

# 現実に近い社会シミュレーション実施のための pix2pix を用いた建造物の用途の判別

田口 栞 ○原田 拓弥 大内 紀知 (青山学院大学)

## Examination of Discrimination of Building Usage

### Using pix2pix for Conducting a Near-Realistic Social Simulation

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

**概要一** 従来の社会シミュレーションは KISS 原理に基づいてモデル化されている。より具体的なシミュレーションの実現を目指すために必要な市民のデータは利用が困難である。建造物の用途の情報がないために公開されているデータを用いても建造物に割り当てられる世帯数が正確でない地域がある。本研究では世帯数の割り当てを行うために建造物の用途の判別を行いその評価の手法を検討する。

**キーワード:** 深層学習, pix2pix, 社会シミュレーション, 地図タイル

## 1 はじめに

社会シミュレーションの研究は1990年代ごろから世界中で始まった。社会シミュレーションの分析により社会現象の理解や未来の可能性を可視化することができる。多くの関心を集めている。特に個人に焦点をあてるモデルが広く使われている。社会シミュレーションの応用は多岐にわたり、社会的原理の解析や現実に似た状況を再現することに貢献している。この分野の技術は物理学や人工知能などを駆使し、一般的には KISS 原理 (Keep It Simple, Stupid) に基づいてモデル化される。KISS 原理<sup>1)</sup>とは、複雑さや難解さを避けて誰にでもわかるように簡潔にする取り組みのことである。しかし結果の解説には専門的な知識が必要であり、社会現象をより深く理解するためには、KISS 原理を逸脱することが必要である<sup>2)</sup>。したがって人口的・仮想的な集団を対象にした社会シミュレーションから、より具体的な組織レベル、社会レベルの社会シミュレーションを可能とする基盤の整備が求められる。

現実に近い社会シミュレーションの実現には環境や市民のデータを組み込み、可能な限り現実に忠実に再現しモデル化する必要がある。環境のデータは地理情報システムの活用が可能であるのに対して、市民のデータは個人の特定ができてしまうために利活用が困難である。近年、実際の個票を用いずに仮想的な個票を作成し人口個票を合成する研究がなされている<sup>3)</sup>。原田らは個人の年齢や親子の年齢差の統計表と仮想個票の誤差を最小化させ仮想個票を生成した。仮想都市を作り上げるために生成した仮想個票に世帯が住む位置情報の属性を付加する研究があるが、これは町ごとにパラメータの調整が必要である<sup>4)</sup>。

建造物の用途の判別を行う方法として Google が提供しているストリートビュー画像を用いた判別方法が存在する<sup>5)</sup>。しかし Google の利用規約変更により、ストリートビュー画像から情報を取得することができない。汎用的な手法で日本全国の人口を建造物ごとに割り当てるために深層学習のモデルの一つである pix2pix を用いて建造物の用途の推定を試みる。

pix2pix<sup>6)</sup>は深層学習の DCGAN (Deep Convolutional

Generative Adversarial Networks) の生成モデルの一つであり、例えばモノクロ画像とカラー画像などといった対となる画像のモノクロ画像を入力するとカラー画像との変換方法を入力し、カラー画像に生成される。それは画像を生成する Generator とその生成された画像が本物なのか偽物なのかを判断する Discriminator を互いに競わせることで画像が生成される。

本研究は具体的な社会シミュレーション実現のために必要とされる世帯の割り当てのための建造物の用途の判別を目標とする。既存研究のようにストリートビュー画像を用いた実験を行えないため pix2pix を用いた実験により作成された建造物の用途の判別を試みる。pix2pix では評価を行えないため画像の評価を行う既存の手法を紹介し、本研究では画像ではなく画像に含まれる建造物の用途の判別を行うために異なる評価手法を行う。

