

公立はこだて未来大学 2022 年度 システム情報科学実習
グループ報告書

Future University Hakodate 2022 Systems Information Science Practice
Group Report

プロジェクト名

Let' s リモートセンシング

Project Name

Let' s Remote Sensing

グループ名

グループ B

Group Name

Group B

プロジェクト番号/Project No.

18-B

プロジェクトリーダー/Project Leader

野原広大 Kohdai Nohara

グループリーダー/Group Leader

渡部日向 Hinata Watanabe

グループメンバ/Group Member

池田幸星 Kousei Ikeda

成田龍斗 Ryuto Narita

野原広大 Kohdai Nohara

指導教員

和田雅昭 長崎健

Advisor

Masaaki Wada Takeshi Nagasaki

提出日

2022 年 1 月 19 日

Date of Submission

January 19, 2022

概要

私たちは、都市開発のために地方自治体に利用してもらうことや、個人が研究で用いることを目的として、入力された都市の衛星画像から将来の都市予測を画像として出力するモデルの製作を目標とした。従来のモデル作成と異なり、複雑な入力や現地調査が必要無いといった利点がある。前期では、2010年から2020年の10年間で人口変動が増加、減少、変化なしで分類された26都市の衛星データから人工物のみの面積を求め、人口との関係性を考察した。考察は相関関係の分析と単回帰分析にて行った。人口減少の相関は-0.783であり強い負の相関となった。これは、人口が減少すると人工物の面積が増える、ということの意味している。単回帰分析でも同様の結果を得られた。変化なしは相関が0.356であり弱い正の相関となった。しかし、単回帰分析では負の相関が得られ、矛盾が生じてしまった。データを増やすなど、さらなる調査が必要だと考える。人口増加について、算出された面積データが他都市の測定値と比較し大きく乖離しているため、データ不良であると判断し、求めることはできなかった。よって、人口増加のデータ不良や、変化なしの分析結果の矛盾より、データの正確性に問題があると考えられる。このことから、正確なデータの取得やその方法の模索が課題となった。

後期では、前期で求めた人口変動と都市面積の関係性をもとに、前期で定めた分類条件に基づいて改めて人口変動が増加、減少、変化なしに当てはまる都市をそれぞれ70都市、108都市、151都市選出し、2010年と2020年のデータを収集した。その後、それらデータを教師データに加工した。また、未来の都市予測を画像で生成するモデルを作成した。この画像生成モデルに作成した教師データを入力することで、10年後の都市を予測した結果を出力することができた。しかし、2010年のデータから2020年の都市を予測する出力結果は画像がぼやけていたり、建物と田畑の判別が困難であったりと、いくつかの問題が発生した。この原因として考えられることとして、2010年と2020年のデータでは画像の色調が異なることや、2010年と2020年のデータにおいてダウンロード時の解像度が異なることがあげられる。このことから、教師データの見直しやさらなる調整、画質向上などが今後の課題となった。

また、これまでのグループの活動を振り返って、スケジュール見積り甘さから思うような成果を出せなかった。スケジュール管理の徹底やグループ内での作業の調整など、プロジェクトを行う上で重要なことを学ぶ良い機会となった。

(※文責: 野原広大)

Abstract

Our goal was to create a model that can output future city predictions as images from input satellite images of cities, to be used by local governments for city development or by individuals in their research. Unlike conventional modeling, it has advantages in that it does not need complex inputs or field surveys. In the first semester, we calculated the area of artificial objects from the satellite data of 26 cities that were classified as increasing, decreasing, or no change in population over the decade from 2010 to 2020, and examined the relationship with population. We used correlation analysis and single regression analysis. The correlation for population decline was -0.783, which is a strong negative correlation. This means that as the population decreases, the area of artificial objects increases. We could obtain the same result in the single regression analysis. The correlation for no change was 0.356, which is a weak positive correlation. However, the single regression analysis showed a negative correlation, which resulted in a contradiction. We consider that we need to increase the number of data and conduct more research. We could not obtain the data for the cities with population increasing, because the calculated area data was far off from the measured values of other cities, and we concluded that there was a problem with the data. Therefore, we considered that the data was not accurate because of the problem of data from cities that population increasing, and the contradiction result from analyzing the data from cities that classified as no change in population. Therefore, we had to obtain accurate data and find a way to do so. In the second semester, based on the relationship between population change and city area obtained in the first semester, we selected 70, 108, and 151 cities with increasing, decreasing, or no change in population, and collected these cities data of 2010 and 2020. After that, we modified those data into training data. So, We created a model to generate the predicted future city into image. By inputting the training data created in this image generation model, we output the results of predicting the city 10 years from now. However, some problems occurred in the output result of predicting the city in 2020 from the data of 2010, such as the image is blurred and it is difficult to distinguish between the building and the field. The reason for these are that the color tones of the images differ between the 2010 and 2020 data, and that the download resolution differs between the 2010 and 2020 data. For this reason, reviewing training data, making further adjustments, and improving image quality are issues for the future. Also, we could not achieve the desired results due to the lack of schedule estimation looking back on the activities of the group so far. It was a good opportunity to learn important things in conducting a project, such as thorough schedule management and coordination of work within the group.

(※文責: 野原廣大)

目次

第 1 章	はじめに	1
1.1	背景	1
1.2	目的	1
1.3	従来例	1
1.4	従来の問題点	2
1.5	課題	2
第 2 章	前期の活動	4
2.1	前期プロジェクト学習の概要	4
2.1.1	問題の設定	4
2.1.2	課題の設定	4
2.1.3	到達目標	4
2.1.4	課題の割り当て	4
2.2	課題解決のプロセス	5
2.3	結果	5
2.3.1	成果	5
2.3.2	解決手順と評価	6
2.4	前期の結果を踏まえた今後の課題と展望	7
第 3 章	中間発表	8
3.1	ポスター作成	8
3.2	動画作成	8
3.3	質疑応答	8
3.4	担当者による評価	8
3.5	考察	9
第 4 章	後期の活動	10
4.1	後期プロジェクト学習の概要	10
4.1.1	課題の設定	10
4.1.2	到達目標	10
4.1.3	課題の割り当て	10
4.2	アプローチ	11
4.2.1	データセット	11
4.2.2	教師データ	12
4.2.3	pix2pix	13
4.2.4	サーバ	13
4.2.5	学習上の試行錯誤	14
4.3	後期プロジェクトの結果	15

4.3.1	成果	15
4.3.2	解決手順と評価	16
4.4	今後の課題と展望	16
第5章	成果発表会	18
5.1	ポスター作成	18
5.2	動画作成	18
5.3	質疑応答	18
5.4	担当者による評価	18
5.5	考察	19
第6章	学び	20
6.1	オンラインでの活動(工夫)	20
6.1.1	Zoom	20
6.1.2	Microsoft Teams	20
6.1.3	議事録(共有メモ)(Google Document)	20
6.2	対面での活動	20
6.3	技術の利用・習得	21
6.3.1	衛星画像	21
6.3.2	QGIS	21
6.3.3	Python	21
6.3.4	発表	21
6.3.5	担当教員との連携	22
第7章	まとめ	23
第8章	インターワーキング	24
付録A	新規習得技術	25
付録B	活用した講義	26
参考文献		27

第 1 章 はじめに

1.1 背景

リモートセンシング技術とは、電磁波（X 線、紫外線、赤外線など）によって「もの」に触れずに海や都市などを知る技術であり、天気予報や漁業分野で使われている。国土交通省（2020）によると、日本では東京一極集中や、地方の過疎化が問題となっている。地方創生推進事務局（2021）によると、その解決のために現在、国や地方自治体が、地方創生のための様々な方策を行っている。その結果に対して変動した人口が都市にどのような影響を与えるのかを知ることが重要である。

（※文責: 野原広大）

1.2 目的

都市開発の際に参考として地方自治体などに利用してもらうため、また個人が研究で用いるために、過去 10 年から現在までの人口と都市のデータから将来の都市を予測するモデルの作成を目的とした。前期では、2010 年から 2020 年の 10 年間の衛星データから人工物のみの都市面積を求め、人口との関係性を考察した。後期では、前期で求めた人口変動と都市面積の関係性をもとに、改めて人口変動が増加、減少、変化なしに当てはまる都市を選出し、それらのデータを収集し、未来の都市予測を画像で生成するモデルを作成した。

（※文責: 野原広大）

1.3 従来例

国土技術政策総合研究所では、土地利用と交通の相互作用と調整速度の差異（施策投入から実現までの時間的ラグ）に基づき、ゾーンごとの土地利用や交通の状況を逐次的に扱う集計型の土地利用交通モデルを将来都市構造予測モデルとして用いていた。このモデルの中には土地利用モデルと交通モデルのサブモデルが存在し、さらにそれぞれのモデルのなかにも複数のモデルが存在した。それぞれのモデルに都市計画基礎調査や国勢調査、PT（パーソントリップ）調査等のゾーン別集計値を入力数値とすることで様々な地方自治体での利用を可能としていた [2]。

