

Pix2Pix を用いた建物の用途判別における土地利用ラベルの簡素化 ー合成人口の改良に向けて

○福嶋竜希 山崎幹矢 原田拓弥 大内紀知（青山学院大学）

To Simplify Geographical Information Labels for Building Usage Classification via Pix2Pix Toward Improving Synthetic Population

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

概要ー 合成人口の住居への割り当てを精緻化するため、先行研究では航空写真と宅地利用動向調査を用い、建物の用途をクラスとして判別するモデルが提案されている。しかし、そこで用いられたクラスラベルは複雑である。本研究では、建築物面積の占有率や人間の住環境という観点から、このクラスラベルを単純化することで、一部のクラスにおいて、判別精度の改善が見られた。

キーワード: Pix2Pix, 社会シミュレーション, 航空写真, 合成人口

1 研究背景

社会科学のような、人間社会を対象とする問題では、社会そのものや構成員に悪影響がある実験は許されない¹⁾。そこで、コンピューター上に仮想的な社会（人工社会）を再現し、その中でシミュレーションを行う社会シミュレーションが注目を集めている。社会シミュレーションには、例えば世帯の構成や位置情報といったレベルまで、可能な限り現実を即した人工社会が求められる場合がある。Kelso et al.²⁾ は、新型インフルエンザの流行に対し、ワクチンや治療薬などの医療対策が講じることができない場合に、Social Distancing の有効性をシミュレーションにより評価している。具体的には、オーストラリアのアルバニーを対象とし、その国勢調査結果の人口や世帯、職業の分布に合うような人口を仮想的に作成した上で、学校閉鎖・通勤の規制・外出自粛・罹患者の隔離の 4 つの施策の効果を罹患者率や 1 日の最大有症者数などで評価している。

このように、人工社会の再現のために人間のデータが必要となる場合があるが、それらは個人情報のため、入手及び利用が困難である。原田ら³⁾ は人口統計を基に、年齢や職業、世帯構成の分布などの統計的性質が反映された仮想的な人口データ（合成人口）を日本全体に渡って合成した。合成人口を用いた社会シミュレーションに関する研究は広く行われている。例えば江尻ら⁴⁾ は、大阪府阿倍野区の合成人口を用い、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、それぞれの施策の効果をシミュレーションにより評価している。

しかし、原田ら³⁾ は合成人口の世帯に対して位置情報、つまり地図上の住居を与える際、対象地域の建物をランダムに選択して割り当てている。このため、人が住むことができないほどに小さい建物に居住する世帯や大規模な建物にごく少数の世帯しか居住していない場合がある。加えて、学校や病院のような住居では

ない建物に居住者がいるという問題もある。そこで、Harada et al.⁵⁾ は建物の面積を考慮することでこれらの問題を緩和した。しかし、この手法は建物面積に関するパラメーターを調整する必要があり、その結果は対象地域でのみ有効である。他の地域で同手法を用いる場合は、このパラメーターを調整する必要がある。また、高層マンションのように、面積は小さいが多数の世帯が居住する建物は考慮できない。

建物の大まかな階数や用途が判明していれば、それに基づいて世帯の割り当てを行うことで、合成人口の位置情報をより精緻なものにできる。田口ら⁶⁾ は、Pix2Pix⁷⁾ という深層学習のモデルを用いて、航空写真から地理情報の画像を生成し、そこから建物の用途を抽出する手法を提案している。同研究では、地理情報として、国土地理院が提供する宅地利用動向調査を用いており、これは土地の利用形態を提供するものである。この研究に使われている土地利用のラベルには建物をほとんど含まないラベルが多く存在している。この意味で、タスクが複雑化していると考えられ、これを単純化することができればより精度の高い結果が期待できる。

本研究では、土地に占める建築物の面積の割合や、人間の居住地として適しているかを勘案して、過不足のない土地利用ラベルを選定する。そして、それを用いて、先行研究と同じく Pix2Pix を訓練し、判別精度の向上を試みる。