## 2 画像評価手法

既存の画像評価手法は MSE (平均二乗誤差) や PSNR (ピーク信号対雑音比)、SSIM (Structural Similarity)<sup>7)</sup> などがある。MSE とは画素値の差分を 2 乗し足していき、その合計の平均を求めたものである。PSNR とは信号が取りうる最大のパワー (精度比較の場合、元の画像の画素値) と劣化をもたらすノイズの比率を表すものである。SSIM とは、画像内の局所領域毎に平均、標準偏差、共分散を計算して SSIM を表し局所領域の平均値がその SSIM である。しかしこの 3 つはすべて元の画像と比較画像に対する画像の評価方法である。しかし本研究では、一つの画像に含まれる複数の建造物に対して個別に評価を行う。本研究では、pix2pix によって生成された画像に含まれる建造物の座標の色と宅地利用動向調査の凡例の色の比較を行う (Fig.1)。その建造物の色がどの凡例の色に近いかに判定し建造物の判別を行う。タイル画像の建造物用途の色は建造物ごとの重心の色を取得した。

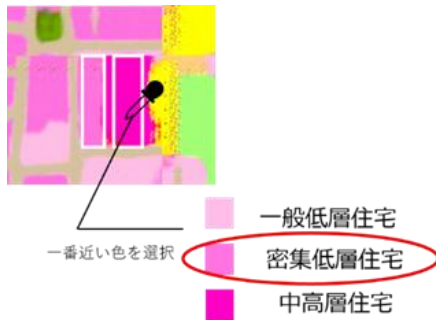


Fig. 1 : 色の判定方法

変換した地図タイル画像の建造物の色は RGB 形式で取得し、凡例の色をそれぞれ色との距離を式 (1) で計算し、値が一番小さい凡例をその建造物の分類結果とする。

$$d_{AB} = \sqrt{(R_A - R_B)^2 + (G_A - G_B)^2 + (B_A - B_B)^2} \quad (1)$$

色の判定を行い、モデルごとに混同行列を作成しモデルの評価を行う。結果の評価尺度には正解率と適合率と再現率の調和平均である f 値を使用する。正解率とは、すべてのクラスに対する予測の正解率である。また f 値を求めるために必要な適合率とは、予測されたもののうち本当に正解であった割合を示す。また、再現率とは本当に正解であるものに対してどれくらい正解と予測されたのかを表す指標である。そして、適合率と再現率はどちらも重要なためその調和平均をとって表したのが f 値といわれる。

### 3 実験

本研究では pix2pix のモデルを用いて航空写真から地図タイルへの変換を試み、建造物の用途の判別を行う。入力に必要な対のデータは国土地理院の空中写真<sup>8)</sup>と宅地利用動向調査<sup>9-11)</sup>のタイル画像を用いる。空中写真はズームレベル 18 で 2007 年以降の画像を用いる。宅地利用動向調査は三大都市圏(東京都を除く首都圏[2005]<sup>9)</sup>, 近畿圏[2008]<sup>10)</sup>, 中部[2003]<sup>11)</sup>)を調査対象としている。宅地利用動向調査の土地利用データは陸域 15 種類の項目と海, 対象地域外の 17 種類に分類されている。住宅は一般低層住宅, 密集低層住宅と中高層住宅の 3 つに分類される。ズームレベルはすべて 18 とする。初めに Table1 に示すように 2 種類の地域でバッチサイズとエポック数を変化させて実験をおこなった。

Table 1: 実験の種類

実験地域	バッチサイズ	エポック数
首都圏 (一部)	1	150
首都圏 (一部)	15	150
首都圏 (一部)	30	300
首都圏 (一部)	15	150
首都圏 (一部)	30	300
中部, 近畿	48	150
中部, 近畿	96	150
中部, 近畿	96	300

バッチサイズとは学習用データセットをいくつかのサブセットに分割し、このいくつかに分けたサブセットの中のデータの数をバッチサイズと呼び、この学習方法をミニバッチ学習と言う。サブセットに分けて学習する理由は、ミニバッチ確率的勾配降下法の要素が入るので学習効果が高まる。ミニバッチ確率的勾配降下法とは、学習データの中からランダムにバッチサイズの数だけデータを取り出し、そのデータを学習して学習データに対する誤差を計算し、パラメーターの更新をしていく仕組みのことである。またエポック数とは一つの訓練データをエポック数の数だけ繰り返して学習させることである。深層学習のようにパラメーターの数が多いと訓練データも繰り返し学習させないとパラメーターがうまく学習できないため適切な数だけ繰り返す必要がある。実験の結果、首都圏 (一部) では 30 バッチサイズ, 150 エポックの時に最も良い結果が得られた。また、中部, 近畿地方では 96 バッチサイズ, 150 エポックが最も良い結果が得られた。これより、バッチサイズは大きい方がよい結果だが、エポック数は 150 エポックがよい結果となることが分かった。この 2 地域はともに画像数が少ない実験だったが、次は画像数の多い宅地利用動向調査の対象地域すべてのタイル画像を用いてバッチサイズは 96, エポック数 150 で実験を行った。