課題解決に向けた衛星リモートセンシングデータ利用モデル実証プロジェクト [17] では、様々な分野におけるリモートセンシング利用モデルの実証を目指していた。中でも 2020 年度には「宇宙と地上の新しい計測手法による都市計画基本図更新手法の実証」を目指したチームが存在し、都市計画基本図の更新頻度の問題を解決していた。それに伴い、導入前に行っていた空中写真の撮影工程を衛星画像の撮影に置き換えており、これにより天候・気象条件による撮影時間の遅延を解消している。現地調査にも全方位カメラ・ビューアを導入し、負担軽減を実現していた。

（※文責: 池田幸星）

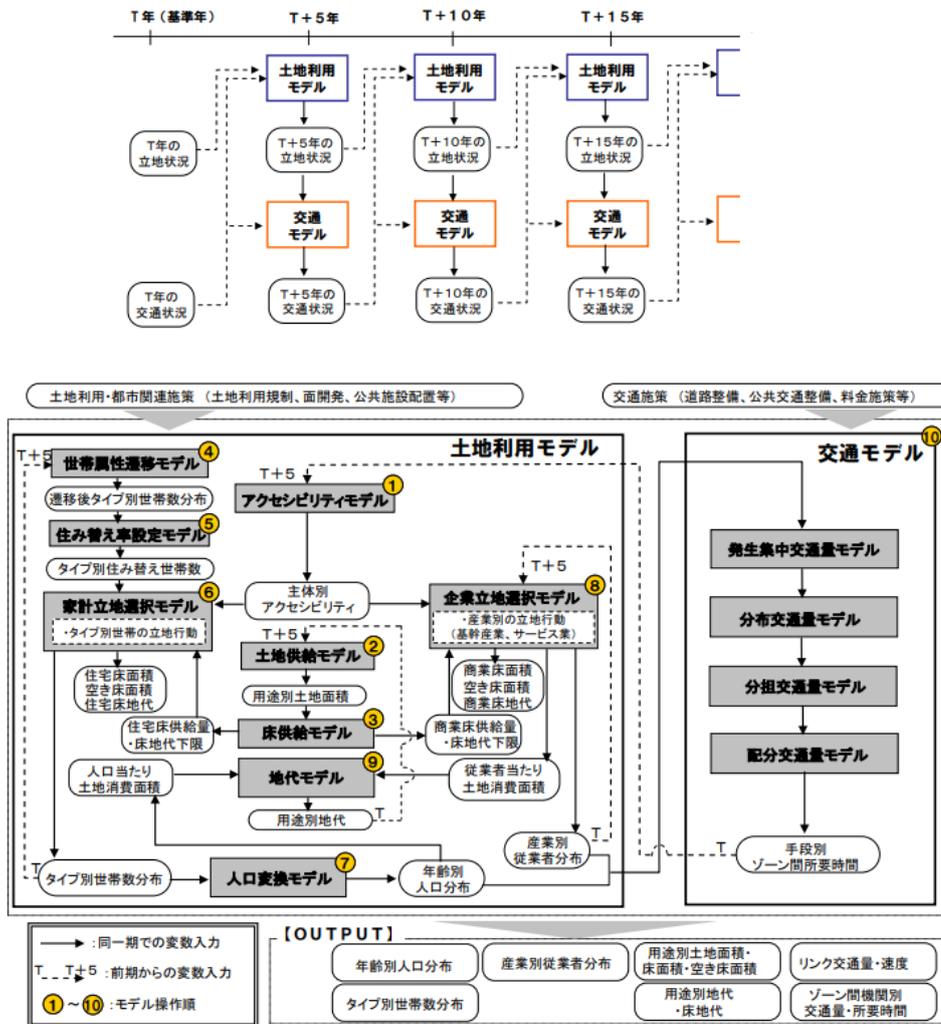


図 1.1 従来例の将来都市構造予測モデル (画像は国土技術政策総合研究所のウェブサイト [2] から引用)

1.4 従来の問題点

国土技術政策総合研究所の手法ではモデルの利用に当たり、人口分布や世帯属性、交通量など必要データの量が多く複雑であり、利用者がデータを収集することが負担であった [2]。また、課題解決に向けた衛星リモートセンシングデータ利用モデル実証プロジェクト [17] の 2020 年度に存在したプロジェクト「宇宙と地上の新しい計測手法による都市計画基本図更新手法の実証」では、衛星画像の高分解能化、撮影頻度の向上、より安価な衛星画像の提供を要望として挙げていた。また、現地調査がいまだに必要な点や民間が気軽に利用できない点が問題だと思われる。

(※文責: 野原廣大)

1.5 課題

前期プロジェクトでは、都市面積は「人工物や建築物があり、人が生活している空間かつ、田んぼや森林を除いた面積」と定義し、リモートセンシング技術を活用して割り出すことを課題とした。そして、そこから人口との関係性について考察することとした。後期プロジェクトでは、pix2pix

Let' s Remote Sensing

という GAN を利用した画像生成アルゴリズムを使用して将来の画像を生成した。人口の変動が現在と過去 10 年の間で、増加、減少、変動なしに分類した都市の 2010 年と 2020 年の画像を教師データとして機械学習を行い、2020 年の画像から 2030 年の予測画像を実際に出力した。

(※文責: 野原広大)

第 2 章 前期の活動

2.1 前期プロジェクト学習の概要

2.1.1 問題の設定

本グループでは、1.4 節で述べたように、あらかじめ利用者がデータを集めなくてはならない点、現地調査が必要な点、民間の利用が容易ではない点が問題として挙げられる。それらの問題の解決を目指す。

(※文責: 野原広大)

2.1.2 課題の設定

2.1 節で述べた問題を、プロジェクトのテーマに従い、私たち学生が自分たちの学びにつながるように、という条件の下で解決することを考えた。その結果、本プロジェクトで学んだ QGIS と、これまで培ってきたプログラミング技術を用いて、26 都市の 2010 年と 2020 年の都市面積を割り出し、人口との関係性を明らかにする。

(※文責: 野原広大)

2.1.3 到達目標

国勢調査 [8] から得た人口データと本グループが算出した都市面積のデータから、将来の都市面積を予測するモデルを作ることを目標とした。

(※文責: 野原広大)

2.1.4 課題の割り当て

2010 年から 2020 年の人口の変化を人口増加、減少、変化なしの傾向に都市を分け、そこからそれぞれ 5~7 都市のデータを取るようメンバーに割り当てた。人口変動の定義は人口の 10% 以上かつ 10000 人以上の増加変動を人口増加とし、人口の 10% 以上かつ 5000 人以上の減少変動を人口減少、人口の ± 1.0% 以内の変動を変化なしとした。なお、人口減少地域の中には被災による大きな変動が含まれる地域も存在したため、それらは除いて考えることとした。割り当ては人口増加を成田龍斗、人口減少を渡部日向、変化なしを池田幸星、野原広大が担当した。

(※文責: 池田幸星)

2.2 課題解決のプロセス

衛星データを提供している EO-browser より Sentinel2、JAXA の ALOS オルソ補正画像プロダクトよりだいち・だいち 2 号から得た都市の衛星データをもとに QGIS を用いて解析を行った。そこで得られたデータを用いて 2010 年と 2020 年の都市面積の変化を求め、人口との比較を行った。人口のデータは国勢調査のデータを用いた。

(※文責: 成田龍斗)

2.3 結果

2.3.1 成果

収集した 26 都市のデータについて、都市面積を算出した。人口の増減率と、都市面積の増減率を変数とした散布図を作成した。不良であるデータを取り除いたものの人口減少は図 2.3 に、変動なしは図 2.4 に示す。相関関係の分析と単回帰分析を行った。単回帰分析については、人口変動の傾向ごとに分類された各データについて、面積増減率を目標変数、人口増減率を説明変数として算出した。人口が減少している都市についての相関は-0.783 と、強い負の相関が表れた。これは、人口が減少すると人工物の面積が増える、ということの意味している。これについて単回帰分析では”面積増減率 (%) = -157.189-9.096*人口増減率 (%)”との結果が得られ、同様に負の相関を得られた。人口の変動が少ない(変化がない)都市においては相関が 0.356 となり、弱い正の相関が表れた。しかし、単回帰分析を行った結果、”面積増減率 (%) = -0.9411 -2.7507*人口増減率 (%)”と、こちらでは負の相関が見られ、矛盾が生じている。人口が増加している都市に関しては、データ収集の際に算出された面積データが他都市の測定値と比較して大きく乖離していて、都市面積として扱うような数値ではないためデータ不良と判断し、相関の算出や単回帰分析を行うことができなかった。