2 先行研究

田口ら⁶⁾ は国土地理院が提供する空中写真⁸⁾ と数値地図 5000（土地利用）⁹⁾ を用いて、Pix2Pix⁷⁾ を訓練することで、航空写真から地理情報の画像を生成し、そこから建物の用途を抽出する手法を提案している。空中写真は、日本のほぼ全土を網羅している。また、数値地図 5000 は首都圏・近畿圏・中部圏それぞれの宅地利用動向調査の結果がまとめられた地理情報であり、一般低層住宅や田、工業用地など 18 の土地利用のラ

ベルをもつ。空中写真とは対照的に、Fig. 1 に示すように、対象地域は日本全土に比べると、限定的である。田口らは、同調査を利用する際、Fig. 2 に示す18の色を土地利用ラベルとした上で、画像に変換している。加えて、対象地域外を表す画素数が、画像全体の画素数の内 90 %以上を占める画像は除外している。

Pix2Pix は GAN¹⁰⁾ (Generative Adversarial Network) の一種であり、一般的な GAN で用いられる学習フレームワークと、多くの機械学習の学習フレームワークである尤度最大化を組み合わせて学習されるモデルである。

GAN と呼ばれるモデルには、Generator と Discriminator という 2 つの機構がある。入力は Generator が受け取り、その時点でのパラメーターに基づいて、出力を計算する。Discriminator は、教師ラベルと Generator の出力を入力にとり、それらが Generator が生成した出力であるかどうかを判断する。つまり、Discriminator にとっては、Generator が生成した出力であるかどうか教師ラベルになり、この判別精度が高くなるように、そのパラメーターを更新する。また、Generator は Discriminator が生成されたものと判断できなくするように、そのパラメーターを更新する。全体として、Generator と Discriminator それぞれが、競合する形で学習が進行する。最終的に、精度が高い Discriminator を以てしても生成されたものと判別できないような出力を与える Generator が得られると、精度の高い出力が期待できる。なお、Discriminator は学習時のみに使用され、予測タスクを行うときは、Generator の出力のみが興味の対象となる。

Pix2Pix は、画像とそれに対応する画像が全体に対する入力である。本研究では、航空写真とその土地利用データの画像がこれに当たる。以下に、Generator と Discriminator それぞれの目的関数を示す。

Discriminator:

$$\max_{W_D} \frac{1}{|I|} \sum_{(x,t) \in I} \ln[D(x,t;W_D)(1-D(x,y;W_D))] \quad (1)$$

Generator:

$$\min_{W_G} \frac{1}{|I|} \sum_{(x,t) \in I} [\ln(1-D(x,y;W_D)) + \lambda \|y-t\|_1] \quad (2)$$

where $y = G(x;W_G)$

W_D, W_G はそれぞれ Discriminator, Generator のパラメーターをまとめたものを表し、これらが学習の過程で更新されていく。I は学習用画像の集合であり、 x が入力画像、 t が教師ラベルとなる画像である。G は Generator を表す関数で、入力画像に対して生成画像を出力する。一方で、D は Discriminator を表す関数であり、入力画像 x と別の画像 y' の組に対して、 y' が x に対応する教師画像である信念の度合いを返す関数である。従って、 $D(x,t;W_D)$ は、 t が x の教師画

像である信念の度合いを表すので、Discriminator はこれを最大化しようとする。 $\|\cdot\|_1$ は L1 norm を意味し、 $\|y-t\|_1$ が小さいほど、L1 距離の意味で生成された画像が教師画像に近いということである。通常の GAN の誤差関数には、この項がない。最後に、 λ はこの項の影響の度合いを調節するハイパーパラメーターである。

LSGAN¹¹⁾ は、誤差関数に二乗和誤差関数を使用して学習する GAN である。これにより、勾配消失などの問題を緩和することが経験的に知られている。Pix2Pix の目的関数にこれを用いると、以下のようになる。

Discriminator:

$$\min_{W_D} \frac{1}{|I|} \sum_{(x,t) \in I} [\|D(x,t;W_D) - 1\|_2^2 + \|D(x,y;W_D) - 0\|_2^2] \quad (3)$$