実験結果は Fig.2 に示す。Fig.2 の (a) や (b) より、色が似ていると変換が困難であるとわかる。また共同住宅や商業施設など、航空写真から判別が比較的容易な分類は良い結果となった。正解率は評価の時に 0.7243 で f 値はすべてのクラスを平均して 0.3055 であった。Fig.3 はクラス別の混同行列である。混同行列とは分類結果をまとめた表のことで機械学習モデルの性能を測る指標として用いられる。縦軸は建造物のクラスごとの正解を表し横軸はクラスごとの予測を表し、表内の値は本当に正解であったもののうち、予測で正解とあてられたものの割合を正規化して示している。

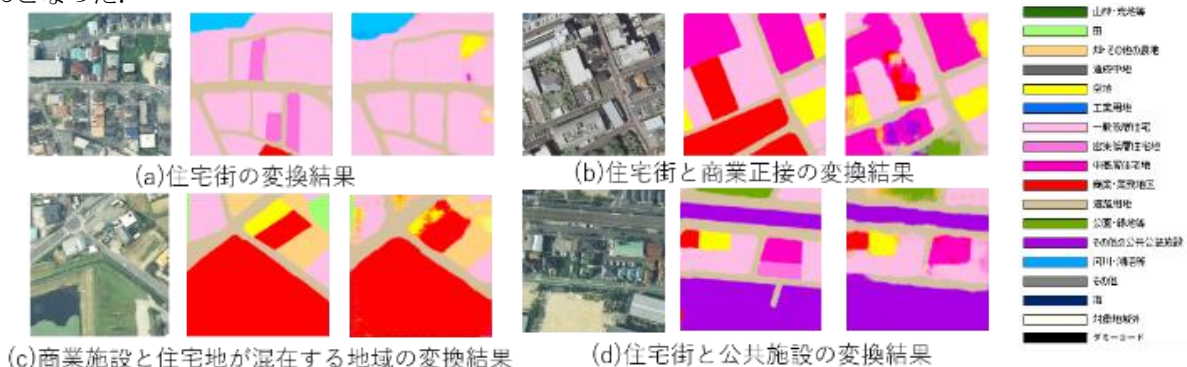


Fig.2 出力画像の一例

Fig.3 からわかるようにクラス別の f 値は一般低層住宅が約 9 割と最も高い値であった。ひとつの画像の中に複数の建造物が存在するため、Table 1 にはクラスごとの建造物数を示す。Table 1 よりクラスごとの建造物で一般低層住宅が一番多いため一般低層住宅に予測されると考えられる。また、Fig.3 よりほかのクラスも一般低層住宅に割り当てられやすいということが分かった。

本研究の目的は世帯数の調整を行うために建造物の用途判別によって宅地利用動向調査の分類の中の住宅に分類された建造物、一般低層住宅、密集低層住宅、中高層住宅の判別精度が一番重要である。一般低層住宅と密集低層住宅の区別は難しく学習の時から密集低層住宅は一般低層住宅に分類されやすい。しかしどちらも居住地区で低層住宅のため世帯数に大きな違いがないと考えられる。中高層住宅は半数以上を判別可能で中高層住宅と判別できなかった残りは主に一般低層住宅と商業施設に分類される結果となった。