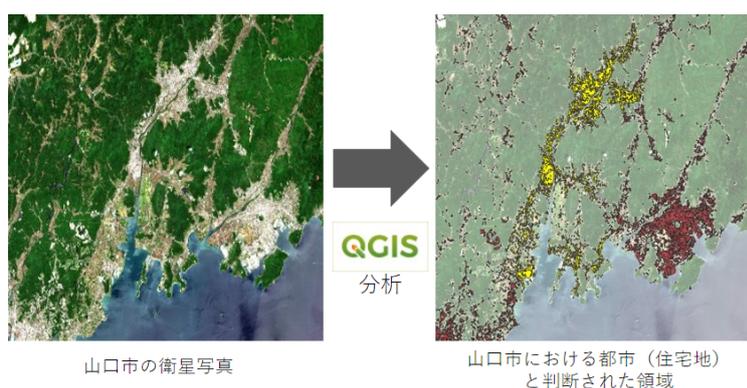


図 2.1 QGIS の分析結果 (衛星画像は EO-browser のウェブサイト [16] から引用)
都市と判定された箇所：赤 山口市であると判定された箇所：黄

(※文責: 渡部日向)

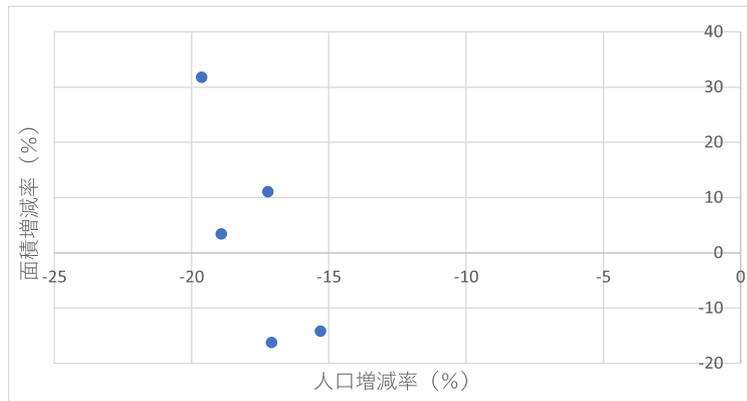


図 2.2 人口減少に分類された都市の人口増減率と面積増減率の関係

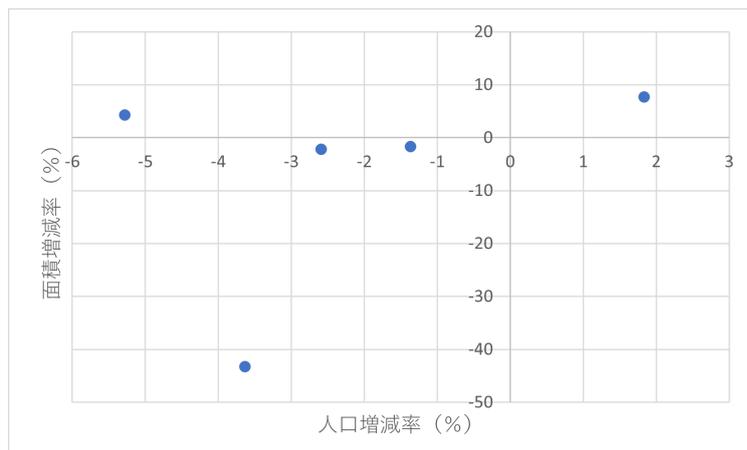


図 2.3 人口変動なしに分類された都市の人口増減率と面積増減率の関係

2.3.2 解決手順と評価

相関係数から、人口と都市面積の変動の関係を把握することができた。ただ、単回帰分析を用いた分析を行った結果、人口減少、変化なし、どちらも両側 5%水準では統計的に有意ではなかった。回帰分析モデルの精度は人口減少で 60%、変化なしで 12%であったため、変化なしの結果の精度は特に低いと言える。このことから、データの正確性やデータのサンプル数、説明変数そのものに問題がある可能性があった。また、人口増加の条件では正しくデータを扱うことができなかった。

人口の変動が少ない都市のデータにおける分析結果について、相関係数では正の相関と出たが、単回帰分析では負の相関となり矛盾が生じているため、さらなる調査、検証が必要であることがわかった。また、人口が増加している都市のデータが不良であったことから、収集したデータの正確性に問題があることも考えられる。このことから、正確なデータ取得やその手法の模索が今後の課題となる。

(※文責: 渡部日向)

2.4 前期の結果を踏まえた今後の課題と展望

今回は、衛星から取得したデータを手作業で QGIS を用いて分析したが、手作業の解析には時間がかかるため、機械学習などを用いて自動的に解析を行っていきたい。また、人口が増加している都市のデータでは、算出した 2010 年の都市面積数値がとて小さくなってしまい正確なデータを取ることが出来なかった。原因として、衛星画像から都市部分を算出する条件式に問題があった可能性が考えられ、利用した条件式ではデータを取得した衛星による衛星画像の差異によって都市の判定ずれが生じてしまったと考察する。これらを踏まえて正確なデータの取得方法を模索していきたい。

(※文責: 成田龍斗)

第 3 章 中間発表

3.1 ポスター作成

中間発表のために、前期の活動内容をまとめたポスターを作成した。ポスターにはリモートセンシングの概要説明と年間スケジュール、それぞれのグループの背景・目的とアプローチ、結果を記した。ポスター作成はプロジェクトグループ全体での作業とし、デザインの統一や画像の挿入、文章の内容などをお互いで確認しながら作業を行った。中間発表時点では主に人口と都市面積の関係性や、実際に QGIS で解析を行った結果などについて記した。

(※文責: 野原広大)

3.2 動画作成

中間発表のための動画制作をポスター作成と並行で行った。ポスターの概要で記載した内容やそれぞれの目的、結果をさらに深く説明したものとなった。PowerPoint にて全部で 38 枚のスライドを作成し、その内のプロジェクト全体の概要説明と年間スケジュール 7 枚をグループ A と B の共同で担当し、グループ B のスライド 7 枚をグループ B で担当した。スライドに合わせてグループの担当者が原稿を読み上げてそれを録音し、全てのスライドで音声による説明を行った。

(※文責: 野原広大)

3.3 質疑応答

中間発表では前半後半でそれぞれ各 3 回の質疑応答の時間が 15 分ほど与えられ、動画やポスターを見た評価担当者や教員から多くの質問を受け付けた。質問には担当者がそれぞれ返答し、質問内容と回答を別の担当者が逐一メモを取り、Zoom の画面共有で常に確認できる状態とした。

(※文責: 成田龍斗)

3.4 担当者による評価

中間発表会の後、Google フォームにて「発表技術」「発表内容」の 2 つの項目を評価担当者の学生や教員に評価と、自由記述によるコメントをいただいた。それぞれの項目は 1 (非常に悪い) から 10 (非常に良い) までの 10 段階評価であった。評価の数は学生と教員を含む 37 名であった。発表技術についての評価の平均は 7.16 (回答数 $n=37$ 、標準偏差 $SD=1.85$) となった。発表内容についての評価の平均は 7.32 (回答数 $n=37$ 、標準偏差 $SD=1.43$) となった。いただいたこれらのコメントをもとに、発表内容の改善や今後の活動を検討した。

(※文責: 成田龍斗)

3.5 考察

最初に、当日答えることが出来なかった二つの質問について回答する。まず「都市だけ抽出する利点はあるのか?」について回答する。これは人の生活圏の面積の変化を知りたいからである。生活圏の面積と人口との関わりから将来の都市がどう変化するかを知りたいと考えたために、都市だけを抽出しようと考えた。次に、「人口減少、都市面積増加で空き家が増加しているのではという考察。これはあき空き家の増加予測に使える?」について回答する。空き家との関係性は分析していないため、そのままのデータで予測する(関連付ける)ことはできないが、実際の空き家の数を調べることで、空き家の増加を予測することができると考える。

次に担当者から頂いた評価について考察する。発表技術に関しては、「質疑応答の際に、今までにあった質問とその回答を提示していた。」や項目に分けて、「質疑応答をしていたように感じたので良いと思いました。」、「主として回答する質疑応答担当者を決めておくことでスムーズに回答に移行できると感じました。」、「質問を一人が答えられる人に振り分けているのは効率的だと感じた。」等の好意的な意見が多くあった。しかし、「説明が一切なく、質問が来なければ特に話すこともなく、無言の時間がとても長かった。」や「概要の説明に関して2,3分説明して再確認を促す時間があるとより分かりやすいと思いました。」の批判的な意見も目立ち、質疑応答の際に、全体を通しての説明の時間を設け、無言の時間を減らし、質問をしやすい環境を作る必要があると考える。また、動画については「スライドは一般の人にもわかりやすいが、説明の声が明瞭でない。」や「動画時間が少し長いと感じたため、もう少し短くしたりパート分けしたりすると良いと感じた。」等の指摘をいただいた。そのため、動画をパート分けし、動画の視聴者が確認しやすいようにする。さらに、動画の声をくっきりさせることでよりよい発表動画になると考える。