Generator:

$$\min_{W_G} \frac{1}{|I|} \sum_{(x,t) \in I} [\|D(x,y;W_D) - 1\|_2^2 + \lambda \|y-t\|_1] \quad (4)$$

where $y = G(x;W_G)$

田口らは、式(3), (4) を目的関数として用いている。そして、空中写真を入力として受け取った Pix2Pix は、各画素に対して、土地利用ラベルの色を基にして推定した画素値を出力する。この結果、1枚の空中写真に対して、Fig. 3 に示すような画像形式の出力が得られる。しかし、Pix2Pix が出力する画像には、建物の位置情報がないため、これだけでは建物の用途を判別することはできない。そこで、同研究では、実際の建物の位置情報と Generator の出力を照合して、その精度を評価する。具体的には、建物の重心の緯度経度を取り、そこに対応する Generator の出力の画素値を取得する。ここで、Generator の出力する画素値が、各クラスに応じて設定した RGB 値と等しいとは限らないので、Euclid 距離の意味で最も近い色のクラスを判別結果としている。つまり、Generator の出力から取得した建物の重心座標の RGB 値を $\mathbf{p}$ としたとき、以下によって、判別結果を決定している。

$$\underset{c}{\operatorname{argmin}} \|\mathbf{p} - \mathbf{c}\|_2 \quad (5)$$

$\mathbf{c}$: 各ラベルの RGB 値



Fig. 1: 宅地利用動向調査の対象地域（赤）



Fig. 3: 航空写真（左）と田口らのモデルにより生成された土地利用の画像（右）

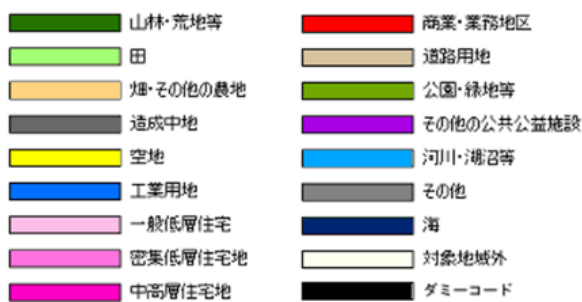


Fig. 2: 田口らが用いた教師ラベルの一覧

3 提案手法

合成人口の世帯への住居割り当てを考えるならば、建物の用途だけが判別できれば十分であり、田口らが用いた教師ラベルは、この意味で、モデルに必要以上に複雑なタスクを強いていると考えられる。Table 1 に宅地利用動向調査の各土地利用の面積と、その内、建築物が占める面積を示した。また、建築物が占める面積を土地利用の面積で除することで、建築物の占有率を算出した。結果、対象地域外、畑・その他の農地、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、海の 11 の土地利用分類は、建築物の占有率が 10% にも満たないことがわかる。従って、本研究ではこれらを、その他という単一の分類にまとめる。世帯を住居に割り当てるといことは、対象地域が市街地や住宅街など、人間の居住環境として適していることが前提となる。都市計画制度によれば、住環境の観点から、住居専用地域あるいは住居地域に工場を建てることはできない。従っ

Table 1: 宅地利用動向調査の土地利用と建築物の面積

土地利用名	建築物の面積 (m ²)	宅地利用動向調査の面積 (m ²)	建築物の占有率
一般低層住宅地	810,076,746.6	1,967,276,384.2	0.41178
商業・業務用地	282,991,295.8	698,299,519.7	0.40526
工業用地	204,549,354.3	530,104,696.1	0.38587
対象地域外	175,632,186.9	2,851,030,749.3	0.06160
その他の公共公益施設用地	124,582,458.8	670,663,231.9	0.18576
中高層住宅地	76,090,492.4	223,360,492.8	0.34066
密集低層住宅地	56,924,803.6	96,429,770.0	0.59032
畑・その他の農地	38,260,177.0	1,300,785,090.3	0.02941
空地	31,511,714.7	613,311,524.5	0.05138
公園・緑地等	22,071,529.6	425,228,418.5	0.05191
田	6,236,420.3	1,583,217,750.5	0.00394
山林・荒地等	6,123,815.9	3,307,499,892.9	0.00185
その他	3,543,039.4	52,548,176.7	0.06742
造成中地	2,739,406.5	66,044,707.8	0.04148
道路用地	2,588,050.8	1,444,170,586.9	0.00179
河川・湖沼等	1,547,733.3	656,614,077.0	0.00236
海	16,626.5	1,468,398,134.0	0.00001

て、宅地利用動向調査の土地利用のひとつである工業用地もその他に含めることとする。また、宅地利用動向調査では、一般低層住宅と密集低層住宅の違いは敷地面積のみであるため、これらはひとつにまとめて、低層住宅とする。