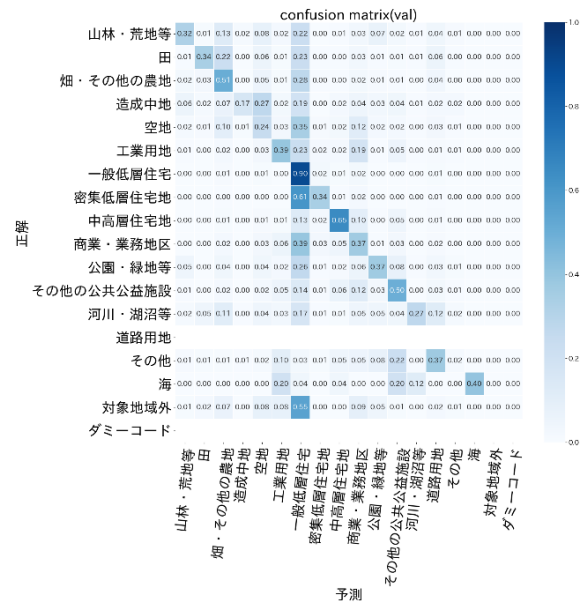


Fig. 3: 混同行列

Table 2: クラスごとの建造物数

	実験 id 9	
	train	validation
山林・荒地等	65,756	28,130
田	56,439	24,194
畑・その他の農地	229,176	97,552
造成中地	9,657	4,324
空地	251,052	106,120
工業用地	508,745	217,657
一般低層住宅	7,282,882	3,113,302
密集低層住宅地	776,275	336,110
中高層住宅地	247,391	106,205
商業・業務地区	1,212,158	521,235
道路用地	0	0
公園・緑地等	145,144	63,259
その他の公共公益施設	337,227	145,867
河川・湖沼等	6,719	2,965
その他	9,782	4,374
海	76	25
対象地域外	187	87
ダミーコード	0	0
合計	11,138,666	4,771,406

したがって航空写真から高さを読み取るのは困難であり、中高層住宅は多数の市民が居住するので大きな作りの建造物も多いためにこのような結果になったのだと考えられる。また、公園や公共施設は国が提供する統計データが存在するためにそのデータを考慮することによってモデル全体の f 値が上がるだろうと考えられる。航空写真と宅地利用動向調査の調査年が異なるために二つのタイル画像すべてが同じとは限らない。航空写真では工事中の建造物で造成中地なのに対して、宅地利

動向調査ではすでに建築が終わり一般低層住宅と分類されていたり、航空写真では商業施設が存在するのに対して、宅地利用動向調査では取り壊されて空地に分類されていたりする可能性があると考えられる。

また生成される画像には Fig.4 のようにノイズが発生している場合がある。そのためにノイズの部分に建物の重心があった際に、色の取得時に誤った色を取りかねない。このようなことがあるためにも f 値が上がらないと考えられる。



Fig. 4: 生成画像にみられるノイズ

本研究で得られた結果を用いて神奈川県相模原市中央区淵野辺五丁目で建造物に世帯数の調整を行った。本研究の結果を用いずに世帯数を割り当てようとすると Fig.5 のようになる。この図は国土地理院が提供する空中写真の上に淵野辺五丁目のみの基盤地図情報<sup>1,2)</sup>を重ねあわせ、その上に世帯数を割り当てている。こ



これは建物のあるところすべてに世帯が均等に分類されてしまうことがわかる。



Fig. 5: 研究結果を用いない淵野辺五丁目の世帯数の割り当て

次に本研究の結果を用いて世帯数を割り当てると Fig.6 のようになる。これは一般低層住宅と密集低層住宅に分類されたものに対しては世帯数 1 を割り当て、中高層住宅には残りの世帯数を均等に割り当てた。Fig.7 は建造物の割り当ての統計情報から世帯数を一般低層住宅と密集低層住宅地と中高層にわけたものである。この統計データは、宅地利用動向調査と基盤地図情報を使用して作成したデータである。Fig.6, Fig.7 の黄色の小さな丸は一般低層住宅と密集低層住宅で赤色の大きな丸は中高層住宅を示す。どちらも比較的左上に住宅地が集まっていて、大学のところには居住地区がないことをしめす。



Fig. 6: 研究結果を用いた淵野辺五丁目の世帯数の割り当て



Fig. 7: 公開されている建造物の用途を用いた世帯数の割り当て

Fig.8 は Fig.6 と Fig.7 の左上の住宅地が密集しているところを拡大した図で左が Fig.6 の拡大図で右が Fig.7 の拡大図である。赤色の中高層住宅地は数は同じだが場所が異なるため、誤って判断されたことがわかる、



(a) Fig.6 の拡大 (b) Fig.7 の拡大

Fig. 8: 居住地区の拡大

また、Fig.9 は Fig.6 と Fig.7 の左下、大学の敷地内の拡大で左が本研究結果を用いた図で右が宅地利用動向調査のデータからの割り当てである。大学の敷地内に黄色の小さな丸がいくつか存在している。正解のほうにはなく、大学の敷地内のため居住している人はいないと考えられる。そのため誤った分類だったことが示唆される。しかし、ほかの建物には割り当てられていないために比較的高い精度で分類ができていると考えられる。建造物の分類がされていない地域への応用も期待される。



