

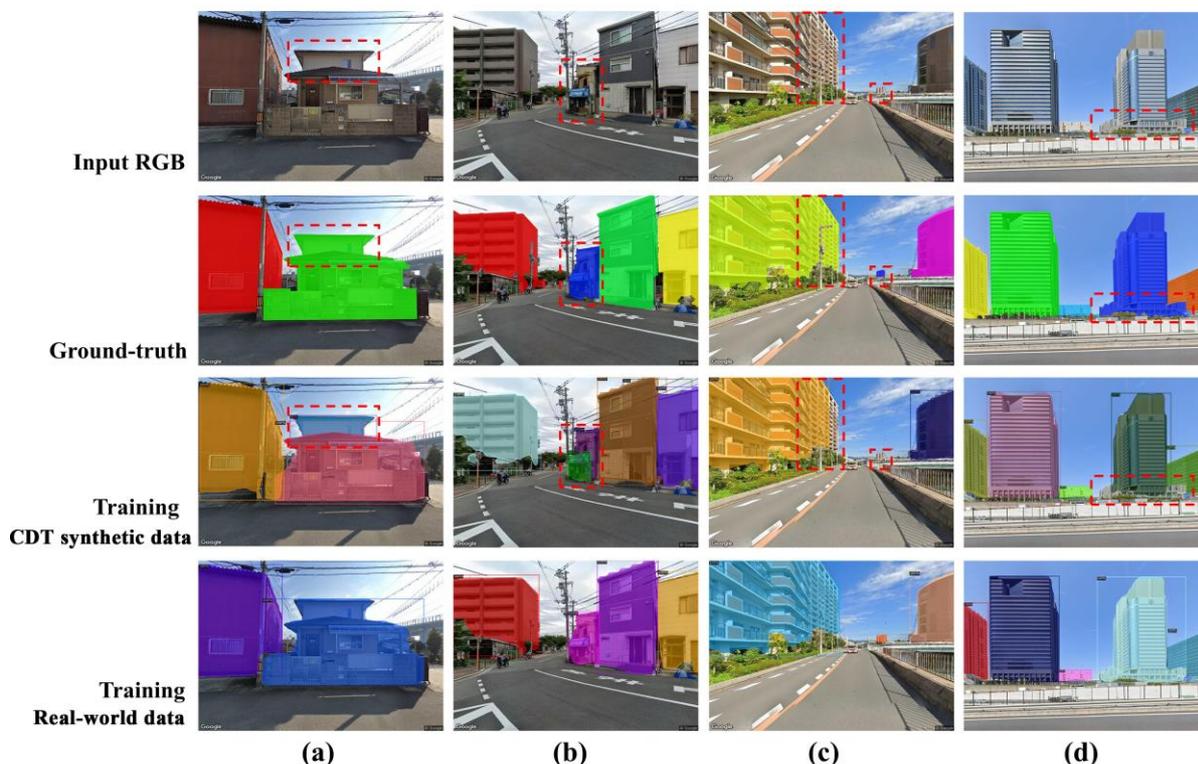
分野：工学系

キーワード：合成データセット、都市デジタルツイン、インスタンスセグメンテーション、建物ファサード、深層学習、SDGs

都市デジタルツインから合成データセットを自動生成 —建物ファサードのインスタンスセグメンテーションの精度向上—

【研究成果のポイント】

- ◆ 都市デジタルツイン(CDT:City Digital Twin)^{※1}を用いて、建物ファサードのインスタンス注釈付き合成データ^{※2}を自動生成する方法を開発しました。
- ◆ CDT 合成データでネットワークを学習させた場合、実世界に対応する建物が存在しない仮想データで学習した場合よりも建物ファサードのインスタンスセグメンテーション^{※3}の抽出精度を向上させることができました。
- ◆ 提案した CDT 合成データを拡張して現実世界のデータと混合することでインスタンスセグメンテーションの性能を向上させることができました。
- ◆ 提案した自動化システムによりデータセットの作成コストを大幅に削減することができます。



都市デジタルツイン(CDT)による合成データと現実世界のデータを学習させ、建物カテゴリにおいて個々の建物ファサードをインスタンスセグメンテーションした結果:(a)戸建て住宅、(b)低層建築物、(c)中層集合住宅、(d)高層建築物、上段:入力写真、中上段:正解画像、中下段:CDT 合成データで学習させインスタンスセグメンテーションした結果、下段:現実世界のデータで学習させインスタンスセグメンテーションした結果 (赤破線の矩形:街並みの画像で、建物ファサードをインスタンスセグメンテーションする際に失敗した箇所を示す)。