発表内容に関しては、「補足説明として難しい用語の説明を行っているのが良いと思いました。」や、「実際のデータが載っているのでもわかりやすいと思った。」、「紹介動画内でプロジェクトの内容に加えて、技術的な説明もされている点が良いと思った。背景がわかりやすく説明されているのでその後に続く説明も理由がしっかりしていて良いと思いました。」等の好意的な意見が多くあった。しかし、目的について「今後の展望をより細かく記載してほしいです。どのようなことをするのか少しあいまいでした。他の点は十分に説明されていて良いものでした。」や「グループBの背景・目的が少しざっくりとしているような気がしたため、もう少し詳細を知りたい。」、「グループBの方の目標が不明瞭な気がします。」、「自治体にどのように利用してもらえるか考えることも必要かと、ちょっと投げやりな様な専門用語が多くて分かりづらいところがありました。」等の指摘があった。このことから、今後の展望、専門用語について詳しく記述し、完成品をどのように利用してもらうかをイメージしやすくする必要があると考える。

(※文責: 成田龍斗)

第 4 章 後期の活動

4.1 後期プロジェクト学習の概要

4.1.1 課題の設定

前期では 10 年間の人口の変動と都市面積の変動の関係を求めたが、その変化を数値のみで表現しても、利用者には伝わりにくいと考えられる。そのため、利用者がどのように変化するか視覚的に理解するために、10 年後の都市を予測した画像を生成するモデルを作成する。また、従来の問題点から利用者が多量のデータを収集することなく手軽に利用できるシステムとして提供する。

(※文責: 渡部日向)

4.1.2 到達目標

標準目標

前期の結果をもとに人口の増加、減少、変化なしの 3 種類に分類した都市規模の変遷を学習し予測するモデルを 3 つ作成する。

(※文責: 成田龍斗)

発展目標

作成したモデルを使用して利用者が手軽に利用できるアプリケーションを作成する。

(※文責: 成田龍斗)

4.1.3 課題の割り当て

GAN を利用した画像生成アルゴリズム「pix2pix」をもとに、Python を用いてモデルの作成を行った。また、衛星データを教師データとして扱うためにデータを加工するプログラムを Python で作成した。これらは渡部日向が担当した。

また、教師データは日本の市区町村を 2010 年から 2020 年の人口の変化を基準に人口増加、減少、変化なしの都市に分類し、それぞれ 70、108、151 都市を選出して各都市 2010 年と 2020 年の衛星データを収集した。2010 年データを野原広大、2020 年データを池田幸星、成田龍斗が担当した。

(※文責: 池田幸星)

4.2 アプローチ

利用者が都市の変化を視覚的に理解できるよう、10年後の都市を予測した画像を生成するモデルを作成した。また、従来の問題点として挙げていたあらかじめ利用者がデータを集めなくてはならない点、現地調査が必要な点、民間の利用が容易ではない点を踏まえて、利用者が多量のデータを収集することなく手軽に利用できるシステムとして提供した。

衛星データを提供している EO-browser より Sentinel2 や、JAXA の ALOS オルソ補正画像プロダクトよりだいち・だいち2号から衛星データを収集した。収集したデータを加工し、学習に使用する都市の形にした。pix2pix を必要に応じて書き換え、学習用の環境に最適化した。変更した pix2pix で、人口の増加、減少、変化なしの3種類に分類した10年間の都市規模の変化を学習し、予測するモデルを3つ作成した。

(※文責: 池田幸星)

4.2.1 データセット

収集方法

教師データに用いるデータセットを2010年と2020年でそれぞれ用意するために、異なるサイトから衛星画像を収集した。2010年の衛星画像は、JAXA が運営する ALOS オルソ利用推進研究プロジェクトのサイトから衛星「だいち」、「だいち2号」の ALOS 補正画像プロダクトを取得した。無償で提供されているため、ユーザー登録を行い、利用規約に同意することで誰でも入手可能である。入手方法だが、画像が欲しい地域の付近にあるピンを選択すると、撮影日時の古い順に並んだ実際の画像とともにダウンロードする画面が開く。そこで欲しい画像のダウンロードを選択するとダウンロードすることができる。2020年の衛星画像は、Sentinel Hub というサイトの EO Browser から衛星「Sentinel2」の衛星画像を収集した。こちらのサイトも会員登録することで誰でも無償でデータを入手可能である。入手方法としては、画面範囲内のデータをダウンロードするため、ウィンドウサイズを全画面にしたうえで対象地域を画面内に収める。今回は画像粒度を統一するため、右下に表示される縮尺が1kmになるよう拡大率を調整した。その後、画面右側ダウンロードアイコンを選択し、Analytical タブを開く。Image format を TIFF(32-bit float)、Image resolution を HIGH、Coordinate system を WGS 84(EPSS 4326) に設定。その後画面下部の Download を選択することでダウンロードすることができる。上記の方法以外でもデータを取得することは可能だが、GeoTiff 形式でダウンロードを行うため、今回の活動では上記の方法で取得を行った。

(※文責: 野原広大)

加工

人口の変動によって都市が分類されるため、その都市のみを抽出する必要があった。QGIS を用いて、国土交通省が公開している「行政区域データ」を市区町村の座標、形状とし、衛星データからその市区町村の切り抜きを行った。2010年の画像は、各波長のモノクロ画像のままで切り抜きを行い、抽出後の画像にそれぞれ赤青緑に着色し、フルカラーに合成した。2020年の画像は、ダウンロード時から、フルカラーであるため、そのまま切り抜きを行った。切り抜きを行った

後、2つの画像を同じ形状に整えた。2010年と2020年の衛星データは、地球上の位置の表現方法が異なった。QGIS上では、地球上の位置の表現方法の違いは自動的に補正されているが、切り抜き後は補正のない画像データとして扱われる。そのため、その補正を行い、2つの画像を同じ形状に整える必要があった。そこでEPSGコードを変更して対応した。EPSGコードとは、EPSG (European Petroleum Survey Group) という団体が規格化した地球上の位置の番号による表現方法 (空間座標系) である。2つの空間座標系は、どちらもWGS84系という測地基準を用いているが、それを2次元平面に表す投影 (地理) 座標系が異なる。2010年の投影座標系は、メルカトル図法を利用したUTM (ユニバーサル横メルカトル) 座標系であり、EPSGコードは、32654である。2020年の座標系は、緯度経度を利用した地理座標系であり、EPSGコードは、4326である。今回は2010年の画像データをEPSG32654からEPSG4326へ変更するための補正を行い、2020年の衛星データの座標系で統一した。次に、2010年と2020年の画像データの解像度 (ピクセル数) を揃えた。2つの画像のピクセルは1対1で対応することが必要だが、ダウンロードした衛星データの解像度が異なるため、形状を揃えるだけでは、対応しない。そのため、画像の大きさ (ピクセル数) を揃える。1ピクセルの情報量を減らさないために小さい方の画像を大きい方の画像のピクセル数まで拡大した。拡大には、品質を高めるためLanczos法を用いてサンプリングした。この加工によって、2010年と2020年の都市の形状と大きさが一致し、各ピクセルが1対1で対応するようになった。10年間の人口の変動が同じ分類をされている都市群で画像の大きさを揃えた。学習する際に、入力する画像の大きさが異なると処理が複雑化するため、入力する画像の大きさを揃えた。全画像の各辺をすべての画像の中で最も長い辺の長さより大きい中で最小の2の累乗の長さまで伸ばし、伸ばした分の領域はすべて黒で埋めた。先ほど変えた大きさや形に手を加えず、不足分を黒で補い、画像を正方形に変形した。上記の過程によって各都市は、同じ形、大きさを揃え、すべての画像は同じ大きさの正方形となるように加工した。次に、2010年と2020年の同じ都市の画像を結合した。学習に使用する画像は2つの画像が左右に結合した状態で入力を行った。2010年の画像を左、2020年の画像を右にして結合した。

(※文責: 渡部日向)

4.2.2 教師データ

ファイルの分類

pix2pixの学習にはtrain, val, testの三種類に分類されたファイルが必要となる。これらの詳細な説明は後の4.2.3 pix2pixにて行う。4.2.1.2にて加工された都市全体が含まれているデータをこれら三種類のファイルに分類した。各ファイルへの配分は82:9:9の割合で行った。

(※文責: 野原廣大)