以上より、本研究では、低層住宅、中高層住宅、公共・公益施設、商業・業務地区、その他の 5 つの土地利用分類を設定し、それぞれに Table 2 の色を RGB 値で与える。結果として、Fig. 4 にあるように航空写真の例とそれに対応する座標の数値地図 5000 のデータを画像にしたものが得られる。

式 (4) にあるように、Generator の目的関数には、L1 Loss の項が含まれるため、Euclid 距離で判別を行うよりも、Manhattan 距離を用いた方が良いということも考えられるため、両方の距離を用いて判別を行い、結果を比較する。

Table 2: 本研究で用いるラベルとその RGB 値

色	R	G	B	ラベル
	227	26	28	低層住宅
	51	160	44	中高層住宅
	31	120	180	公共・公益施設
	255	255	255	商業・業務地区
	0	0	0	その他 (道路や山林など)



Fig. 4: 航空写真²⁾ (左) と対応するラベル変更後の数値地図 5000³⁾ (右) の例

4 実験結果

数値地図 5000 の用いる画像数を Table 3 に示す。ただし、その他というラベルが、画素数の 90% 以上を占める画像は除いてある。全体の内、およそ 7 割程度の 322,880 枚を訓練用画像、残り 138,420 枚の内、50,000 枚を評価用画像、その残りの 88,420 枚をテスト用画像として用いる。Table 4 に訓練用・評価用・テスト用画像それぞれに含まれる土地利用クラス別建物数を示す。

本研究では、Table 5 のハイパーパラメーターの候補を設定し、それぞれの組み合わせで学習及び評価用画像による精度評価を行い、ハイパーパラメーターの最良の組み合わせを採用する。なお、パラメーターの更新則には、Adam を使用した。Adam は勾配降下法の一つである。それゆえ、各反復で全ての学習データに対する誤差を計算するには莫大な時間を要する。加えて、過剰適合の恐れもある。そこで、バッチサイズという値を固定し、学習データをバッチサイズと同数の画像の部分集合 (以下、ミニバッチ) に分割し、1 反

復毎に 1 ミニバッチに対する誤差を計算してパラメーターを更新する、確率的勾配降下法によって学習を行う。また、Pix2Pix のパラメーター数は莫大であり、大量の学習データが必要となる。従って、次の 2 つの対策を組み合わせる。1 つ目は、学習用データ全てを用いて 1 度の学習が終わったら、同じ学習用データを用いてもう一度学習することを繰り返すことである。ここで、学習用データを使う回数をエポック数といい、各エポックにおいてミニバッチが新たに作り直される。2 つ目は、ミニバッチ内の画像を 0.5 の確率で水平方向に反転させることである。これにより、異なった画像が作られるため、過剰適合を抑えることができる¹²⁾。

Table 6 に示す計算機環境を用いて、Table 5 に示すハイパーパラメーター全ての組み合わせ (100 通り) で訓練データを用いて学習を行い、評価データで、それぞれの性能を算出した結果のうち上位 10 を Table 7 と Table 8 に示す。前者は、Manhattan 距離で判定を行なった場合、後者は Euclid 距離で判定を行なった場合の結果である。ここでは、ラベルが複数あることを考慮して、まずクラス毎の正解率 (全建物数に対し、予測結果と実際の建物用途が合致したものの割合) を算出し、その全クラスの結果を平均したものを用いる。その結果が正解率平均である。列名であるハイパーパラメーターの組みは、(Discriminator Dropout Rate, Generator Dropout Rate, Lambda) の順でハイパーパラメーターを設定したことを意味する。また、GAN の訓練では Mode Collapse という、入力を変えてもモデルが同じ出力ばかりをする問題があり、本研究でもそれが確認されたハイパーパラメーターの組みが存在した。これらは無視する。よって、ハイパーパラメーターの組み合わせは Euclid 距離、Manhattan 距離いずれを用いた場合でも、(Discriminator Dropout Rate, Generator Dropout Rate, Lambda) = (0.3, 0.5, 100.0) が最良の結果となったので、これを採用する。