❖ 概要

大阪大学大学院工学研究科の Jiaxin Zhang さん(博士後期課程)、福田知弘准教授、矢吹信喜教授らの研究グループは、都市の 3 次元デジタルアセットをゲームエンジンでレンダリングし、建物ファサードをインスタンスセグメンテーションするための画像とそのアノテーションを自動生成する合成データ作成システムを開発しました。

都市環境の情報基盤を整備するために、建物ファサードデータの抽出と統合は必要な作業のひとつです。しかしながら、既存のセマンティックセグメンテーションに基づく建物ファサードの解析方法では、つながっている建物を個々のインスタンスに区別することが困難でした。また、大規模なデータセットに含まれる建物ファサードデータを手作業で収集し、アノテーションを施すことは時間と労力がかかる作業です。

近年、都市デジタルツイン(CDT)の開発・活用により、国土交通省の PLATEAU [プラトー]など、大量の高品質な建物デジタルアセットが作成されています。これらのデジタルアセットを使用することで、現実世界を対象としたデータセットに代わる高品質で費用対効果の高い合成データセットを生成し、教師あり学習による建物ファサードのインスタンスセグメンテーションの学習セットとすることが可能になります。そこで本研究では、CDT から合成データセットを自動的に生成するための新しいフレームワークを開発しました。実験では、2 種類の合成データ(CDT ベースと仮想ベース=実世界に対応する建物が存在しない)を比較して、CDT 合成データは仮想合成データに比べて、現実世界の画像を用いたディープラーニングの学習を後押しする効果があることを明らかにしました。また、現実データの一定量を提案する CDT 合成データと入れ替えた場合に、現実世界の学習セットを使用した場合に達成できる性能とほぼ一致することを確認しました。

図 1 は、従来の手動ラベリングと自動ラベリングによる建物ファサードのインスタンスセグメンテーションの学習プロセスを比較したものです。本研究では、手動によるラベリング作業を自動生成システムに置き換えて、CDT データをより効率的に作成し、DCNN(畳み込みニューラルネットワーク)の学習を行うことを目指しています。

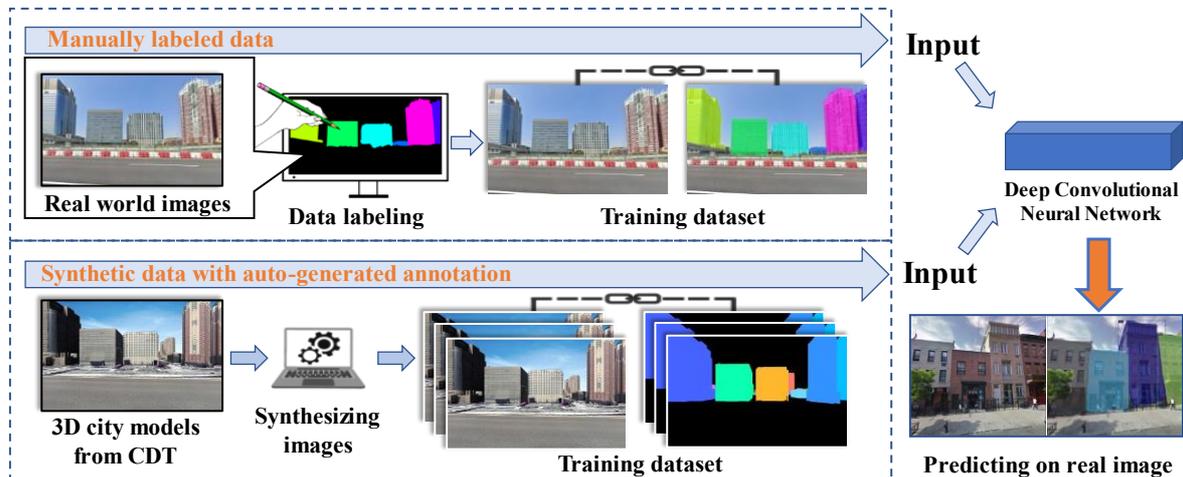


図 1 手動でアノテーションしたデータセットと自動生成した合成データセットの比較。従来方法では学習セットを作成する際に手作業でラベル付けする必要があったが、提案システムでは CDT のデジタルアセットを利用してインスタンスアノテーションを付与した合成データを自動生成できる。

図 2 は、東京のある地域の 3 次元都市モデルを俯瞰したものです。選択した 3 次元都市モデルは、国土交通省の PLATEAU [プラトー]プロジェクトのデータベースから CityGML 2.0 と Filmbox (fbx)形式で無償ダウンロードすることができ、後者は本研究で使用したフォーマットです。3 次元モデルには地理情報が含まれているため、衛星写真や地形図をロードすることで仮想世界と一致させることができます。