サイズ変更

結合後画像は、サイズが非常に大きくなり、pix2pixにそのまま入力することができなかった。そのため、256×256ピクセルに画像を分割した。左上から順番に256ピクセル単位で切り抜きを行い、右側の画像からそれと同じ座標の部分を切り抜き、その2つを同様に結合した。生成された画像は、縦256×横512ピクセルの大きさとなった。

(※文責: 成田龍斗)

データの選別

結合後画像は、サイズが非常に大きくなり、pix2pix にそのまま入力することができなかった。そのため、256 × 256 ピクセルに画像を分割した。左上から順番に 256 ピクセル単位で切り抜きを行い、右側の画像からそれと同じ座標の部分を切り抜き、その 2 つを同様に結合した。生成された画像は、縦 256 × 横 512 ピクセルの大きさとなった。

(※文責: 成田龍斗)

4.2.3 pix2pix

画像を生成するモデルには、GAN を利用した画像生成アルゴリズム pix2pix を利用した。学習に必要な画像データ train と、検証用の val、test の 3 つのファイルに分類した。train(train dataset) は学習の基になる画像であり、各分類条件の約 82% の都市がここに該当した。val(validation dataset) は、学習中に学習の進み具合 (精度) を判断する際に用いる検証用のデータとしての画像であり、各分類条件の約 9% の都市がここに該当した。test(test dataset) は、学習後に実際に生成した画像を人の目で比較する際に用いる検証用の画像であり、各分類条件の約 9% の都市がここに該当した。pix2pix の書き換えについてだが、Google Colaboratory 上で公開されている pix2pix のコードのうち、画像を読み込む機能と、画像を出力する機能を書き換えた。画像の読み込み機能は、インターネット上のファイルをダウンロードして読み込むという箇所を、分類後の画像を Google Drive から Zip ファイルでダウンロードするように変更した。画像出力の機能は、実行画面に出力する箇所を画像ファイル (png ファイル) で保存するように変更した。その他の変更点として、学習するシステムは変更を加えず、学習回数を 40000 回から 170000 回に変更し、学習を行った。また、10000 回ごとにモデルの保存と test フォルダのすべての画像から予測画像を出力するように変更を加えた。

(※文責: 渡部日向)

4.2.4 サーバ

初期段階では、Google が無料で公開している Google Colaboratory 上でコードの書き換えや、実行を行っていた。しかし、試行錯誤を続けるうちに、無料で使用できる 10GB 程度の GPU ではメモリ不足が発生し、さらに実行時間に制限があるため、長時間の使用ができない問題が発生した。そこで、担当教授のアドバイスに基づき、本学の GPU 計算サーバを使用することとした。このサーバのクラスタ構成は、ファイルサーバと内部ネットワークで接続した計算ノードが 4 台であった。今回は、メモリが 11GB × 4 枚の GPU を搭載している 2 台の計算ノードを使用した。ファイルサーバへの接続は、リモートデスクトップ接続で行った。また、pix2pix のコードの転送は SCP コマンドを使用した。学習用の画像データは、pix2pix に変更を加えて Google Drive からダウンロードするように設定した。pix2pix の実行のために、TensorFlow や GoogleDriveDownloader などのライブラリのインポートをファイルサーバに行った。pix2pix を実行するシェルスクリプトを作成し、計算スケジューラである Torque を利用して計算ノードで学習を行った。生成された画像は、ファイルサーバから Google Drive へアップロードした。

4.2.5 学習上の試行錯誤

pix2pix の書き換え

初めの構想では、都市の画像は 256 × 256 ピクセルの画像に分割せず、4096 × 4096 ピクセルの都市全体の画像をそのまま学習の入力に使用する計画だった。そのため、pix2pix の学習のシステムを大きな画像に対応させるため、画像の特徴量を畳み込むエンコーダと、逆畳み込みで画像を生成するデコーダのそれぞれの畳み込みの回数や量を増やすコードの書き換えを行った。これを Google Colaboratory で実行した際に、エンコーダで 1024 × 1024 まで畳み込んだ段階で GPU のメモリが不足してしまい、より多くの GPU のメモリが扱える環境が必要になった。

(※文責: 野原廣大)

複数 GPU 環境での実行

担当教員へ GPU のメモリ不足について相談を行った結果、複数 GPU を搭載したサーバが本学にあり、それを利用するようにアドバイスをいただいた。本学のサーバで実行するにあたり、動作環境を整えていた。しかし、ファイルサーバ内に複数の Python2 や Python3 が存在し、ファイルサーバの環境変数がどの Python に繋がっているか、計算ノードでの環境変数でどの Python が実行されているのかが不明であり、それらの究明と問題解決に苦戦した。担当教員の指導の下、計算ノードで実行される Python に TensorFlow などのライブラリをインポートすることに成功し、大きな画像を入力するよう改変した pix2pix のコードの実行を行った。しかし、同じ部分でメモリが不足するエラーが発生した。原因として、TensorFlow の初期の仕様では 1 枚の GPU しか使わないことがわかった。そこで並列化を導入し、複数の GPU を利用するための書き換えを行った。しかし、TensorFlow の並列化は、全ての GPU に同じモデルをコピーし、学習を早める「データ並列化」しか対応していないため、メモリが不足する問題の解決には至らなかった。

(※文責: 成田龍斗)

Mesh TensorFlow の導入

メモリ不足の解決策として GPU を複数用いて一つの巨大なモデルを作成する「モデル並列」を行いたい。しかし、TensorFlow のみでは解決できないため、別のライブラリを用いる必要があった。調査を行った結果、Mesh TensorFlow であればモデル並列に対応していることが分かり、導入の検討を行った。Mesh TensorFlow とは、分散ディープラーニング用のライブラリであり、モデル並列ができることのであった。そのため、畳み込み部分に Mesh TensorFlow を利用し、モデル並列を行うよう書き換えを試みた。ここで問題となったのが、畳み込み部分に用いられている Sequential モデルであり、これを利用することで 1 度の畳み込みで行う 3 つの処理をまとめて定義している。この Sequential モデルが Mesh TensorFlow には存在しない機能であるため、単純な書き換えだけでは動作しないことが判明した。 Mesh TensorFlow の関数で Sequential モデルが行っていた処理を再現するには、記述が長くなる可能性があった。また、一度の畳み込みにおける三つの処理の最後にバッチ正規化を行うのだが、モデル並列を行う上で別々の GPU に分散したデータをどのように正規化すればよいのか、という箇所でも Mesh TensorFlow でどのように実装

するのが課題となった。分散させるための定義を十分に理解できていなかった点も原因のひとつだと思われる。以上のことから、Mesh TensorFlow を導入する上で行うべきことが分かったが、その解決に際して、新たに見えてきた問題点も存在した。

(※文責: 渡部日向)

問題点・懸念点

Mesh TensorFlow を導入する場合、動作の詳細な理解と書き換えに時間がかかり、期日までの完成が見込めない懸念があった。また、Mesh TensorFlow を用いたモデル並列を行う場合には、入力データの形式や畳み込み層の見直しが必要となり、調整が必要な上、畳み込み部のメモリ量の試算からメモリ不足が解消しない可能性もあった。また、メモリ不足が解消し、学習が可能になったとしても教師データの数が少なく、1都市のデータを1つの教師データとして扱う場合には、十分な学習ができない可能性があった。さらに、教師データは対象都市の形状で衛星画像を切り抜いているため、都市形状をGAN(敵対的生成ネットワーク)が学習してしまうことによる出力への影響が懸念された。以上のことから、Mesh TensorFlow の導入は断念し、pix2pix の中身をなるべく改変せずそのまま利用できるよう、教師データをさらに細分化したものを学習に用いる手法を採用した。

(※文責: 渡部日向)

4.3 後期プロジェクトの結果

4.3.1 成果

人口の増加、減少、変化なしの3種類に分類した都市について、2010年と2020年の衛星データをそれぞれ収集して適切な加工を施し、教師データセットを作成した。作成したモデルにそれら3種類の教師データセットを入力し、学習を行うことで3種類のモデルを作成した。

この3種類のモデルは、都市画像を入力することで実際に10年後の変化を予測した画像を生成することができた(図4.1)。図の左側にある「Input Image」は入力する画像であり、この画像をジェネレータにかけて画像を生成する。図の右側にある「Predicted Image」は、生成された予測画像である。図の真ん中にある「Ground Truth」は、実際の10年後の画像であり、出力結果として期待される画像である。

出力画像の精度を評価するために、L1_loss について考える。L1_loss とは、正解画像と生成された画像の差の絶対値の平均である。グラフの4万回から8万回の学習結果ではL1_loss が小さい値をとる。よって当該箇所での出力はtarget画像により近いものであるといえる。