次に、訓練データに評価データを加え、計 372,880 枚の画像を用い、上記のハイパーパラメーターを設定し、Pix2Pix を改めて学習させる。その結果の単純な正解率と、正解率平均を Table 9 に示す。また、この学習データに対する結果を Table 10 に示す。田口らは 0.7234 という正解率を記録している。一方で、Table 9 より、本研究では、0.8126 を記録している。しかし、前述の通り、用いているラベルが異なるため、単純に比較をすることはできない。以降は、クラス毎の精度に着目する。Table 11, Table 12 はテスト画像に対し、Manhattan 距離、Euclid 距離をそれぞれ用いた場合の混同行列である。また、Table 13, Table 14 は訓練画像に対し、Manhattan 距離、Euclid 距離をそれぞれ用いた場合の混同行列である。そして、比較のため、田口らの混同行列を Table 15 に示す。これらの混同行列は何れも、縦方向が実際のクラス、横方向がモデルの予測、そして各行は対応するクラスの建物数で除して正規化してある。例えば、Table 11 の (1, 2) 成分は 0.0061 であるが、これは実際に低層住宅であるもののうち、中高層住宅と判別されたものの割合が 0.0061 であることを意味する。加えて、田口らの混同行列の作成にあたり、一般低層住宅は 3,113,302 軒、密集低層住宅は 336,110 軒が用いられており、一般低層住宅の方が圧倒的に多い。このため、以降の議論は一般低層住宅に焦点を当てる。Table 16 に Manhattan 距離、

Euclid距離それぞれを用いた際の F1 score を示す。全体的に大きな差はないが、中高層住宅において、Euclid 距離の方が 0.005 程度大きい結果を示した。次に、本研究の訓練データの精度と、テストデータの精度に大きな差はないため、過剰適合の恐れは少ないと言える。次に、世帯の割り当てにとって、最も重要である低層住宅と中高層住宅に着目する。中高層住宅の精度は、田口らと同程度であったが、低層住宅の精度については、3% 程度改善された。ただし、ここでは Table 15 で実際に一般低層住宅であるもののうち、一般低層住宅に分類されたもの、密集低層住宅に分類されたものの割合を合計した値と比較している。一方で、商業・業務地区は、低層住宅への誤分類が大きく増え、精度は 6% 程度下がった。最後に、公共・公益施設は 1.5% 程、精度が改善された。

Table 1 で示した通り、低層住宅地の面積は、本研究で用いた他の分類よりも遥かに大きい。このため、モデルが低層住宅地を積極的に出力することを助長してしまう。結果として、実際に低層住宅である建物のうち、低層住宅であると判別される建物の割合がより大きくなるが、それに伴い、他の用途を低層住宅と誤判別する割合も大きくなると考えられる。実際、Table 15 の田口らの混同行列で一般低層住宅と推定されたもののうち、実際は中高層住宅、公共・公益施設、商業・業務地区であったものの割合はそれぞれ、0.13, 0.14, 0.39 であるが、本研究では Table 12 より、0.16, 0.17, 0.52 程度であり、増加している。田口らは、Fig 2 に示すすべてのラベルを用いたため、本研究の方が土地利用ラベルの偏りが大きかったと考えられる。これにより、その結果として住宅街にも多く存在する商業・業務地区を誤判別するケースが多くなってしまったと考えられる。この偏りへの対処として、少ないクラスの画像を水増しして、訓練画像を増やすことが挙げられる。

Table 3: 本研究で用いる数値地図 5000 の画像枚数（地域別）

地域（調査年）	画像枚数（枚）
近畿圏（2008年）	122,887
首都圏（2005年）	264,713
中部圏（2003年）	73,700
合計	461,300

Table 4: 訓練・評価・テストデータに含まれる建物数の集計

土地利用	訓練データ	評価データ	テストデータ
低層住宅	7,494,800	1,157,009	2,037,702
中高層住宅	235,476	36,742	64,744
公共公益施設	314,203	49,046	85,994
商業業務地区	1,130,522	173,988	312,876
その他	876,235	135,072	237,883
合計	10,051,236	1,551,857	2,739,199

Table 5: ハイパーパラメーターとその候補

ハイパーパラメーター	値
バッチサイズ	32
エポック数	59
学習係数	0.0002
lambda	0, 1, 10, 100
Discriminator dropout rate	0.1, 0.2, 0.3, 0.4, 0.5
Generator dropout rate	0.1, 0.2, 0.3, 0.4, 0.5

