# INTISARI

Bangunan sebagai data spasial berupa *footprint* merupakan fitur yang dibutuhkan untuk berbagai macam keperluan. Salah satu manfaat *footprint* bangunan adalah untuk menghasilkan model bangunan 3D melalui proses pemodelan berbasis ekstrusi *footprint*. *Footprint* bangunan biasanya diperoleh dari proses digitasi manual. Namun, digitasi manual umumnya membutuhkan waktu yang lama. Selain itu, hasil digitasi juga inkonsisten karena tergantung pada kemampuan operator. Untuk mengatasi hal tersebut, terdapat teknik yang mampu menghasilkan *footprint* secara otomatis, cepat, dan konsisten, yaitu *deep learning* berbasis Mask R-CNN. Dibandingkan dengan model *deep learning* yang lain, Mask R-CNN memiliki beberapa keunggulan, yaitu arsitektur yang sederhana, fleksibilitas untuk mengekstraksi objek yang bervariasi, dan kemampuan dalam membedakan objek-objek sejenis yang relatif rapat. Pada penelitian ini, dilakukan ekstraksi *footprint* bangunan secara otomatis menggunakan Mask R-CNN untuk pemodelan bangunan 3D berbasis ekstrusi *footprint*.

Penelitian dilakukan di wilayah kampus Universitas Riau, Kota Pekanbaru. Ekstraksi *footprint* bangunan dilakukan pada *orthomosaic* hasil pengolahan foto udara yang diakuisisi menggunakan UAV. Proses ekstraksi dilakukan dengan terlebih dahulu mendesain sejumlah model Mask R-CNN berdasarkan kombinasi jumlah *input channel*, dimensi *feature map*, dan jumlah *hidden layer*. Model bangunan 3D diperoleh dengan memanfaatkan *footprint* hasil ekstraksi terbaik dari seluruh model Mask R-CNN yang telah didesain. Untuk meningkatkan tingkat kedetailan model, dilakukan perbaikan LoD secara manual dengan memanfaatkan data foto udara yang diakuisisi dengan kaidah fotogrametri jarak dekat terhadap 3 sampel bangunan yang dipilih, yaitu Gedung Rektorat, Gedung Dekanat FAPERIKA, dan Gedung "C" Fakultas Teknik. *Footprint* bangunan hasil ekstraksi dievaluasi melalui kriteria nilai *precision*, *recall*, *f1-score*, dan indeks IoU, sedangkan model bangunan 3D yang dihasilkan dievaluasi berdasarkan tingkat kedetailan model, segmentasi atap (berdasarkan nilai *correctness*, *completeness*, dan *quality*), serta secara geometrik dengan menghitung nilai RMSE yang mengacu kepada standar CityGML dalam LoD 2 dan 3.

Dari 21 model yang telah didesain, Model Mask R-CNN dengan kriteria 3 *input channels* (RGB), 1024 dimensi *feature map*, dan 50 *hidden layers* (ResNet-50 menghasilkan *footprint* yang paling akurat. Model tersebut memiliki akurasi *training* senilai 96,80% setelah melewati 25 *epochs*, dengan akurasi ekstraksi *footprint* berdasarkan nilai *precision* 91,78%, *recall* 90,29%, *f1-score* 91,03%, dan indeks IoU 92,18%. Secara keseluruhan, *footprint* mampu mengikuti bentuk bangunan dengan cukup presisi. Proses ekstrusi *footprint* menghasilkan model bangunan 3D pada LoD 2 dengan evaluasi segmentasi atap berdasarkan *correctness* senilai 93,72%, *completeness* 89,44%, dan *quality* 83,86%. Perbaikan model bangunan 3D LoD 2 pada 3 sampel bangunan menghasilkan model LoD 3 dengan detail atap dan dinding yang lebih akurat. Model LoD 2 memiliki RMSE senilai 1,112 m, dan Model LoD 3 memiliki RMSE senilai 0,312 m, 0,222 m, dan 0,203 m untuk masing-masing Gedung Rektorat, Gedung Dekanat FAPERIKA, dan Gedung "C" sebagai sampel. Model bangunan 3D LoD 2 dan LoD 3 yang dihasilkan memenuhi standar CityGML, berdasarkan tingkat kedetailan dan geometriknya.

Kata kunci: Bangunan, Mask R-CNN, LoD, Model 3D, UAV

# ABSTRACT

Buildings as spatial data in the form of footprints are features that are needed for various purposes. One of the benefits of building footprints is to generate 3D building models through a footprint extrusion-based modeling process. Building footprints are usually obtained from the manual digitizing process. However, manual digitizing generally takes a long time. In addition, the digitization results are also inconsistent because it depends on the ability of the operator. To overcome this, there is a technique that can generate footprints automatically, quickly, and consistently, namely Mask R-CNN-based deep learning. Compared to other deep learning models, Mask R-CNN has several advantages, specifically a simple architecture, flexibility to extract a variety of objects, and the ability to distinguish objects that are spatially close to each other. In this study, the building footprint was extracted automatically using Mask R-CNN for footprint extrusion-based 3D building modeling.

The research was conducted in the campus area of Riau University, Pekanbaru City. Building footprints extraction was executed on orthomosaic as a result of aerial imageries processing that was acquired using a UAV. The extraction process was carried out by first designing several Mask R-CNN models based on combinations of the number of input channels, feature map dimension, and the number of hidden layers. The 3D building models were obtained by utilizing the best-extracted footprint from all Mask R-CNN models that have been designed. To increase the level of model detail, the LoD was improved manually by utilizing aerial imageries acquired with close-range photogrammetry rules for the 3 selected building samples, namely the Rectorate Building, Dekanat FAPERIKA Building, and the "C" Building of Engineering Faculty. The extracted building footprints were evaluated through the criteria of precision, recall, f1-score, and IoU index, while the resulting 3D building models were evaluated based on the level of model detail, roof segmentation (based on correctness, completeness, and quality values), and geometrically by calculating the RMSE value which refers to CityGML standard in LoD 2 and 3.

From 21 models that have been designed, the Mask R-CNN Model with the criteria of 3 input channels (RGB), 1024 feature map dimensions, and 50 hidden layers (ResNet-50) extracted the most accurate footprint. The model has a training accuracy of 96.80% after passing 25 epochs, with an accuracy of footprint extraction based on the precision of 91.78%, recall of 90.29%, f1-score of 91.03%, and IoU index of 92.18%. Overall, the footprint can follow the shape of the building with considerable precision. The footprint extrusion process produced 3D building models at LoD 2 with an evaluation of roof segmentation based on the correctness of 93.72%, completeness of 89.44%, and quality of 83.86%. Improvements to the LoD 2 models on 3 building samples resulted in LoD 3 models with more accurate roof and wall details. LoD 2 models have RMSE of 1.112 m, and Model LoD 3 has RMSE of 0.312 m, 0.222 m, and 0.203 m for the Rectorate Building, Dekanat FAPERIKA Building, and Building “C” as samples, respectively. The resulting LoD 2 and LoD 3 building models meet CityGML standards, based on their level of detail and geometry.

Keywords: Building, Mask R-CNN, LoD, 3D Model, UAV