本プロジェクトでは、従来例のモデルがあらかじめ利用者がデータを集めなくてはならない点、現地調査が必要な点、民間の利用が容易でない点を取り上げた。この成果では、利用者が集めなければいけないデータが都市面積のみであること、衛星データのみでの予測なので現地調査が不要なこと、無償で公開されている衛星データを利用しているので容易に利用できることといえる。

以上のことから標準目標は達成できたといえる。しかし、発展目標として定めていたアプリケーションとしての完成形にはたどり着くことができなかった。

(※文責: 成田龍斗)



図 4.1 2010 年から 2020 年の「変化なし」の出力結果 (学習 70000 回)

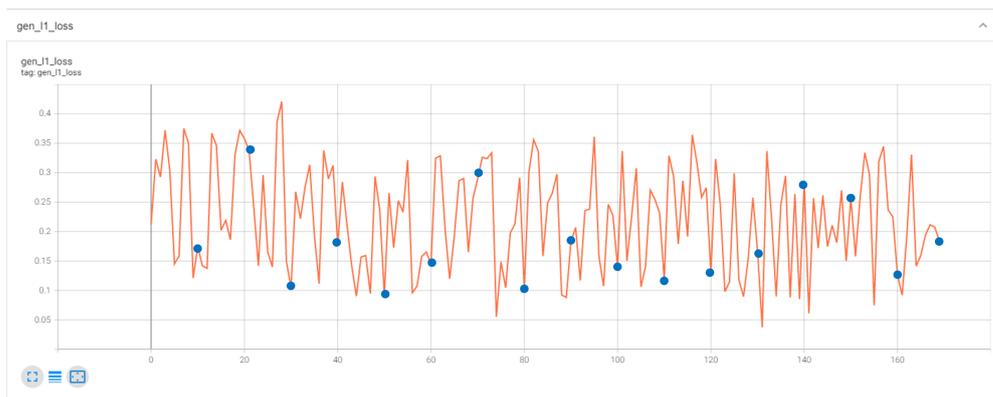


図 4.2 L1.loss のグラフ 縦軸は L1Loss、横軸は学習回数を表す (単位は千回) 青点は画像が出力される 1 万回ごとの学習の区切り

4.3.2 解決手順と評価

学習が 6~8 万回程度で最も良い結果が得られた。また、学習が 9 万回を超えるとノイズが激しくなり精度が低下した。出力全体を通して「建物がぼやけて都市部の境界の判別が難しい」「田畑が山のように出力されてしまい人の活動圏が曖昧になってしまった」「建物と田畑の判別が難しく、混ざって出力されているように見える」といった問題が結果に表れた。原因は、2010 年と 2020 年でダウンロード時の解像度が異なることが考えられる。同様に 2020 年のデータを入力し 2030 年の予測を行ったが、都市の画像としては良い結果が得られなかった (図 4.2)。原因としては、用いたデータが 2010 年と 2020 年では色調が異なることから、2020 年のデータを入力するとその色を入力するパターンを学習していないため、正しく動作していないと考えられる。

(※文責: 池田幸星)

4.4 今後の課題と展望

今後の課題としては、教師データを見直して、学習に悪影響を与える可能性のある画像を取り除き精度を向上させること、2010 年と 2020 年の画像との彩度の差を調整して汎用性を向上させること、取得する衛星データの画質を向上させることが挙げられる。

また、当初の目的である自治体や一般の方々などに利用してもらうという点において、教師データの取得に難航したことやスケジュール見積りの甘さから、サービスの提供方法や、UI の設計と

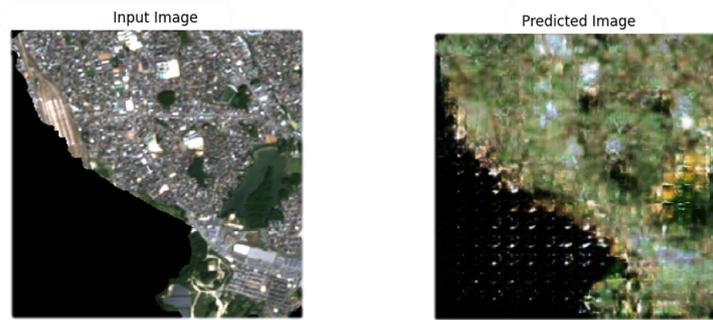


図 4.3 2020 年から 2030 年の出力結果、右が予測画像 (学習 70000 回)

いった面に着手することが厳しい状況であった。スケジュール管理や作業時間の調整など、まだまだ改善の余地があったと思われる。

(※文責: 成田龍斗)

第 5 章 成果発表会

5.1 ポスター作成

成果発表会のためにポスター作成を行った。前期の中間発表のときに用いたポスターをもとに、内容の修正や変更、画像などを変更した。結果の記載方法を簡条書きで簡潔にすることで、ポスターを見た人に容易に伝わるようにした。また、それぞれの項目の文章量に合わせて、全体的なポスターのデザインにも変更を加えた。

(※文責: 野原廣大)

5.2 動画作成

成果発表会のための動画作成をポスター作成と並行で行った。前期と同様に、ポスターで記載されている背景や結果の項目を詳細に説明した。動画作成の手法は前期と同様に PowerPoint を用いて行った。スライド作成は前期と同様にグループ A と B で分かれて行い、全体で 54 枚のうち、グループ B は 14 枚作成した。また、概要説明などの全体の内容についてのスライド 9 枚は共同で作成を行った。

(※文責: 野原廣大)

5.3 質疑応答

成果発表会では前半後半でそれぞれ各 3 回の質疑応答の時間が 15 分ほど与えられ、動画やポスターを見た評価担当者や教員から多くの質問をいただいた。質問には担当者がそれぞれ返答し、質問内容と回答を別の担当者が逐一メモを取り、Zoom の画面共有で常に確認できる状態とした。

(※文責: 成田龍斗)

5.4 担当者による評価

成果発表会の後、Google フォームにて「発表技術」「発表内容」の 2 つの項目を評価担当者の学生や教員に評価と自由記述によるコメントをいただいた。それぞれの項目は 1 (非常に悪い) から 10 (非常に良い) までの 10 段階評価であった。評価の数は学生と教員を含む 35 名であった。発表技術についての評価の平均は 7.03 (回答数 $n=35$ 、標準偏差 $SD=1.84$) となった。発表内容についての評価の平均は 7.86 (回答数 $n=35$ 、標準偏差 $SD=1.22$) となった。いただいたこれらのコメントをもとに、今後の活動を検討した。

5.5 考察

最初に、当日答えることが出来なかった「具体的にどこかの都市の変化が分かったのか？」という質問に対して回答する。実際に未来の都市の予測画像を生成することは成功した。しかし、画像が精度よく出力されていないため、どのように都市が変化したか言語化するには多少の考察が必要であり、変化が具体的に分かったとは言えない結果となった。

次に担当者から頂いた評価について考察していく。発表技術に関しては、「今までにきた質問とその回答を都度書き出して表示（画面共有）していたのが良かった。」という意見があった一方で、「発表の準備が少し足りていなかったかと思います。事前に、プロジェクトの概要説明の準備をしていればよりよい発表になるかと思います。」や「質疑応答時間には初めに手短でいいので説明が欲しかった。」という意見があったため、質疑応答の際には5分ほど使って全体の説明をする必要があると考える。動画スライドに関しては、「動画のスライドはクオリティが高く、とても聞きやすいものだった。」という意見があった一方で、「スライドの作りが工程順に見えたので、最初に最終成果物を見せてしまってイメージを持ってもらう流れでも良かったかも知れませんね。」という意見があったため、最初に最終成果物を見せ、イメージを持ちやすくする必要があると考える。ポスターに関しては、「ポスターに関してプロジェクトメンバの名前が白文字で背景の地球の雲の白色と同化して見づらいと感じました。」や「ポスターが Let's というイメージに対して背景が暗い。都市の配色がすこし汚いと感じた。」、「ポスターの配色についてコントラストが強すぎて見づらさを感じました。また、文字ばかりで読みづらさを感じました。」という批判的意見が多くあった。そのため、ポスターを見る人の立場になり、文字の見やすさや配色についてもっと吟味する必要があると考える。

発表内容に関しては、「10年後の都市の変化が予測できるのが、おもしろいと思った。」や「内容自体は今後発展していけば役に立ちそうだった。」という好意的な意見が多かった。しかし、目的に関しては、「目的もおもしろかったし、やっている内容も独特で面白いものだと感じた。」という意見があった一方で、「動画を見ても、それぞれの想定する利用者と利用方法、利用するとどのようなことが得られるのかが分かりづらい。」という意見もあったため、だれが見てもわかりやすい発表が出来たとは言えないと考える。改善点として、目的や利用方法についてもっと詳しく説明する必要があると考える。