Table 6: 使用した計算機のスペック

CPU	Intel Xeon Gold 6126 (Skylake / 2.6 GHz 12コア)
GPU	NVIDIA Tesla P100 (NV-Link)
メモリ	192 GB

Table 7: Manhattan 距離で判定した際の正解率平均上位 10

ハイパーパラメーターの組み	正解率平均
(0.3, 0.5, 100.0)	0.5590
(0.5, 0.5, 100.0)	0.5572
(0.4, 0.5, 100.0)	0.5526
(0.1, 0.4, 100.0)	0.5518
(0.2, 0.4, 100.0)	0.5512
(0.2, 0.5, 100.0)	0.5510
(0.1, 0.5, 100.0)	0.5456
(0.3, 0.3, 100.0)	0.5446
(0.2, 0.3, 100.0)	0.5429
(0.4, 0.4, 100.0)	0.5428
(0.3, 0.5, 100.0)	0.5590

Table 8: Euclid 距離で判定した際の正解率平均上位 10

ハイパーパラメーターの組み	正解率平均
(0.3, 0.5, 100.0)	0.5592
(0.5, 0.5, 100.0)	0.5574
(0.4, 0.5, 100.0)	0.5527
(0.1, 0.4, 100.0)	0.5520
(0.2, 0.4, 100.0)	0.5511
(0.2, 0.5, 100.0)	0.5511
(0.1, 0.5, 100.0)	0.5450
(0.3, 0.3, 100.0)	0.5447
(0.2, 0.3, 100.0)	0.5427
(0.4, 0.4, 100.0)	0.5423
(0.3, 0.5, 100.0)	0.5592

Table 9: テストデータの正解率と正解率平均

	Manhattan	Euclid
正解率	0.8127	0.8126
正解率平均	0.5689	0.5691

Table 10: 訓練データの正解率と正解率平均

	Manhattan	Euclid
正解率	0.8210	0.8209
正解率平均	0.5884	0.5887

Table 11: テストデータにおける正規化された混同行列 (Manhattan 距離)

実際	予測				
	低層住宅	中高層住宅	公共・公益施設	商業・業務地区	その他
低層住宅	0.9546	0.0061	0.0020	0.0112	0.0261
中高層住宅	0.1600	0.6553	0.0446	0.0763	0.0638
公共・公益施設	0.1756	0.0799	0.5157	0.0969	0.1319
商業・業務地区	0.5198	0.0354	0.0432	0.3096	0.0920
その他	0.4283	0.0250	0.0506	0.0869	0.4092

Table 12: テストデータにおける正規化された混同行列 (Euclid 距離)

実際	予測				
	低層住宅	中高層住宅	公共・公益施設	商業・業務地区	その他
低層住宅	0.9542	0.0059	0.0020	0.0111	0.0268
中高層住宅	0.1591	0.6537	0.0448	0.0761	0.0663
公共・公益施設	0.1746	0.0777	0.5157	0.0965	0.1355
商業・業務地区	0.5201	0.0341	0.0427	0.3088	0.0944
その他	0.4266	0.0233	0.0504	0.0869	0.4128

Table 13: 訓練データにおける正規化された混同行列 (Manhattan 距離)

実際	予測				
	低層住宅	中高層住宅	公共・公益施設	商業・業務地区	その他
低層住宅	0.9569	0.0058	0.0018	0.0103	0.0253
中高層住宅	0.1559	0.6765	0.0388	0.0685	0.0603
公共・公益施設	0.1756	0.0756	0.5491	0.0792	0.1205
商業・業務地区	0.5187	0.0329	0.0386	0.3275	0.0822
その他	0.4274	0.0253	0.0459	0.0692	0.4322

Table 14: 訓練データにおける正規化された混同行列 (Euclid 距離)

実際	予測				
	低層住宅	中高層住宅	公共・公益施設	商業・業務地区	その他
低層住宅	0.9565	0.0056	0.0017	0.0102	0.0261
中高層住宅	0.1549	0.6752	0.0390	0.0683	0.0626
公共・公益施設	0.1746	0.0735	0.5492	0.0788	0.1239
商業・業務地区	0.5190	0.0317	0.0380	0.3268	0.0844
その他	0.4257	0.0239	0.0455	0.0691	0.4357