(a) Fig.6 の拡大 (b) Fig.7 の拡大

Fig. 9: 大学内の拡大

#### 4 結論と今後の課題

本研究では、現実に近い社会シミュレーション実施のために必要な市民のデータを、個人の特性を避け実用させるために航空写真から建造物用途の判別を行うことにより世帯数の調整を試みた。深層学習のモデルの一つである pix2pix を用いることで航空写真から宅地利用動向調査の用途画像へのスタイル変換を行い、画像に含まれる建造物の色を凡例の色と比較させることで生成された画像の精度を明らかにした。実験に使用された画像のクラスごとの割合に偏りがあるために正解率と、適合率と再現率の調和平均である f 値の二つを用いて評価を行った。結果として正解率は高くなったものの f 値は 3 割程度であった。これは入力に使用した航空写真と宅地利用動向調査の用途画像の調査年が異なることによって生成された画像と正解画像との間で異なるものと判断されてしまう可能性があるとして示唆される。

今後の課題として実験に使用した航空写真と宅地利用動向調査の調査年が異なることによってモデルの精度にも大きく影響を及ぼしたため入力に使用する画像の調査年をそろえることができれば精度が上がるだろう。

また、ノイズの発生により重心の点の色が誤って取得される可能性が考えられる。そのため本研究で提案した評価手法のような建造物の一点を取得するのではなく、その建造物の複数の場所から色を取得することによって精度の向上が見込める。

#### 謝辞

本研究の一部は JSPS 科研費 19K23229 と AOYAMA VISION「AI 研究拠点形成プロジェクト」の助成を受けたものです。

#### 参考文献

- 1) R. Axelrod : *Advancing the Art of Simulation in the Social Sciences*, Springer, Berlin, Heidelberg (1997)
- 2) 寺野 隆雄 : エージェントベースモデリング : KISS 原理を超えて, 人工知能学会誌, **18-6**, 710/715 (2003)
- 3) 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)
- 4) T. Harada , T. Murata : Projecting Households of Synthetic Population on Buildings, *SICE Journal of Control, Measurement, and System Integration*, **10-6**, 505/512 (2017)
- 5) J. Kang, M. Körner, Y. Wang, H. Taubenböck , X. X. Zhu : Building instance classification using street view images, *ISPRS Journal of Photogrammetry and Remote Sensing*, **145**, Part A, 44/59 (2018)

- 6) P. Isola, J. -Y. Zhu, T. Zhou , A. A. Efros : Image-to-Image Translation with Conditional Adversarial Nets, *Conference on Computer Vision and Pattern Recognition 2017*, 1125/1134 (2017)
- 7) Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli : Image quality assessment: From error visibility to structural similarity, *IEEE Transactions on Image Processing*, **13-4**, 600/612 (2004)
- 8) 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)
- 9) 国土地理院 : 宅地利用動向調査 首都圏 2005, [https://cyberjapandata.gsi.go.jp/xyz/lum4bl\\_capital2005/{z}/{x}/{y}.png](https://cyberjapandata.gsi.go.jp/xyz/lum4bl_capital2005/{z}/{x}/{y}.png) (2005)
- 10) 国土地理院 : 宅地利用動向調査 近畿圏 2008, [https://cyberjapandata.gsi.go.jp/xyz/lum4bl\\_kinki2008/{z}/{x}/{y}.png](https://cyberjapandata.gsi.go.jp/xyz/lum4bl_kinki2008/{z}/{x}/{y}.png) (2008)
- 11) 国土地理院 : 宅地利用動向調査 中部圏 2003, [https://cyberjapandata.gsi.go.jp/xyz/lum4bl\\_chubu2003/{z}/{x}/{y}.png](https://cyberjapandata.gsi.go.jp/xyz/lum4bl_chubu2003/{z}/{x}/{y}.png) (2003)
- 12) 国土地理院 : 基盤地図情報, <https://fgd.gsi.go.jp/download/terms.html> (2020 年 1 月閲覧)