第6章 学び

6.1 オンラインでの活動 (工夫)

6.1.1 Zoom

本グループでは、1.4節で述べたように、あらかじめ利用者がデータを集めなくてはならない点、現地調査が必要な点、民間の利用が容易ではない点が問題として挙げられる。それらの問題の解決を目指す。

(※文責: 池田幸星)

6.1.2 Microsoft Teams

2.1節で述べた問題を、プロジェクトのテーマに従い、私たち学生が自分たちの学びにつながるように、という条件の下で解決することを考えた。その結果、本プロジェクトで学んだ QGIS と、これまで培ってきたプログラミング技術を用いて、26都市の2010年と2020年の都市面積を割り出し、人口との関係性を明らかにする。

(※文責: 池田幸星)

6.1.3 議事録 (共有メモ)(Google Document)

国勢調査から得た人口データと本グループが算出した都市面積のデータから、将来の都市面積を予測するモデルを作ることを目標とした。

(※文責: 池田幸星)

6.2 対面での活動

2010年から2020年の人口の変化を人口増加、減少、変化なしの傾向に都市を分け、そこからそれぞれ5~7都市のデータを取るようメンバーに割り当てた。人口変動の定義は人口の10%以上かつ10000人以上の増加変動を人口増加とし、人口の10%以上かつ5000人以上の減少変動を人口減少、人口の±1.0%以内の変動を変化なしとした。なお、人口減少地域の中には被災による大きな変動が含まれる地域も存在したため、それらは除いて考えることとした。割り当ては人口増加を成田龍斗、人口減少を渡部日向、変化なしを池田幸星、野原広大が担当した。

(※文責: 池田幸星)

6.3 技術の利用・習得

6.3.1 衛星画像

衛星は基本的に地表面から放出される電磁波の波長を計測してデータとして扱うことを学んだ。衛星データは直接利用することは難しい。単純なトゥルーカラー画像として利用するだけでも赤色、緑色、青色に当たる波長データを合成することでようやく入手することができた。私たちは物体と波長の関係性を調べつつ実際に操作することで、波長を組み合わせ四則計算することによる植生分布や土壌、都市化指数など、様々なデータを求めることが可能であることを学び、衛星データの性質や取り扱い方について理解することができた。

(※文責: 池田幸星)

6.3.2 QGIS

プロジェクト当初は、誰も使用したことのないツールであったため手探りの状態であったが、グループ全員で使い方を模索しつつ共有することで、すべての機能とまではいかないが、最終的に本プロジェクトで必要となる操作に関しては一通り使いこなすことができた。具体的には波長ごとの画像合成、色付け、衛星画像の合成、トリミング操作についてである。

(※文責: 池田幸星)

6.3.3 Python

教師データを作成する過程で、衛星データを加工するプログラムを Python で記述した。また、学習プログラムの pix2pix が Python で記述されているため、こちらも Python を用いて書き換えを行った。書き換えを行う際には、機械学習 (ディープラーニング) や TensorFlow, Mesh TensorFlow についての具体的な動作を理解する必要があるため、それらについての学習を行った。全体を通して作成したプログラムはすべて画像を扱うものであったため、画像を扱うライブラリである OpenCV2 と PIL についても学習を行った。

(※文責: 池田幸星)

6.3.4 発表

ポスター

プロジェクト紹介用のポスター作成では文字量を減らすことを意識して、見やすく、読みやすいポスターを目標としたが、可能な限り削減を行ったが、見やすいポスターとしては不十分であった。その結果、情報量が多くなり認知負荷が高いと思われるデザインとなってしまったことは反省点である。また、内容記入もポスターデザインも各メンバーの裁量で同時並行していたため、客観で確認することが難しかったことも反省点として挙げられる。このことから、ポスターの作成自体を少数精鋭で行う、もしくは監修する役割を設けることでポスター全体を客観から俯瞰して適宜確認する重要性を学んだ。

(※文責: 池田幸星)

動画 (スライド)

発表資料は前期発表、最終発表ともに動画形式をとった。発表動画作成にあたり活動内容を効果的に伝えるスライドを作成した。スライド作成では一目でわかる簡潔さを求められるため、ポスター同様、文量の削減やデザインの重要性を再確認した。また、資料に利用する素材の著作権管理には苦戦し、適宜著作権の確認を怠らないことが必要であると学んだ。

動画作成ではスライドだけではなく発表音声の収録も重要となった。収録する際には聞き取りやすさを意識し、声の大きさや進行速度に気を配った。また、原稿文章の校正も念入りに行い、伝わりやすく、かつ不自然のない日本語となるよう注力した。

(※文責: 池田幸星)

質疑応答

前期発表、最終発表における質疑応答では、私たちの間で生まれなかった貴重な意見や視点を持つ疑問を多くいただいた。これにより、客観的な視点からプロジェクト成果を見直す良い機会となった。また、発表技術への意見もいただき、改善すべき点を認識することができた。このように、私たち自ら最善を尽くしても他者からの評価では指摘すべき点多々あるということが分かった。よって、より広い視点を持ち、さらなる改善に努めたいと考える。また、発表では他プロジェクトへの評価も良い機会となり、発表手法やプロジェクト進行のノウハウなど、我々にも通じる学びがある有意義な時間となった。

(※文責: 池田幸星)

6.3.5 担当教員との連携

プロジェクトでは各回において開始・終了の二度、メンバー全員と担当教員二名が集まり作業内容と進捗の報告を行った。その際に担当教員からプロジェクトに関するアドバイスや提案をいただいた。また、学習用プログラム作成の際には、発生する課題の解決策提案にご助力いただき、プログラムの実行に用いるサーバの扱い方をご教授いただいた。

(※文責: 池田幸星)

第7章 まとめ

本プロジェクトではリモートセンシングを用いて、環境や情勢に適したシステムを作り、利用者の助けとなることを目的とした。情勢に着目すると、日本の大きな問題の中には東京一極集中による地方の過疎化が挙げられる。そのような中、政府や各自治体は地方創世のための方策を常日頃行っている。そこで、本グループはリモートセンシングを用いた都市の分析をテーマにプロジェクトを進行し、地方創世の方策による影響を調べるとともに都市開発の助けや自治体住民の興味を引くようなシステムとして、衛星画像から将来の都市を予測するモデルを作成することを最終目標とした。前期の活動としては、2010年から2020年間の人口の変動に増加、減少、変化なしの特徴がある市町村を26都市選び、それら年のデータを収集した。リモートセンシング技術を用いて各市町村のエリアから田畑や森林を除く人工物のある地域を特定し、それらを都市面積として割り出した。そして最終的にこれらの面積と各市町村の人口を照らし合わせ、関係性を求めることを目標とした。従来例との差別化として、複雑な入力や現地調査を必要とせず、手軽に利用できることを意識した。結果としては、衛星画像からQGISを用いて田畑や森林を除いた都市面積を算出することができた。都市面積と人口変動の関係性について相関係数を用いて算出すると、人口が減少している都市では強い負の相関(-0.783)、人口の変動があまりない都市では弱い正の相関(0.356)があることがわかった。人口が増加している都市はデータ不良で算出不可であった。求めることができなかった人口増加データに関してはもちろんのこと、求めることができた二つのデータに関しても、取得するデータの正確性やデータの母数が不足していたことが課題として挙げられた。これら前期の内容を踏まえて、後期では前期同様2010年から2020年の人口の変動に従い増加、減少、変化なしの三種類に分類された市区町村をそれぞれ70、108、151都市に分類し、2010年と2020年のデータを収集した。それらから都市規模の変化を学習することで将来の都市を予測するモデルを作成することを目標とした。結果としては、教師データを学習し都市画像を出力することができた。しかし、出力全体を通して建物がぼやけて山間部と都市部の境界が判別できないことや、田畑が山と同様に出力されてしまい、人の活動圏を正確に捉えられないことなどが問題として挙げられた。また、学習回数が増えると出力画像のノイズが激しくなることや、2010年と2020年のデータ収集に用いた衛星が異なるため色調に差異が生まれ、期待した結果が得られない事態も発生した。このことから、教師データの見直しと精度向上や色調の調整、衛星データ自体の画質向上などが今後の課題として挙げられる。活動全体を通して、教師データ収集の難航やスケジュール見積り甘さにより、当初の目的である自治体や一般の方々などに利用してもらうという目標が達成できず、サービスの提供やUI設計などに着手することができなかった。スケジュール管理や作業時間の調整について改善の余地があり、プロジェクトを行う上で良い学びの機会となった。

(※文責: 成田龍斗)

