

# 合成人口の精緻化に向けた U-Net を用いた建物用途推定

○福嶋竜希 原田拓弥 大内紀知 (青山学院大学)

## U-Net Estimating Building Usages for Finer Synthetic Population

\* T. Fukushima, T. Harada, and N. Ouchi (Aoyama Gakuin University)

**概要**— 社会シミュレーションにおける Facsimile Model の構築には、しばしば対象地域の世帯やその住所情報が求められる場合がある。先行研究では、実際の統計の特徴と整合する世帯のデータが作られた。しかし、世帯に住所を与える際、無作為に建物へ割り当てている。本研究では、世帯を精緻に建物へ割り当てるため、深層学習モデルのひとつである U-Net を用いて航空写真内の建物の用途を推定する。

**キーワード:** Real Scale Social Simulation, Deep Learning, Synthetic Population

### 1 研究背景

社会シミュレーションのモデルは、Abstract Model, Middle Range Model, Facsimile Model の3つに大別できる。Facsimile Model はシミュレーション対象地域の情報を可能な限り組み込んだ人工社会を用いるため、施策の意思決定支援が可能となる。しかし、人工社会構築のために、実世界の人間のデータを用いることはプライバシーを侵害し得る。

原田らは年齢や職業、世帯構成の分布などの国勢調査の統計的性質を有する日本全国の仮想的な人口データ(合成人口)を合成した。原田らの合成人口を用いた社会シミュレーションとして、江尻ら<sup>2)</sup>の研究がある。江尻ら<sup>2)</sup>は、大阪府阿倍野区の合成人口を用いて、AEDの使用率向上のための施策として、傷病者が発見される確率・発見者がAEDの運搬を試みる確率・AEDを新たに設置する建物の割合を変化させて、各施策の効果を評価している。

しかし、原田らは合成人口に住所を与える際、世帯に建物を無作為に割り当てているという問題がある。結果、居住が困難なほどに小さい建物に住む世帯や、住居ではない建物に居住する世帯が存在する。

本研究では、合成人口の住所として、適切な建物を割り当てるために、航空写真内の建物の用途を推定する。

### 2 建物用途の推定手法

国土地理院が提供する航空写真<sup>3)</sup>と宅地利用動向調査<sup>4)</sup>をデータとして用いる。ズームレベルは18とした宅地利用動向調査は、首都圏・近畿圏・中部圏の土地利用をまとめたもので、その用途の種類は17である。本研究で用いるデータの概要をFig. 1に示す。Fig. 1の左が航空写真、中央が対応する宅地利用動向調査のラスターデータである。これを右のようにグリッドで区切って、各セルに土地用途に応じた値をTable 1に従って与える。実際には、ピクセル単位で、Table 1右列のクラスラベルを与えている。その他のセルを90%以上含むラスターは除外した。これにより、計507,844のラスターデータが得られる。本研究は、航空写真を入力にとり、各ピクセルの土地利用を推定する機械学習モデルを学習させる。航空写真のピクセルとラスターデータのセルが対応することに注意されたい。推定器として、深層学習のモデルのひとつであるU-Net<sup>5)</sup>を用

いる。U-Netは畳み込み演算を繰り返し、入力画像の各ピクセルのラベルを推定する。つまり、入力と出力のサイズが同じである。誤差関数はCross Entropy Lossを用いる。

Cross Entropy Lossはクラス $c$ に対する重み $w^{(c)}$ を考慮できる。重みを全く考慮しない場合(Unweighted)と、訓練画像の教師ラベルから決定する場合(Weighted)の2通りを検証する。後者では、以下の式で重みを定める。

$$w^{(c)} = \frac{\sum_{c' \in S} N_{c'}}{|S|N_c} \quad (1)$$

ここで、 $S$ は全てのクラスの集合を表し、 $N_c$ はクラス $c$ のデータ数である。本研究では、画像の各ピクセルが教師ラベルをもつので、分子は全訓練画像のピクセル数の合計に等しい。 $w^{(c)}$ の値が大きいくほど、判別の誤差を大きく評価して学習する。

