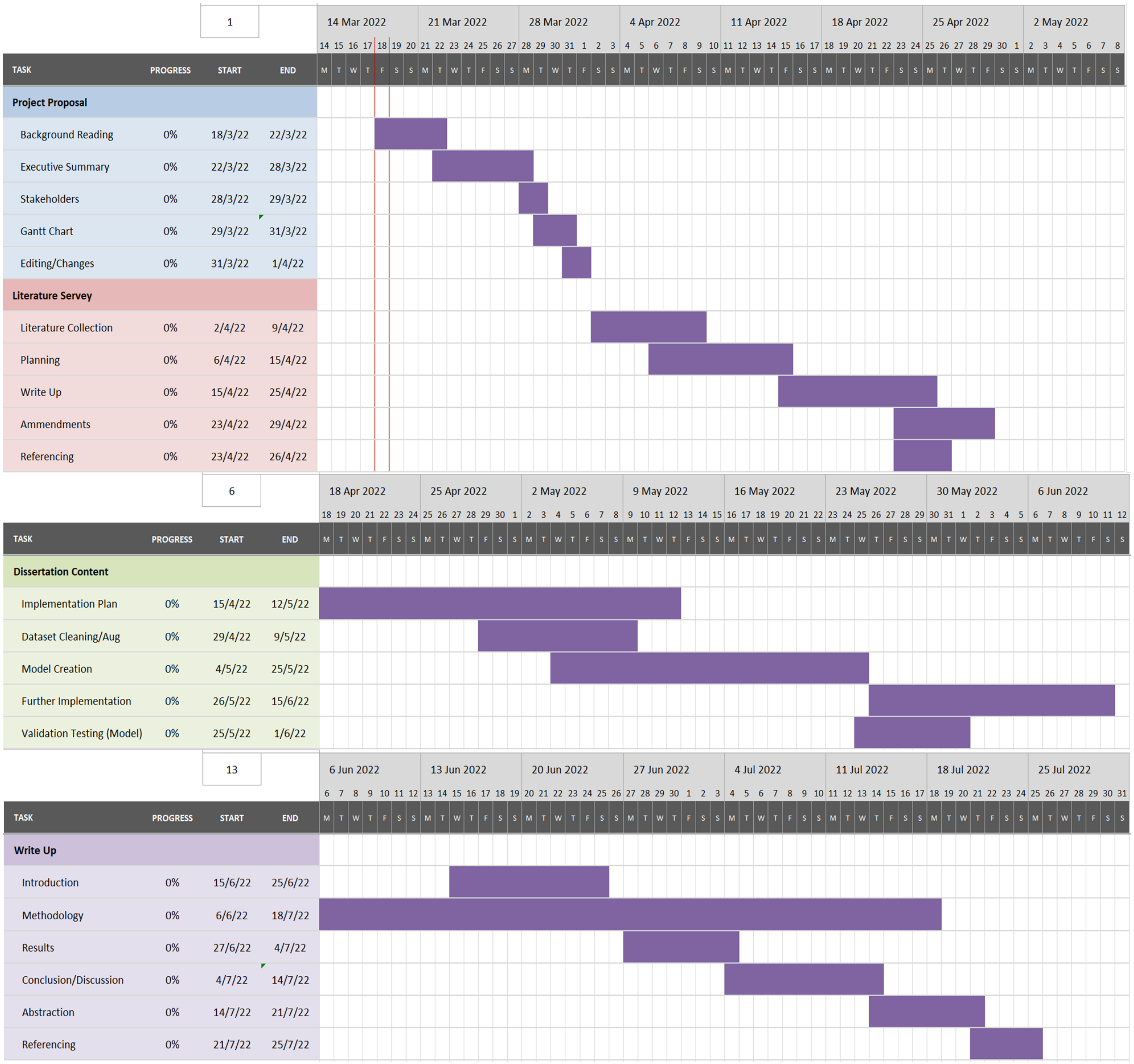

Figure 3.1: Project Timeline - Gantt Charts

3.2 List of Resources and Availability

- University GPU Cluster (depending on model implementation and availability)
- CITY-OSM Dataset (3000 - image label pairs)

3.3 Dataset Specification

The CITY-OSM dataset employs the open street map (O.S.M.) to provide a simple classification of items within a satellite image scene; the classes include pixel-wise building locations, roads and background labels (Figure 3.2). The dataset consists of 3000 image segmentation pairs representing satellite imagery of areas worldwide: Potsdam, Berlin, Chicago, Paris, Tokyo and Zurich. The significant upside of this dataset is the number of items within the data; this is not only rare within satellite segmentation but in all segmentation data. This is due to the need to source both input imagery and segmentation annotation/labelling, which is both costly and time-consuming. This dataset overcomes the issue of enormous annotation work but utilises the crowd-sourced mapping system known as Open-Street-Map to segment the data into three classes defined earlier. The drawback of this data is the lack of complexity, as more meaningful research will want to extract a broader range of information from a satellite image. The dataset is also known to have accuracy errors with the segmented label imagery, but this is overcome the quantity of the data.