第 8 章 インターワーキング

前期では、2010 年から 2020 年にかけて人口変化が± 1.0% 以内である 5 都市の面積を衛星画像から算出した。その際、都市を割り出すフィルターの数値や計算式について、他のグループメンバーと相談して決定した。しかし、他メンバーとの認識の違いから、正しく面積を取れず、グループ内での認識のすり合わせの重要性を学んだ。後期では、機械学習に用いる教師データの収集に尽力し、2010 年の都市のデータを主に担った。また、pix2pix の運用や変更、サーバの導入などについてグループメンバーと方向性を共有し、前期のような失敗を繰り返さないよう意識した。ポスターや発表資料制作の際には、都市グループ内のみでなく双方のグループメンバーから意見を取り入れ、デザインや見出し、内容についてまとめた。発表で用いた動画については、デザインや演出を監督し、グループ間で違いが出ないように統一し、見やすいものを意識して制作した。

(※文責: 野原廣大)

2010 年から 2020 年にかけて人口が減少した 5 都市、群馬県下仁田町、高知県仁淀川町、高知県馬路村、神奈川県箱根町、徳島県つるぎ町の面積を算出した。都市の面積を計算するための計算式やその処理について方策を求めた。課題について中心をなして決定した。全員のデータを集約してその関係性を求めた。中間発表用のスライドについてその内容を決定した。機械学習へのアプローチを担当した。pix2pix の改変、MeshTensorFlow の導入の検討、サーバの環境構築や、学習の実行や管理を行った。教師データ作成のための、加工を行うプログラムの作成を行った。衛星データの取得方法の検討に貢献した。成果発表会への、スライドやポスター作成の原稿や、想定質問への回答などを作成した。

(※文責: 渡部日向)

2010 年から 2020 年にかけて人口が増加した福岡県新宮町などの衛星写真を Sentinel-2 とだいちを用いて取得した。取得したデータを QGIS で解析した。その際、Sentinel-2 とだいちから取得したそれぞれのデータの都市と植生を区別する 2 つの計算式をグループメンバーと相談して決めた。しかし、同じ式を用いたにも関わらず 2010 年の愛知県長久手市の都市面積が 31ha ととても狭くなってしまいデータ不良となった。また、グループ B 発表動画の音声を入れた。後期では 2010 年と 2020 年の人口減少、増加、変動なしの都市の衛星データをだいちと Sentinel-2 を用いて収集した。また、その収集したデータを市区町村ごとの形に切り抜いた。また、成果発表会の動画に音声付けを行った。

(※文責: 成田龍斗)

前期活動では、2010 年から 2020 年にかけて人口変化が± 1.0% 以内である都市のうち、大津市、盛岡市、福島市、水戸市、福井市、山形市、佐賀市、山口市の 8 都市の面積を 2020 年、2010 年それぞれにおいて衛星画像から算出した。また、中間発表用資料の作成に尽力した。後期活動では、2020 年の衛星データを共同で収集し、都市形状に切り抜く作業を中心に言い、手が空き次第別作業の手伝いに回った。また、前期同様最終発表用資料の作成を率先して行った。

(※文責: 池田幸星)

付録 A 新規習得技術

Sentinel-2 とだいち、だいち 2 号より衛星画像を収集する技術。QGIS を用いてその収集したデータを解析する技術。機械学習についての知識。学習プログラムを作成する際に用いた Python 技術。Python にて画像を扱う際に用いた画像処理ライブラリについての技術。実行環境として用いたサーバ技術。

(※文責: 成田龍斗)

付録 B 活用した講義

認知システム論で学習した R の使い方で、人口と都市面積の増減率のデータより単回帰分析を行った。その結果、それらの関係を 1 次関数で表すことができた。

(※文責: 渡部日向)

参考文献

- [1] 宇宙航空研究開発機構 地球観測研究センター ALOS 利用研究プロジェクト. 「ALOS オルソ補正画像プロダクト (ALOS-ORI)」. 2018. <https://www.eorc.jaxa.jp/ALOS/alos-ori/index.html>. (最終閲覧日:2021年7月14日)
- [2] 国土技術政策総合研究所. 「国土技術政策総合研究所版都市の将来像アセスメントツール技術資料」. 2013. <http://www.nilim.go.jp/lab/jdg/doc/assessment-tdocj-v1.0.pdf>. (最終閲覧日:2021年7月14日)
- [3] 国土交通省. 「国土交通白書 2020 第1節 我が国を取り巻く環境変化 ■ 2 東京一極集中と地方への影響」. 2020. <https://www.mlit.go.jp/hakusyo/mlit/r01/hakusho/r02/html/n1112000.html>. (最終閲覧日:2021年7月14日)
- [4] 小口高, 奥貫圭一, 河端瑞貴, 齋藤仁, 瀬戸寿一, 高橋信人, 谷謙二, 西村雄一郎, 山田育穂, 米澤千夏, 村山祐司, 近藤康久, 早川裕弐, 古橋大地, 湯田ミノリ, 貞広幸雄, 黒木貴一, 高阪宏行, 佐藤英人, 中山大地, 岡部篤行, 石川徹, 田中靖, 松山洋, 矢野桂司, 若林芳樹, 久保田光一, 有川正俊, 太田守重, 藤田秀之. 「GIS 実習オープン教材」. 2018. <https://gis-oer.github.io/gitbook/book/>. (最終閲覧日:2021年7月14日)
- [5] 総務省. 「政府統計の総合窓口 (e-Stat)」, 国勢調査 (総務省), 都道府県・市区町村別統計表 (国勢調査). 2017. <https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200521&tstat=000001049104&cycle=0&tclass1=000001049105&tclass2val=0>. (最終閲覧日:2021年7月14日)
- [6] 総務省. 「政府統計の総合窓口 (e-Stat)」, 社会・人口統計体系 (総務省), 統計でみる市区町村のすがた 2020. <https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200502&tstat=000001141146&cycle=0&year=20200&month=0&tclass1=000001141147&tclass2val=0>. (最終閲覧日:2021年7月14日)
- [7] 総務省. 「政府統計の総合窓口 (e-Stat)」, 社会・人口統計体系 (総務省), 統計でみる市区町村のすがた 2020. <https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200502&tstat=000001141146&cycle=0&year=20200&month=0&tclass1=000001141147&tclass2val=0>. (最終閲覧日:2021年7月14日)
- [8] 総務省. 「住民基本台帳に基づく人口、人口動態及び世帯数」. 2021. https://www.soumu.go.jp/main_sosiki/jichi_gyousei/daityo/jinkou_jinkoudoutai-setaisuu.html. (最終閲覧日:2021年7月14日)
- [9] 総務省統計局. 統計学習の指導のために先生向け. 2020. <https://www.stat.go.jp/teacher/glossary-sa-so.html>. (最終閲覧日:2021年7月14日)
- [10] 宙畑編集部. 「課題に応じて変幻自在? 衛星データをブレンドして見えるモノ・コト #マンガでわかる衛星データ」. 2019. <https://sorabatake.jp/5192/>. (最終閲覧日:2021年7月14日)
- [11] はとば. 「【実習編】 ~Landsat8 衛星画像で植生を映えさせた地図を描こう. 2020.

https://note.com/kinari_iro/n/nfc6ee22bcd11. (最終閲覧日:2021年7月14日)

- [12] 内閣府地方創生推進事務局. 地方創生. 2021. <https://www.chisou.go.jp/sousei/>. (最終閲覧日:2021年7月14日)
- [13] リモートセンシング技術センター. Sentinel-2A / 2B / 2C / 2D. 2020. <https://www.restec.or.jp/satellite/sentinel-2-a-2-b>. (最終閲覧日:2021年7月21日)
- [14] Esri ジャパン. 「ついにキタ! 日本の衛星だいち (ALOS) の衛星画像が無償でダウンロード可能に!」. 2018. <https://blog.esrij.com/2018/04/20/post-29871/>. (最終閲覧日:2021年7月14日)
- [15] nigo1973. 「無料で最新の衛星画像を入手する方法.」. 2019. <https://qiita.com/nigo1973/items/9bb6a11caac8e3e1e850>. (最終閲覧日:2021年7月14日)
- [16] Sentinel Hub. 「EO-browser」. 2018. <https://apps.sentinel-hub.com/eo-browser/>. (最終閲覧日:2021年7月14日)
- [17] s-net. 「令和3年度課題解決に向けた衛星リモートセンシングデータ利用モデル実証プロジェクト」.2021. <http://www.uchuriyo.space/model/>. (最終閲覧日:2021年7月14日)
- [18] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros. 「Image-to-Image Translation with Conditional Adversarial Networks」.

・衛星画像の提供について

Contains modified Copernicus Sentinel data 2020 and 2021 processed by Sentinel Hub
本プロジェクトに使用したデータは、JAXA の ALOS オルソ補正画像プロダクトより提供を受けました。

本プロジェクトに使用したデータは、JAXA の GCOM-C プロダクトより提供を受けました。

(※文責: 渡部日向)