


現実らしい架空都市の合成データセットの自動生成法を開発

工学研究科 環境エネルギー工学専攻

准教授 福田 知弘

 <https://researchmap.jp/tomohirofukuda>

研究の概要

AIの性能や期待される精度が向上するにつれて、学習に必要なデータセットの規模は大きくなる傾向があります。一方、データセットの規模が大きくなるにつれて、データセットを作成するためにかかるコストが増え、学習に用いるデータの不足が予想されています。

筆者らは、バーチャルオブジェクトや3Dデジタルツインモデルと画像を組み合わせた、合成データセットの生成方法を研究開発してきました (Zhang et al. 2022)。しかし、実在する都市データを事前に用意が必要があることや、作成できるデータのバリエーションが有限であるという課題がありました。

そこで本研究は、現実的でありながら実在しない都市の3Dモデルを生成できるプロシージャルモデリングと現実的な画像を生成できるディープラーニングの画像変換技術を統合して、建物を個別検出するための合成データ自動生成法を開発しました。

研究の背景と結果

図1は、プロシージャルモデリングと画像変換技術を用いて、大量の学習データを自動的に生成する概要フローを示します。生成したい都市のパラメータを入力するだけで、学習データを生成する基になる3D都市モデルが生成され、そのモデルをもとに大量の学習データが生成されます。生成された大量の学習データを用いてディープラーニングモデルを学習させ、実世界の画像から対象物を検出することが最終ゴールです。

もう少し具体的に説明します。Step 1では3D都市モデルを自動生成するために都市の特徴を定めるパラメータを入力します。Step 2ではStep 1で生成した3D都市モデルと画像サイズなどの設定を入力し、ゲームエンジン上でアノテーションデータとStep 3で景観画像データを生成するために必要なセグメンテーション画像のペアを生成します。Step 3では画像変換技術を用いてStep 2で生成したセグメンテーション画像を現実らしい景観画像に変換します。そして、Step 4でアノテーションデータと景観画像を関連付けます。

図2は、提案したフレームワークを用いて開発したプロトタイプシステムで合成データセットを生成し、それを学習したインスタンスセグメンテーションモデルを用いて建物を個別検出した結果を含んでいます。図2の2行目は筆者らが提案した方法で生成したデータセットを用いて学習したインスタンスセグメンテーションモデルによる現実世界の検出結果であり、従来法である現実世界の画像を手作業により作成したデータセットで学習したモデル (図2の3行目) と同程度、もしくは、より優れた検出結果を得られたことを確認しました。

研究の意義と将来展望

本稿で紹介する方法は、データセットの準備コストを大幅に削減できるとともに、都市景観のような複雑な構造を内包する画像を対象とする場合でも、現実らしいが架空の都市モデルを用いた合成データを適用できることを示唆しています。これは、都市景観分野に限らず、AIの活用が期待される他の分野にも適用可能なアプローチであると考えています。

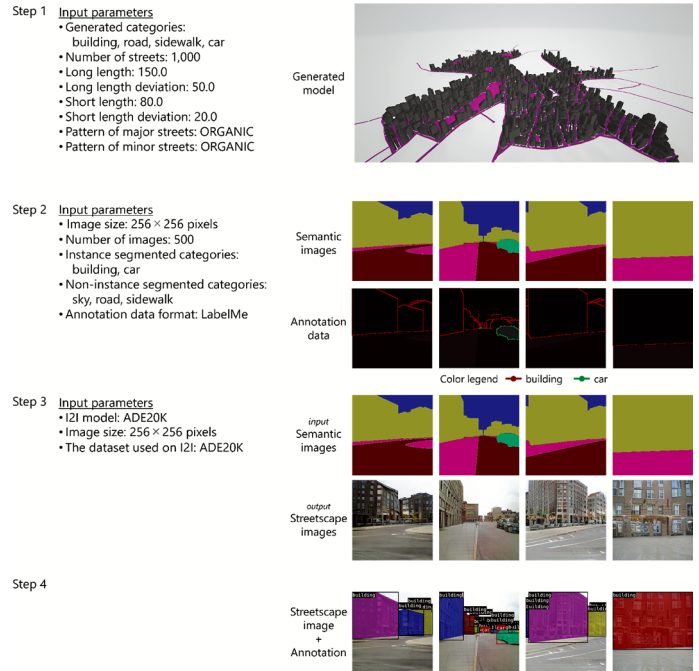


図1 提案方法の概要



図2 提案法で生成したデータセットを用いて学習したモデル (2行目) と、現実画像を用いて学習したモデル (3行目) の検出精度の比較。

特 許	Kikuchi, Takuya; Fukuda, Tomohiro; Yabuki, Nobuyoshi. Development of a synthetic dataset generation method for deep learning of real urban landscapes using a 3D model of a non-existing realistic city. Advanced Engineering Informatics. 2023, 58, 102154. doi:10.1016/j.aei.2023.102154
論 文	Zhang, Jiaxin; Fukuda, Tomohiro; Yabuki, Nobuyoshi. Automatic generation of synthetic datasets from a city digital twin for use in the instance segmentation of building facades. Journal of Computational Design and Engineering. 2022, 9(5), 1737-1755. doi:10.1093/jcde/qwac086
参考URL	https://resou.osaka-u.ac.jp/ja/research/2023/20230921_3 https://y-f-lab.jp/
キーワード	AI、合成データ、建造環境、デジタルアセット、システム化