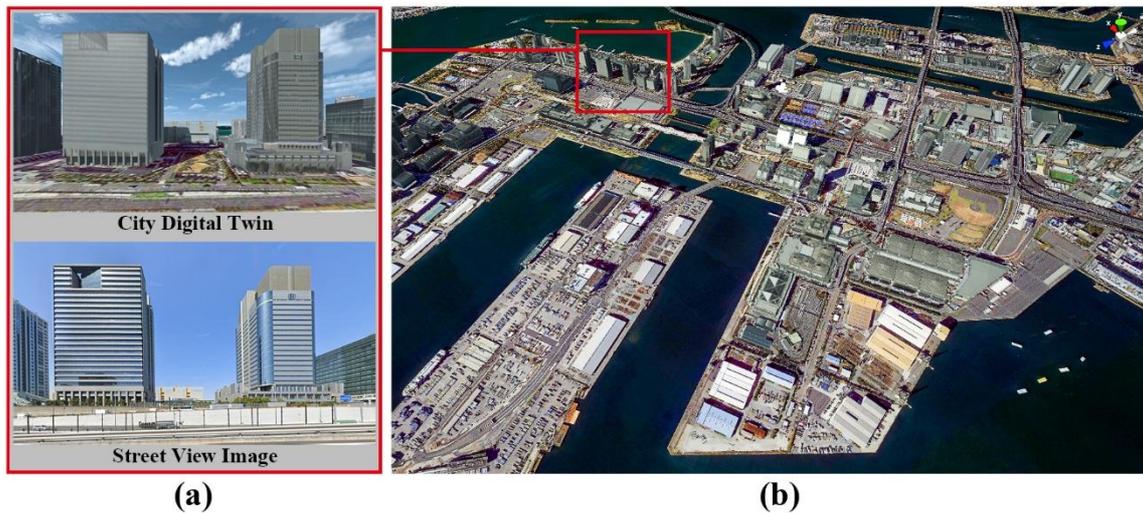


図 2 調査対象地域の 3 次元都市モデル。(a) CDT(上)と現実世界のストリートビュー(下)の同じ視点からのビュー(東京都湾岸道路(緯度:35.6283、経度:139.7782、2021 年 3 月))。(b) 都市デジタルツインの俯瞰。

以前の研究では、合成データセットのみを CNN の学習セットとして使用した場合、現実世界の街並みをセグメンテーションした精度は満足のものではないことが示されていました。そこで本研究では、この問題を解決するために、学習段階で、現実世界と合成データ両方の建物ファサード画像とそのアノテーションを含むハイブリッドデータセット(HSRBFIA: Hybrid collection of Synthetic and Real-world Building Facade Images and Annotations)の作成方法を提案しました。

図 3 は、異なるスケールの現実データと合成データを混合して用いて、建物ファサードをセグメンテーションした定性結果を示しています。CDT で生成した合成データを用いて学習させることで、低層ビル、高層ビル、複合施設などを十分に認識できることを確認しました。また、現実データと合成データの組み合わせ(HSRBFIA-60)により、建物でない要素(図 3a)、低層・中高層建物(図 3b、図 3c)、複雑な建物ファサード(図 3d)まで高精度に認識できることを確認しました。一方、ストリートビュー画像において、カメラから遠い建物や建物ファサードが複雑な構成である場合など、建物ファサードのインスタンスセグメンテーションに失敗しやすい部分を赤色の破線枠で表示しています。