U-Netにより、建物毎ではなく、入力航空写真全体の出力が得られる。よって、建物の重心座標の値を抽出し、それを最終的な判別結果として評価する。

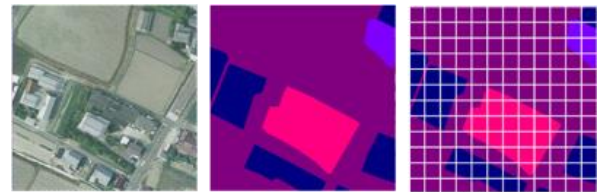


Fig. 1: 用いるデータの概要

### 3 実験結果

得られた計 507,844 データの 7 割を訓練データ、3 割をテストデータとした。Unweighted, Weighted いずれの場合も、バッチサイズは 32、エポック数は 110 とした。また、過剰適合抑制のため、Dropout Rate を 0.5 としている。クラス数の偏りがあるため、訓練データ内の低層住宅地、工業用地、商業・業務用地、公共公益施設用地、中高層住宅地のピクセル数が概ね等しくなるように画像の回転・平行移動を用いて訓練データを水増しした。式 (1) により算出された、Weighted で用いるクラスの重みを Table 2 に示す。

Unweighted の正解率は 0.813, Weighted の正解率は 0.811 であった. 次に, それぞれの混同行列を Table 3, Table 4 に示す. ただし, これらは各クラスの建物数で除してある. 商業・業務用地以外は, Weighted の方が良い結果となっている. 特に, 合成人口の割り当てで重要となる低層住宅地は 95%, 中高層住宅地は 71% 正しく判別している. 住居以外のクラスの精度は改善の余地がある. 公共施設の所在地のオープンデータを評価時に活用することで精度の向上が期待できる.

#### 4 結論と今後の課題

本研究は, 原田らの合成人口の精緻化のため, U-Net を訓練し, 建物用途を判別した. 居住に相応しい低層住宅地の建物は 95%, 中高層住宅地の建物は 71% の精度で判別が可能である.

U-Net の学習に関するハイパーパラメータの調整が今後の課題として挙げられる.

#### 参考文献

- 1) 原田, 村田, 柁井: 家族類型と世帯内の役割を考慮した SA 法による大規模世帯の合成, 計測自動制御学会論文集, 54-9, 705/717 (2018)
- 2) 国土地理院: 空中写真・衛星画像, <https://cyberjapandata.gsi.go.jp/xyz/ort/{z}/{x}/{y}.jpg> (2007)
- 3) 国土地理院: 数値地図 5000 (土地利用), <https://www.gsi.go.jp/kankyochiri/lum-5k.html>
- 4) 江尻, 中村, 原田, 大内: 住宅内における AED 利活用向上のためのシミュレーション分析, 計測自動制御学会システム・情報部門 学術講演会, 501/504 (2020)
- 5) O. Ronneberger, P. Fischer, T. Brox: U-Net: Convolutional Networks for Biomedical Image Segmentation, *Medical Image Computing and Computer-Assisted Intervention*, 234/241 (2015)

Table 1: 宅地利用動向調査の土地利用名と本研究で用いるクラスラベルの対応

宅地利用動向調査	本研究
一般低層住宅地	低層住宅地
密集低層住宅地	
中高層住宅地	中高層住宅地
工業用地	工業用地
商業・業務用地	商業・業務用地
その他の公共公益施設用地	公共公益施設
畑・その他の農地	その他
空地	
公園・緑地等	
田	
山林・荒地等	
その他	
造成中地	
道路用地	
河川・湖沼等	
海	
対象地域外	

Table 2: Weighted における各クラスの重み

ラベル	重み
低層住宅地	0.810
中高層住宅地	2.205
工業用地	1.241
商業・業務用地	1.422
公共公益施設	1.228
その他	0.503

Table 3: テストデータにおける規格化された混同行列 (Unweighted)

実際	予測					
	低層住宅地	中高層住宅地	工業用地	商業・業務地区	公共公益施設	その他
低層住宅地	0.946	0.004	0.005	0.017	0.003	0.026
中高層住宅地	0.134	0.674	0.010	0.099	0.043	0.040
工業用地	0.253	0.005	0.504	0.139	0.038	0.061
商業・業務用地	0.441	0.015	0.086	0.373	0.029	0.057
公共公益施設	0.148	0.026	0.057	0.106	0.583	0.080
その他	0.442	0.012	0.036	0.056	0.037	0.417

Table 4: テストデータにおける規格化された混同行列 (Weighted)

実際	予測					
	低層住宅地	中高層住宅地	工業用地	商業・業務地区	公共公益施設	その他
低層住宅地	0.949	0.005	0.005	0.011	0.002	0.027
中高層住宅地	0.134	0.712	0.010	0.063	0.043	0.038
工業用地	0.261	0.007	0.527	0.091	0.043	0.071
商業・業務用地	0.468	0.023	0.102	0.306	0.033	0.069
公共公益施設	0.154	0.031	0.058	0.072	0.598	0.087
その他	0.440	0.014	0.034	0.038	0.039	0.434