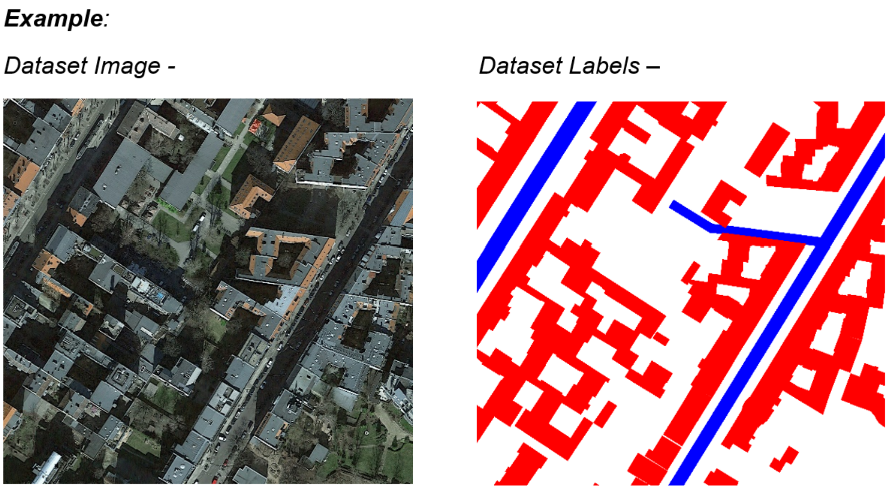

Figure 3.2: CITY-OSM. Image/Label Pair

Legend -

- Blue - Roads
- Red – Building Structures
- White – Background

This dataset will be used first to evaluate the deep learning architecture proposed by this paper as this will provide a strong indication of its ability to semantically segment satellite imagery datasets. Then as defined by research question 3, the analysis will be conducted to attempt to amend this data so that it can be used for architecture planning recommendations.

This will include adding a fourth class (green) to the labelled images depicting the property boundaries of the buildings within the images (Figure 3.3) . Due to the varied nature of the images within this dataset only around 450 of the 3000 images represent residential scenes where property boundaries are easily visible (e.g not high-rise urban settings). This analysis will then investigate the feasibility of adding such annotations and provide options to overcome any difficulty.

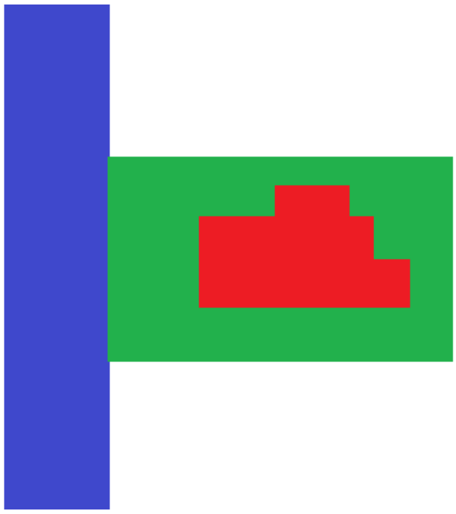

Figure 3.3: CITY-OSM. Dataset Amendment

4. Bibliography

Abdollahi, A., Pradhan, B., Sharma, G., Maulud, K.N.A. and Alamri, A., 2021. Improving road semantic segmentation using generative adversarial network. IEEE Access, 9, pp.64381-64392.

Acharjya, P.P., Das, R. and Ghoshal, D., 2012. Study and comparison of different edge detectors for image segmentation. Global Journal of Computer Science and Technology.

Adams, R. and Bischof, L., 1994. Seeded region growing. IEEE Transactions on pattern analysis and machine intelligence, 16(6), pp.641-647.

Al-Amri, S.S. and Kalyankar, N.V., 2010. Image segmentation by using threshold techniques. arXiv preprint arXiv:1005.4020.

Al-Amri, S.S., Kalyankar, N.V. and Khamitkar, S.D., 2010. Image segmentation by using edge detection. International journal on computer science and engineering, 2(3), pp.804-807.

Badenas, J., Sanchiz, J.M. and Pla, F., 2001. Motion-based segmentation and region tracking in image sequences. Pattern recognition, 34(3), pp.661-670.

Burney, S.A. and Tariq, H., 2014. K-means cluster analysis for image segmentation. International Journal of Computer Applications, 96(4).

Buslaev, A., Seferbekov, S., Iglovikov, V. and Shvets, A., 2018. Fully convolutional network for auoad extraction from satellite imagery. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops (pp. 207-210).

Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K. and Yuille, A.L., 2017. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE transactions on pattern analysis and machine intelligence, 40(4), pp.834-848.

Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F. and Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European conference on computer vision (E.C.C.V.) (pp. 801-818).

Cheriet, M., Said, J.N. and Suen, C.Y., 1998. A recursive thresholding technique for image segmentation. IEEE transactions on image processing, 7(6), pp.918-921.

Cover Technologies Inc, 2021. Cover. [online] buildcover.com. Available at: https://buildcover.com/ [Accessed 2 May 2022].

Espindola, G.M., Câmara, G., Reis, I.A., Bins, L.S. and Monteiro, A.M., 2006. Parameter selection for region-growing image segmentation algorithms using spatial autocorrelation. International Journal of Remote Sensing, 27(14), pp.3035-3040.