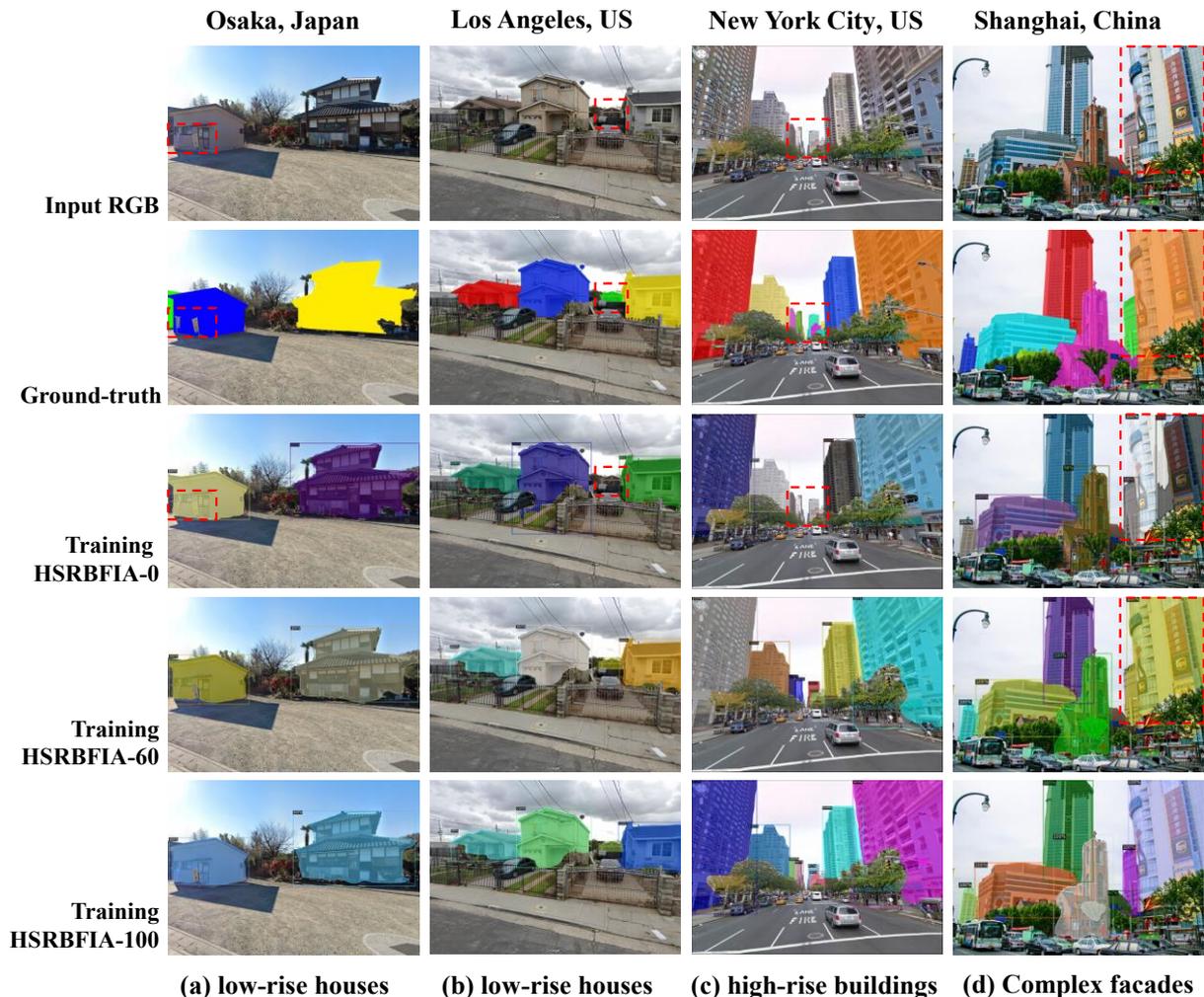


図3 合成データと実データの混合比率を変えた HSRBFIA データセットを用いて、建物の種類や大きさ別にインスタンスセグメンテーションした定性結果：(a) 大阪の戸建住宅、(b) ロサンゼルス市の戸建住宅、(c) ニューヨーク市の中高層建物、(d) 上海市の複合施設。(HSRBFIA-60:現実データ 60%、合成データ 40%を混合したデータセットであることを意味する。赤色の破線:ストリートビュー画像の中で、建物ファサードのインスタンスセグメンテーションで失敗しやすい箇所を示している)。

❖ 本研究成果が社会に与える影響（本研究成果の意義）

- 従来のデータラベリングの方法は、手作業に頼っていましたが、本研究では、DCNN(畳み込みニューラルネットワーク)の学習のために CDT(都市デジタルツイン)合成データを作成する際に、手動によるラベリング作業を自動生成システムで代替することを試みました。提案方法は、各画像のアノテーションを手作業で行う場合に比べて、約 1/2050 の時間で済み、データのアノテーション作業に必要なコストを大幅に削減することができます。
- 現実世界の画像、合成画像、それらを混合したハイブリッドデータセットに対する DCNN の学習結果を比較した結果、提案する合成データで学習セットを拡張することにより、現実世界の画像で建物ファサードをインスタンスセグメンテーションした場合の精度を向上させることができます。すなわち、CDT 合成データセットに一定割合の現実世界のデータを混合すると、インスタンスセグメンテーションの精度が大幅に向上し、現実世界のデータを 100%使用した場合と同程度の性能になることを確認しました。このことは、提案する合成データセットが、現実世界の画像を学習セットと置き換えることができる可能性を示しています。