Table 15: 田口ら⁶⁾の混同行列 (一部抜粋)

実際	予測				
	低層住宅		中高層住宅	公共・公益施設	商業・業務地区
	密集低層住宅	一般低層住宅			
密集低層住宅	0.34	0.61	0.01	0.00	0.02
一般低層住宅	0.02	0.90	0.01	0.00	0.02
中高層住宅	0.02	0.13	0.65	0.05	0.01
公共・公益施設	0.01	0.14	0.06	0.50	0.12
商業・業務地区	0.03	0.39	0.05	0.03	0.37

Table 16: 各用途の F1 score

	Manhattan	Euclid
低層住宅	0.9105	0.9104
中高層住宅	0.5917	0.5965
公共・公益施設	0.5444	0.5455
商業・業務地区	0.4153	0.4146
その他	0.4500	0.4502
平均	0.5824	0.5834

5 結論と今後の課題

本研究は、現実即した社会シミュレーションの実現に向けて、住居に対し、合成人口世帯の割り当ての精緻化を助けることを目的とした。同様の目的で、田口らは、航空写真と宅地利用動向調査の結果を深層学習モデルのひとつである Pix2Pix の学習データとし、航空写真に映る土地の利用形態を判別した上で、建物の用途を判別した。しかし、田口らが用いた土地利用ラベルは必要以上に複雑であった。そこで本研究は、土地利用別の建築物の占有率や、市街地・住宅地に焦点を当て、過不足のないラベルの選定を提案した。その上で、いくつかのハイパーパラメータの組み合わせの候補を用意し、Pix2Pix を訓練した。結果として、(Discriminator Dropout Rate, Generator Dropout Rate, Lambda) = (0.3, 0.5, 100.0) という組み合わせが経験的に最良のものとなった。これらを用いて訓練したモデルを評価したところ、一般低層住宅の正解率は、田口らの結果と比較して 5.5 % ほど改善された。また、テストデータに対する全体の正解率については、およそ 9 % 改善された。

今後の課題として、土地利用別建物数の大きな偏りの対処が挙げられる。これに対しては、比較的建物数が少ない土地利用が映る画像の水増しをすることが対策として考えられる。また本研究では、工業用地をその他に含めたが、都市計画制度によると、小規模の工場であれば、商業地域に建設することが可能である。このため、学習データに工業用地が多く含まれ、その他以外の建物に分類された恐れがある。工業用地に対しても、ラベルを付与して学習を行った結果と本研究の結果の比較が望まれる。

謝辞

本研究の一部は、JSPS 科研費 19K23229 の助成を受けたものです。また、本研究成果の一部は、大阪大学サイバーメディアセンターのOCTOPUS を利用して得られたものです。

参考文献

- 1) 塩沢：経済学にとっての人工市場，人工知能学会誌 **15**-6, 951/957 (2000)
- 2) J. Kelso, G. Milne and H. Kelly: Simulation suggests that rapid activation of social distancing can arrest epidemic development due to a novel strain of influenza, *BMC Public Health*, **9**, 117/126 (2009)
- 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) 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) T. Harada, T. Murata: Projecting Households of Synthetic Population on Buildings, *SICE Journal of Control, Measurement, and System Integration*, **10**-6, 505/512 (2017)
- 6) 田口，原田，大内：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，計測自動制御学会第 22 回社会システム部会研究会資料，124/128 (2020)
- 7) P. Isola, J. Zhu, T. Zhou and A. Efros: Image-To-Image Translation With Conditional Adversarial Networks, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1125/1134 (2017)
- 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：数値地図 5000 (土地利用)，<https://www.gsi.go.jp/kankyochiri/lum-5k.html>
- 10) I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets, *Neural Information Processing Systems*, 2672/2680 (2014)
- 11) X. Mao, Q. Li, B. Lau and S. Smolley: Least Squares Generative Adversarial Networks, *Proceedings of the IEEE International Conference on Computer Vision*, 2794/2802 (2017)
- 12) L. Taylor and G. Nitschke: Improving Deep Learning with Generic Data Augmentation, *IEEE Symposium Series on Computational Intelligence*, 1542/1547 (2018)