F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in 2016 Fourth International Conference on 3D Vision (3DV). IEEE, 2016, pp. 565–571.

Farabet, C., Couprie, C., Najman, L. and LeCun, Y., 2012. Learning hierarchical features for scene labeling. IEEE transactions on pattern analysis and machine intelligence, 35(8), pp.1915-1929.

Finch (2020). FINCH3D Software. [online] Finch3d.com. Available at: https://finch3d.com/ [Accessed 2 May 2022].

G. Wang, W. Li, S. Ourselin, and T. Vercauteren, "Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks," in International M.I.C.C.A.I. Brainlesion Workshop. Springer, 2017, pp. 178–190. H. Noh, S. Hong, and B. Han, "Learning deconvolution network for semantic segmentation," in Proceedings of the IEEE international conference on computer vision, 2015, pp. 1520–1528.

Han, L., Tao, P. and Martin, R.R., 2019. Livestock detection in aerial images using a fully convolutional network. Computational Visual Media, 5(2), pp.221-228.

Iqbal, T. and Ali, H., 2018. Generative adversarial network for medical images (MI-GAN). Journal of medical systems, 42(11), pp.1-11.

Kazeminia, S., Baur, C., Kuijper, A., van Ginneken, B., Navab, N., Albarqouni, S. and Mukhopadhyay, A., 2020. GANs for medical image analysis. Artificial Intelligence in Medicine, 109, p.101938.

Kim, J.U., Kim, H.G. and Ro, Y.M., 2017, July. Iterative deep convolutional encoder-decoder network for medical image segmentation. In 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (E.M.B.C.), pp. 685-688.

L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Semantic image segmentation with deep convolutional nets and fully connected crfs," arXiv preprint arXiv:1412.7062, 2014.

Lafferty, J., McCallum, A. and Pereira, F.C., 2001. Conditional random fields: Probabilistic models for segmenting and labeling sequence data.

Liu, Z., Li, X., Luo, P., Loy, C.C. and Tang, X., 2017. Deep learning markov random field for semantic segmentation. IEEE transactions on pattern analysis and machine intelligence, 40(8), pp.1814-1828.

Long, J., Shelhamer, E. and Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition pp. 3431-3440.

Luc, P., Couprie, C., Chintala, S. and Verbeek, J., 2016. Semantic segmentation using adversarial networks. arXiv preprint arXiv:1611.08408.

Mardia, K.V. and Hainsworth, T.J., 1988. A spatial thresholding method for image segmentation. IEEE transactions on pattern analysis and machine intelligence, 10(6), pp.919-927.

Muthukrishnan, R. and Radha, M., 2011. Edge detection techniques for image segmentation. International Journal of Computer Science Information Technology, 3(6), p.259.

O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in International Conference on Medical image computing and computer-assisted intervention. Springer, 2015, pp. 234–241.

Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE transactions on systems, man, and cybernetics, 9(1), pp.62-66.

Preetha, M.M.S.J., Suresh, L.P. and Bosco, M.J., 2012, March. Image segmentation using seeded region growing. In 2012 International Conference on Computing, Electronics and Electrical Technologies (I.C.C.E.E.T.) (pp. 576-583). IEEE.

Schwing, A.G. and Urtasun, R., 2015. Fully connected deep structured networks. arXiv preprint arXiv:1503.02351.

SearchLand Ltd (2020). SearchLand - All U.K. Property Data in One Place. [online] searchland.co.uk. Available at: https://searchland.co.uk/ [Accessed 2 May 2022].

Shih, F.Y. and Cheng, S., 2005. Automatic seeded region growing for color image segmentation. Image and vision computing, 23(10), pp.877-886.

Souly, N., Spampinato, C. and Shah, M., 2017. Semi supervised semantic segmentation using generative adversarial network. In Proceedings of the IEEE international conference on computer vision (pp. 5688-5696).

TheCodingBug, 2021. Semantic and instance segmentation differences, Available at :https://www.reddit.com/r/learnmachinelearning/comments/kt0hov/ [Accessed April 20, 2022].

V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," IEEE transactions on pattern analysis and machine intelligence, vol. 39, no. 12, pp. 2481–2495, 2017.

Vemulapalli, R., Tuzel, O., Liu, M.Y. and Chellapa, R., 2016. Gaussian conditional random field network for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3224-3233).

Wurm, M., Stark, T., Zhu, X.X., Weigand, M. and Taubenböck, H., 2019. Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks. I.S.P.R.S. journal of photogrammetry and remote sensing, 150, pp.59-69.

Zheng, S., Jayasumana, S., Romera-Paredes, B., Vineet, V., Su, Z., Du, D., Huang, C. and Torr, P.H., 2015. Conditional random fields as recurrent neural networks. In Proceedings of the IEEE international conference on computer vision (pp. 1529-1537).

Zheng, X., Lei, Q., Yao, R., Gong, Y. and Yin, Q., 2018. Image segmentation based on adaptive K-means algorithm. E.U.R.A.S.I.P. Journal on Image and Video Processing, 2018(1), pp.1-10.