Press Release

- ・ 複数の都市を対象として検証した結果、提案したフレームワークの移植性を確認することができました。本データセットは、ほとんどの現代建物の様式に対して、インスタンスセグメンテーションの有効な精度を得ることができます。一方、特徴的な建物様式や高密度に集積する街路環境では、インスタンスセグメンテーションの精度を向上させる必要があります。
- ・ 提案手法は、都市情報モデリングとデジタルアセットを有効に活用した CDT に基づく合成データセットを生成します。CDT がさらに開発・改良されれば、本研究の枠組みを都市環境の他の要素に適用することができ、デジタル基盤のさらなる発展において、それらの意味情報を充実させることが可能となるでしょう。

❖ 特記事項

本研究成果は、2022 年 8 月 26 日(金)(日本時間)に、学術雑誌「Journal of Computational Design and Engineering」(Oxford University Press)にオンライン掲載されました。

タイトル: “Automatic generation of synthetic datasets from a city digital twin for use in the instance segmentation of building facades (建物ファサードのインスタンスセグメンテーションに用いるための都市デジタルツインからの合成データセットの自動生成)”

著者名: Jiaxin Zhang, Tomohiro Fukuda, Nobuyoshi Yabuki (大阪大学 大学院工学研究科 環境エネルギー工学専攻)

DOI: <https://doi.org/10.1093/jcde/qwac086>

なお、本研究は、日本学術振興会 科学研究費助成事業 基盤研究の一環として行われました。

❖ 用語説明

※1 都市デジタルツイン (CDT: City Digital Twin)

都市のデジタルツイン(現実世界のデジタルの双子)。高い詳細度(LOD)を有する都市デジタルツイン(CDT)のデジタルアセットは、物理世界に存在するもののコピーであり、現実世界の情報を正確に反映したものである。

※2 合成データ (Synthetic Data)

コンピュータのシミュレーションやアルゴリズムから人工的に作成されたデータ

※3 インスタンスセグメンテーション (Instance Segmentation)

画像中の物体に対して、クラスラベルを予測し、一意の ID を付与する技術。同じ種類(クラス)に含まれるオブジェクトを別々に認識することができる。

❖ 関連するこれまでの研究成果

- ・建物取り壊し後の景観シミュレーションをリアルタイムに(2022/7/29)
https://resou.osaka-u.ac.jp/ja/research/2022/20220729_1
- ・障害物を除去し建物ファサードの全景を効率的に復元する方法を開発(2021/9/7)
https://resou.osaka-u.ac.jp/ja/research/2021/20210907_1
- ・建物の機能分類(住宅・オフィス・商業施設等)と 建物ファサードカラー測定による大都市スケールの空間分析法(2021/9/6)
https://resou.osaka-u.ac.jp/ja/research/2021/20210906_2
- ・航空写真の薄雲を自動で除去 深層学習用建物マスク画像の自動生成(2021/8/20)
https://resou.osaka-u.ac.jp/ja/research/2021/20210820_1
- ・景観シミュレーション用の MR に新技術(2021/3/24)
https://resou.osaka-u.ac.jp/ja/research/2021/20210324_1
- ・景観シミュレーション用の DR に新技術(2020/6/17)
https://resou.osaka-u.ac.jp/ja/research/2020/20200617_1



大阪大学
OSAKA UNIVERSITY

国立大学法人 大阪大学

〒565-0871 大阪府吹田市山田丘 1-1

TEL: 06-6877-5111 (代)

www.osaka-u.ac.jp

Press Release

❖ SDGs 目標



❖ 参考 URL

福田 知弘 准教授 研究者総覧 URL <https://rd.iai.osaka-u.ac.jp/ja/d2782e4b9c864b39.html>

【研究者のコメント】

本研究は、環境工学分野(環境設計)とコンピュータサイエンス(情報学)にまたがる学際領域で、実問題を解決するために、Zhang Jiaxin(張 嘉新)さんや矢吹信喜教授と取り組んだ内容です。

平たく言えば、今回の研究は「深層学習のための道具」を作ったものです。深層学習(インスタンスセグメンテーションのディープニューラルネットワーク)自体が、都市の建物を個々に検出するための道具ですので、今回の研究はそのための道具、すなわち、二次的道具の研究になります。都市景観検討における課題の整理と、その課題を解決するための、深層学習(インスタンスセグメンテーション)、ゲームエンジン、ストリートビューなど幅広い分野の知見を理解しながら実装と実験を行い、一定の研究成果としてまとめ上げることには苦勞しました。

研究成果が、生活の質向上に向けて、まちづくりや景観・環境アセスメント分野をはじめとする社会実装や、建築・建設分野のDX(デジタルトランスフォーメーション)に関する産業化につながれば幸いです。