**Paper: PolyWorld: Polygonal Building Extraction with Graph Neural Networks in Satellite Images**

**Performance**

PolyWorld presents a notable advancement in the extraction of vector representations of buildings from aerial and satellite imagery, by directly predicting the vertices and constructing precise polygons. Utilizing both CNN and GNN architectures, PolyWorld outperforms similar models, such as the FFL approach, in terms of mean Intersection over Union (IoU) and Combined IoU, scoring 9.12 and 0.88 respectively. These metrics indicate a high degree of accuracy in the model's ability to segment and delineate building polygons accurately, reflecting its efficacy in producing clean and regular polygons. The end-to-end nature of the model, which integrates vertex detection and polygonization into a single process, reduces the propagation of errors typically seen in multi-stage methods, contributing to its superior performance.

**Limitations**

Despite its strengths, PolyWorld exhibits several limitations. Like many deep learning models, it requires considerable computational resources and time for training. The model struggles with buildings that have internal holes or are cut off at image boundaries. Buildings with holes are challenging as they might be misinterpreted or entirely missed in the polygonization process, while those cut off by image edges may inaccurately be represented as complete structures.

**Areas for Improvement**

To enhance PolyWorld's capabilities, several improvements can be considered:

1. Handling Buildings with Holes: Incorporating additional layers or mechanisms that specifically target the identification and correct handling of buildings with internal holes could improve accuracy.

2. Edge Case Management: Improving the model’s ability to recognize and accurately represent buildings that are partially visible at the edges of an image is crucial. This could involve training the model with augmented data that specifically includes these scenarios or developing algorithms that infer the likely continuation of building boundaries beyond the visible area.

3. Multi-Level Structure Discrimination: Addressing the challenge of distinguishing between buildings layered on top of one another in urban environments is crucial for PolyWorld. This improvement could involve incorporating elevation data or depth sensors to differentiate between various building levels accurately. Further, adapting the model with 3D convolutional networks could enhance spatial awareness, enabling it to recognize and delineate structures vertically. Additionally, enriching the training datasets with images of complex, multi-level urban scenarios would allow the model to better learn and generalize in environments with vertical architectural diversity, thus enhancing its effectiveness in detailed urban mapping and planning.

**Conclusion**

PolyWorld is a significant step forward in the field of building extraction and polygonization from aerial images, demonstrating robust performance and generating high-quality outputs. However, addressing its limitations related to the handling of specific architectural features and optimizing its training process are crucial for enhancing its practicality and accuracy in real-world applications.

Figure1, 2, 3: Input image vs output results

A aerial view of houses and a road

Description automatically generatedA group of yellow squares

Description automatically generated

Aerial view of a city

Description automatically generatedA group of purple squares

Description automatically generated

Aerial view of a building

Description automatically generatedA red and white rectangle

Description automatically generated

Figure 4: Model performance

A screenshot of a computer

Description automatically generated

Figure 5:Our mean IOU and CIOU is 9.12, 0.88, which is higher than the FFL approach A table with numbers and a number of polygons

Description automatically generated

**Code Used**

A screenshot of a computer program

Description automatically generated